CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 908

January 2009

International Trade Integration: A Disaggregated Approach

Natalie Chen and Dennis Novy

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C
ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract

This paper investigates the sources and size of trade barriers at the industry level. We derive a micro-founded measure of industry-specific bilateral trade integration that has an in-built control for time-varying multilateral resistance. This trade integration measure is consistent with a broad range of recent trade models including the Anderson and van Wincoop (2003) framework, the Ricardian model by Eaton and Kortum (2002) and heterogeneous firms models. We use it to explore trade barriers for manufacturing industries in European Union countries between 1999 and 2003. We find a large degree of trade cost heterogeneity across industries. The most important trade barriers are transportation costs and policy factors such as Technical Barriers to Trade. Trade integration is generally lower for countries that opted out of the Euro or did not abolish border controls in accordance with the Schengen Agreement. Reductions in trade barriers explain about one-half of the growth in trade over the period 1999-2003 and are therefore a major driving force of the EU Single Market.

JEL Classifications: F10, F15
Keywords: Trade Integration, Gravity, Trade Costs, Multilateral Resistance, Industries, Disaggregation, European Union

This paper was produced as part of the Centre's Globalisation Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements

We are grateful for comments by Richard Baldwin, Pierre-Philippe Combes, Donald Davis, Ronald Findlay, James Harrigan, Thierry Mayer, Steve Redding, David Weinstein and participants at the CEPR ERWIT 2007, Midwest Spring Meetings 2008, the Product Market Integration in the EU 2008 workshop at the Universidade do Minho, Columbia, Fordham, Georgia Tech, HEI Geneva, Nottingham, Oxford, Santa Clara, UC Davis, UC Santa Cruz, Utrecht School of Economics, Vanderbilt and Yale.

Natalie Chen is a Visitor with the Globalisation Programme at the Centre for Economic Performance, London School of Economics and an Associate Professor of Economics at the University of Warwick. Dennis Novy is an Assistant Professor of Economics at the University of Warwick.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

ISBN 978-0-85328-341-6

1 Introduction

Trade costs are a staple ingredient in today's trade literature. Broadly defined, trade costs include any cost of engaging in international trade such as transportation costs, tariffs, non-tariff barriers, informational costs, time costs, different product standards, exchange rate costs and local distribution costs, among others. They feature prominently in the vast majority of theoretical papers, including the models of Eaton and Kortum (2002) and Melitz (2003). In addition, a growing empirical literature, surveyed by Anderson and van Wincoop (2004), is devoted to exploring the sources and size of trade costs. A deeper understanding of the causes of trade costs is of particular importance because it would enable a better evaluation of their welfare implications. These are suspected to be large: on their own, policy-related trade costs may be worth more than ten percent of national income (Anderson and van Wincoop, 2002).

A major challenge faced by empirical researchers is to measure overall trade costs since "direct measures are remarkably sparse and inaccurate" (Anderson and van Wincoop, 2004, p.692). Direct measures are only available for a few components, for instance transportation and insurance costs, usually proxied by the ratio of *c.i.f.* and *f.o.b.* trade values (Harrigan, 1993, Hummels, 2001a, 2007),[1] policy barriers such as specific tariff or non-tariff barriers (Chen, 2004, Harrigan, 1993, Head and Mayer, 2000), informational costs (Rauch, 1999) or time costs (Evans and Harrigan, 2005, Harrigan, 2005, Hummels, 2001b). But even for those components, data coverage is often limited to a few countries and years, and it can be hard to gather disaggregated trade cost data at the industry or product level.

Given those difficulties in obtaining accurate measures of trade costs, some researchers indirectly infer the level of trade impediments from trade flows. This approach has the obvious advantage of extending the analysis to more countries, years and more finely disaggregated data. One way of doing this is to compute trade to output ratios (Harrigan, 1996); another is to estimate "border effects," which mostly reflect the extent of border-related costs.[2]

This paper is part of the research effort that attempts to indirectly infer trade impediments from trade flows. The first contribution of the paper is to develop a micro-founded measure of bilateral trade integration that can be applied to disaggregated panel data. We derive this measure by modeling disaggregated trade flows at the industry level in the gravity framework pioneered by Anderson and van Wincoop (2003, 2004), allowing trade costs and substitution elasticities to be heterogeneous across industries. Anderson and van Wincoop (2003) show that trade flows are determined not only by bilateral trade costs between two countries but also by average trade barriers with other countries, which they refer to as "multilateral resistance." Following the approach of Novy (2008), we derive an analytical solution for multilateral resistance variables that vary across industries and over time. In turn, this enables us to derive a micro-founded measure of bilateral trade integration that has an in-built control for time-varying multilateral resistance and can therefore be applied to panel data.

---

[1] Moreover, Limão and Venables (2001) use the quotes from shipping firms for a standard container shipped from Baltimore to several destinations. Combes and Lafourcade (2005) develop a new methodology to compute transportation costs and apply it to road transport by truck in France.

[2] Examples include Anderson and van Wincoop (2003), Baldwin, Forslid, Martin, Ottaviano and Robert-Nicoud (2003), Chen (2004), Eaton and Kortum (2002), Evans (2003), Head and Mayer (2000), Head and Ries (2001), McCallum (1995), Nitsch (2000) and Wei (1996).

Arguably, the Anderson and van Wincoop model is one of the most parsimonious trade models of recent years. It rests on the Armington assumption that countries produce differentiated goods and trade is driven by consumers' love of variety, leading to the key gravity equation. However, we show that isomorphic trade integration measures can be derived from the Ricardian trade model by Eaton and Kortum (2002), from Chaney's (2008) extension of the Melitz (2003) heterogeneous firms model as well as from the heterogeneous firms model by Melitz and Ottaviano (2008) with linear non-CES demand. This is possible because all these models lead to gravity equations that have a similar structure. Our trade integration measure is therefore consistent with a broad range of the recent theoretical trade literature and is not limited to any one particular framework.

Although we regard our approach as complementary to other research that indirectly infers trade impediments from trade flows, we discuss how our trade integration measure differs from ad hoc proxies commonly used in the literature such as trade to output ratios. In particular, we show that ad hoc proxies without micro-foundations can erroneously pick up changes in *multilateral* trade barriers as changes in *bilateral* trade barriers. We are instead able to focus on bilateral trade integration because our micro-founded measure is not distorted by multilateral resistance effects.

The second contribution of the paper is to bring our measure of trade integration to the data. This enables us to document and explain the variation of trade barriers across 166 manufacturing industries in 11 European Union countries over the period 1999-2003. The case of the European Union is appealing since trade integration is expected to be strong between its member states due to two factors. First, these countries have succeeded in dismantling many restrictions on trade, including tariffs and quotas that were completely eliminated by 1968. Second, the situation has been further reinforced by the implementation of the Single Market Programme (SMP), launched in the mid-1980s.

Consistent with the standard gravity literature, the variation of trade integration across country pairs can to a large extent be captured by typical gravity variables such as distance and adjacency but also by policy-related variables such as membership in the Eurozone or participation in the Schengen Agreement. But our focus lies on the substantial degree of heterogeneity in trade integration across industries. Our results suggest that modeling trade costs as a "one-fits-all" impediment is clearly at odds with empirical evidence.

We investigate the role of several sectoral characteristics in explaining trade integration across industries, with a particular emphasis on policy-related variables such as the extent of Technical Barriers to Trade (TBTs). Such barriers are a predominant concern in today's global trade negotiations, and for the WTO in particular as it precisely seeks to ensure that "technical regulations and standards, including packaging, marking and labelling requirements [...] do not create unnecessary obstacles to international trade."[3] We find that trade integration is indeed lower in countries and industries where TBTs are strong, suggesting that there is room left for policy action and that further gains are possible through the reduction of those barriers. We also show that trade integration tends to be high for industries characterized by high productivity, low transportation costs and a high degree of transparency in public procurement. From a dynamic perspective, average trade integration has

---

[3] Agreement on Technical Barriers to Trade (p.117). This Agreement, negotiated during the Uruguay Round, is an integral part of the WTO Agreement.

improved for most countries over the period 1999-2003, as well as individually for a large number of industries in our sample.

We also contrast our empirical approach of explaining the variation in bilateral trade integration to the standard gravity approach of regressing bilateral trade flows. We believe this is an important exercise that yields two insights. First, our approach allows us to identify the effects of a number of crucial explanatory variables such as weight-to-value that would normally drop out in a correctly specified gravity equation due to perfect collinearity with fixed effects. Second, we show that misspecifying a standard gravity equation by omitting multilateral resistance variables leads to a substantial bias in the estimated trade cost elasticities. Our approach avoids this problem because our trade integration measure has an in-built control for time-varying multilateral resistance.

The third contribution of the paper is to explore to what extent the increase in trade flows over recent years can be explained by the decrease in trade barriers. For that purpose, we use our model to decompose the growth in trade into two main components – (1) the growth of manufacturing output and (2) improvements in trade integration. We find that on average 58 percent of the growth in trade can be accounted for by changes in bilateral trade barriers and multilateral resistance, while the rest can be attributed to the secular growth of manufacturing output. This contribution of improved trade integration is quantitatively larger than the contribution found by Baier and Bergstrand (2001) for a sample of OECD countries. As we focus on the European Union, our results suggest that on average, intra-EU trade integration has progressed faster than integration amongst other developed countries.

The paper is organized as follows. In Section 2 we develop a general equilibrium model with industry-specific trade costs. In that section we also derive an analytical solution for time-varying industry-level multilateral resistance variables and our trade integration measure. Section 3 presents our dataset. Section 4 reports our empirical results, focusing on the determinants and time trend of trade integration. We also present the decomposition of trade growth into output growth and improvements in trade integration. Section 5 provides robustness checks and Section 6 concludes.

2 A Model with Industry-Specific Trade Costs

Our model closely follows the seminal paper by Anderson and van Wincoop (2003). Their general equilibrium model of trade results in a gravity equation that incorporates trade costs. The key insight from their gravity equation is that bilateral trade flows are not determined solely by the absolute bilateral trade barrier between two countries but rather by their bilateral *relative* to their average trade barrier. Anderson and van Wincoop (2003) refer to the appropriate average trade barrier as "multilateral resistance."

As a generalization, Anderson and van Wincoop (2004) model bilateral trade for an individual industry that is characterized by industry-specific bilateral trade costs and an industry-specific elasticity of substitution. We follow Anderson and van Wincoop (2004) in modeling trade flows with heterogeneous trade costs and heterogeneous elasticities of substitution at the industry level. The innovation of our approach is to derive an analytical solution for time-varying industry-specific multilateral resistance variables that can be related to observable data. With this solution at hand, we are able to derive a micro-founded measure of industry-specific bilateral trade integration that nets out multilateral resistance effects. Bergstrand (1989, 1990) also derives gravity equations for industry-level trade flows but does not focus on multilateral resistance.

2.1 The Basic Framework

Denote x_{ij}^k as nominal exports from country i to country j in goods associated with industry k. Suppose that consumers in country j allocate expenditure x_j^k on industry-k goods and that their preferences over these goods can be described by a standard CES utility function as

$$C_j^k \equiv \left(\sum_{i=1}^{J} \left(c_{ij}^k \right)^{\frac{\sigma_k - 1}{\sigma_k}} \right)^{\frac{\sigma_k}{\sigma_k - 1}} \tag{1}$$

where c_{ij}^k is real consumption of industry-k goods from country i by country-j consumers and where the elasticity of substitution σ_k is specific to industry k and assumed to exceed unity, $\sigma_k > 1$. Furthermore suppose that the factory gate price of industry-k goods from country i is denoted by p_i^k and that trade costs associated with the trade cost factor $t_{ij}^k \geq 1$ are incurred when these goods are shipped to country j such that the price faced by country-j consumers, denoted by p_{ij}^k, can be written as $p_{ij}^k = t_{ij}^k p_i^k$. The demand function for exports $x_{ij}^k = p_{ij}^k c_{ij}^k$ then follows as

$$x_{ij}^k = \left(\frac{p_{ij}^k}{P_j^k} \right)^{1-\sigma_k} x_j^k = \left(\frac{t_{ij}^k p_i^k}{P_j^k} \right)^{1-\sigma_k} x_j^k \tag{2}$$

where the price index P_j^k can be derived as

$$P_j^k = \left(\sum_{i=1}^{J} \left(p_{ij}^k \right)^{1-\sigma_k} \right)^{\frac{1}{1-\sigma_k}} \tag{3}$$

2.2 The Gravity Equation

Denote output of industry-k goods by country-i firms as y_i^k and impose market-clearing as

$$y_i^k = \sum_{j=1}^{J} x_{ij}^k \tag{4}$$

Substituting the demand function (2) into the market-clearing condition (4) and rearranging yields

$$p_i^k = \left(\sum_{j=1}^{J} \left(\frac{t_{ij}^k}{P_j^k} \right)^{1-\sigma_k} \frac{x_j^k}{y_i^k} \right)^{\frac{1}{\sigma_k - 1}} \tag{5}$$

Plug equation (5) back into the demand function (2) and, as in Anderson and van Wincoop (2004), define outward multilateral resistance for industry-k goods from country i as

$$\Pi_i^k \equiv \left(\sum_{j=1}^{J} \left(\frac{t_{ij}^k}{P_j^k} \right)^{1-\sigma_k} \frac{x_j^k}{y^k} \right)^{\frac{1}{1-\sigma_k}} \tag{6}$$

where y^k is world output in industry k. Rearrange to arrive at a gravity equation for industry k

$$x_{ij}^k = \frac{y_i^k x_j^k}{y^k} \left(\frac{t_{ij}^k}{\Pi_i^k P_j^k} \right)^{1-\sigma_k} \tag{7}$$

Trade flows x_{ij}^k depend on supply y_i^k of the k-good from country i and expenditure x_j^k for the good in country j. Large bilateral trade costs t_{ij}^k reduce bilateral trade, whereas large average outward trade barriers of country i (i.e., large Π_i^k) and large average inward trade barriers of country j (i.e., large P_j^k) lead to more bilateral trade. Substituting the solution for p_i^k in (5) and the definition of Π_i^k in (6) into the price index (3) yields

$$P_j^k = \left(\sum_{i=1}^{J} \left(\frac{t_{ij}^k}{\Pi_i^k} \right)^{1-\sigma_k} \frac{y_i^k}{y^k} \right)^{\frac{1}{1-\sigma_k}} \tag{8}$$

which is inward multilateral resistance for industry-k goods entering country j.

2.3 Solving for Multilateral Resistance

A problem that arises in empirical work is that we do not have data for the multilateral resistance terms P_j^k and Π_i^k in gravity equation (7). The method we employ here is to solve for these terms analytically as a function of observable trade flows. We exploit the fact that multilateral resistance is related to the amount of trade a country conducts with itself (see Novy, 2008). Intuitively, if a country's trade barriers with the rest of the world are high (i.e., if the country's multilateral resistance is high), the country will trade a lot domestically.

To see this formally, use gravity equation (7) and consider domestic trade flows for industry-k goods

$$x_{ii}^k = \frac{y_i^k x_i^k}{y^k} \left(\frac{t_{ii}^k}{\Pi_i^k P_i^k} \right)^{1-\sigma_k} \tag{9}$$

where t_{ii}^k are domestic trade costs for industry-k goods, for example domestic transportation costs. Equation (9) can be solved for the product of outward and inward multilateral resistance as

$$\Pi_i^k P_i^k = \left(\frac{y_i^k x_i^k}{x_{ii}^k y^k} \right)^{\frac{1}{1-\sigma_k}} t_{ii}^k \tag{10}$$

Note that we do not impose zero domestic trade costs since the trade cost factor t_{ii}^k may exceed unity. Equation (10) implies that for given t_{ii}^k and σ_k, it is easy to measure multilateral resistance since the quantities on the right-hand side are observable.

2.4 A Micro-Founded Measure of Industry-Specific Trade Integration

The solution for multilateral resistance can be exploited to solve the model. Gravity equation (7) contains the product of outward multilateral resistance of country i and inward multilateral resistance of country j, $\Pi_i^k P_j^k$, whereas equation (10) provides a solution for $\Pi_i^k P_i^k$. It is therefore useful to multiply gravity equation (7) by the corresponding gravity equation for trade flows in the opposite direction, x_{ji}^k, to obtain a bidirectional gravity equation that contains both countries' outward and inward multilateral resistance variables. This yields

$$x_{ij}^k x_{ji}^k = \frac{y_i^k y_j^k x_i^k x_j^k}{y^k y^k} \left(\frac{t_{ij}^k t_{ji}^k}{\Pi_i^k P_i^k \Pi_j^k P_j^k} \right)^{1-\sigma_k} \tag{11}$$

Substitute the solution for multilateral resistance given in equation (10) to obtain

$$x_{ij}^k x_{ji}^k = x_{ii}^k x_{jj}^k \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{1-\sigma_k} \tag{12}$$

From equation (12) it is easy to solve for the trade cost factors as

$$\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} = \left(\frac{x_{ii}^k x_{jj}^k}{x_{ij}^k x_{ji}^k} \right)^{\frac{1}{\sigma_k - 1}} \tag{13}$$

It is only possible to infer relative trade costs, in this case bilateral trade costs $t_{ij}^k t_{ji}^k$ relative to intranational trade costs $t_{ii}^k t_{jj}^k$.[4] We do not impose trade cost symmetry so that t_{ij}^k and t_{ji}^k on the left-hand side of equation (13) may be asymmetric ($t_{ij}^k \neq t_{ji}^k$). As Anderson and van Wincoop (2003, footnote 11) point out, it is problematic to infer the degree of trade barrier asymmetry from trade data because there are multiple combinations of t_{ij}^k and t_{ji}^k that can give rise to the same trade flows x_{ij}^k and x_{ji}^k. We therefore take the square root to get an expression for the *average* bilateral trade barrier. Thus, the average relative trade barrier can be expressed as

$$\theta_{ij}^k \equiv \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{\frac{1}{2}} = \left(\frac{x_{ii}^k x_{jj}^k}{x_{ij}^k x_{ji}^k} \right)^{\frac{1}{2(\sigma_k - 1)}} \tag{14}$$

We interpret θ_{ij}^k as a micro-founded measure of bilateral industry-specific trade frictions, or the inverse of bilateral trade integration. The more two countries trade with each other (i.e., the higher $x_{ij}^k x_{ji}^k$), the lower is our measure of relative trade frictions ceteris paribus. Conversely, if the two countries start trading more domestically (i.e., the higher $x_{ii}^k x_{jj}^k$), the higher is our measure of relative trade frictions ceteris paribus.

For the interpretation of θ_{ij}^k it is also helpful to think of two opposite extreme cases – a frictionless world with no trade costs on the one hand, and a closed economy on the other. In a frictionless world, all trade cost factors equal unity ($t_{ij}^k = t_{ji}^k = t_{ii}^k = t_{jj}^k = 1$) and θ_{ij}^k would be one, implying that bilateral and domestic barriers are the same. In the case approaching a closed economy, bilateral

---

[4] On this point also see Anderson and van Wincoop (2004, p.709).

trade $x_{ij}^k x_{ji}^k$ tends towards zero and thus θ_{ij}^k tends towards infinity, implying that bilateral barriers are prohibitive relative to domestic barriers.

Last but not least, we stress that our trade integration measure is valid more generally beyond the particular framework presented above. We have derived θ_{ij}^k in equation (14) from a model based on a CES demand system in combination with the Armington assumption that goods are differentiated by country of origin. But as we show in the technical appendix, we can also derive our trade integration measure from a range of other leading trade theories. These include the Ricardian trade model by Eaton and Kortum (2002), Chaney's (2008) extension of the Melitz (2003) heterogeneous firms model to asymmetric countries as well as the heterogeneous firms model by Melitz and Ottaviano (2008) with linear non-CES demand and endogenous markups.

The reason why our trade integration measure θ_{ij}^k is consistent with a broad range of trade models is related to the fact that they all lead to gravity equations that have a similar structure as equation (7).[5] Intuitively, the gravity equation is an expenditure equation that indicates how consumers allocate their expenditure across countries subject to trade frictions (Baldwin and Taglioni, 2006). Gravity equations arise regardless of *why* consumers want to buy goods from foreign countries. In an Armington world, consumers buy foreign goods because those goods are inherently different and consumers prefer variety. In a Ricardian world, countries produce goods according to comparative advantage and consumers buy foreign goods because they are cheaper. It turns out that the particular motivation behind foreign trade is not crucial to understand the role of trade frictions.[6]

2.5 A Comparison to Alternative Measures

We now contrast θ_{ij}^k with alternative measures of trade integration. First, compare θ_{ij}^k with a bilateral trade to output ratio, $\left(x_{ij}^k x_{ji}^k\right) / \left(y_i^k y_j^k\right)$. Various forms of such ratios have been used as measures of trade openness (Harrigan, 1996). Dividing gravity equation (11) by $y_i^k y_j^k$ and substituting equation (10) yields

$$\frac{x_{ij}^k x_{ji}^k}{y_i^k y_j^k} = \frac{x_{ii}^k x_{jj}^k}{y_i^k y_j^k} \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{1-\sigma_k} \tag{15}$$

Theory thus predicts that apart from the trade cost parameters, the bilateral trade to output ratio is also determined by the two countries' domestic trade shares, $\left(x_{ii}^k x_{jj}^k\right) / \left(y_i^k y_j^k\right)$.

To compare the trade to output ratio with θ_{ij}^k, take the trade cost parameters t_{ij}^k, t_{ji}^k, t_{ii}^k and t_{jj}^k as given. This pins down θ_{ij}^k in equation (14). However, the trade to output ratio can still move. For

---

[5] Feenstra, Markusen and Rose (2001) and Evenett and Keller (2002) also show that gravity equations are consistent with various competing trade models.

[6] Deardorff (1998) argues that in a Heckscher-Ohlin framework, trade frictions prevent factor price equalization so that for the large majority of goods, only one country is the lowest-cost producer. In the presence of trade frictions, trade in a Heckscher-Ohlin world therefore resembles trade in an Armington world. Grossman (1998) provides the following intuition for the gravity equation: "*Specialization* lies behind the explanatory power [of the gravity equation], and of course some degree of specialization is at the heart of any model of trade. Thus, the derivation of the gravity equation need not make reference to any particular trade model at all [...] This is true no matter what supply-side considerations give rise to the specialization, be they increasing returns to scale in a world of differentiated products, technology differences in a world of Ricardian trade, large factor endowment differences in a world of Heckscher-Ohlin trade, or (small) transport costs in a world of any type of endowment-based trade." (Emphasis in the original.)

example, suppose that trade costs with a third country increase, say, t_{im}^k with $m \neq i, j$. The increase in t_{im}^k pushes up the domestic trade ratio x_{ii}^k/y_i^k and therefore also the bilateral trade to output ratio in equation (15).[7] But this is a multilateral resistance movement and does not imply that trade integration between i and j has improved. In contrast, θ_{ij}^k does not react to multilateral resistance movements because θ_{ij}^k only reflects changes in bilateral and domestic trade costs, not changes in third-party trade costs. In conclusion, trade to output ratios are distorted by multilateral resistance effects.

Second, compare θ_{ij}^k with the "*phi*-ness" of trade, which is given by the simple ratio of bilateral to domestic trade flows and is commonly used to estimate sectoral border effects (Baldwin et al., 2003, Head and Mayer, 2004, Head and Ries, 2001).[8] Theory shows that the border effect can arise because a high degree of substitution between domestic and imported goods may lead to a high responsiveness of trade flows even in the case of very modest trade barriers. However, in an industry with a high elasticity σ_k, consumers are so price-sensitive that a high ratio of domestic over bilateral trade does not necessarily reflect high bilateral trade barriers but rather a high degree of competition.[9] The definition of θ_{ij}^k combines the *phi*-ness measure with an exponent that involves the industry-specific elasticity of substitution σ_k so in contrast to *phi*-ness, θ_{ij}^k is able to separate this competition effect from the trade barrier effect. As equation (14) shows, a higher elasticity of substitution implies *lower* trade frictions θ_{ij}^k. Likewise, as we show in the technical appendix, the theoretical models by Eaton and Kortum (2002), Chaney (2008) and Melitz and Ottaviano (2008) also require that the trade ratio be scaled by an industry-specific exponent. We therefore argue that θ_{ij}^k improves on the *phi*-ness measure because it accounts for differences in market power and competition across industries.[10]

3 Data and Descriptive Statistics

To compute our measure of trade integration θ_{ij}^k across industries, countries and time, we need the domestic trade of countries i and j in industry k, x_{ii}^k and x_{jj}^k, as well as their bilateral exports, x_{ij}^k and x_{ji}^k, at time t. As in previous literature (for instance, Chen, 2004, Evans, 2003, Head and Mayer, 2000, Nitsch, 2000, Wei, 1996), domestic trade for country i is given by its gross output in industry k, y_i^k, minus total exports of country i to the rest of the world in that industry. The elasticities of substitution σ_k are taken from Hummels (2001a).[11] Our sample includes trade flows for 166 manufacturing industries across 11 EU countries at the 4-digit Nace rev.1 level between 1999 and 2003. The sample is balanced over time. The 11 countries are Austria, Denmark, Finland, France, Germany, Italy, Ireland, the Netherlands, Spain, Portugal and the United Kingdom. In the data appendix we provide a detailed description of the data and their sources.[12]

---

[7] Formally, the increase in multilateral resistance follows from equation (6), implying an increase in x_{ii}^k through equation (10).

[8] Alternatively, border effects can be estimated from a gravity equation that combines domestic and international trade flows as the dependent variable (Chen, 2004, Evans, 2003, McCallum, 1996, Wei, 1996, Wolf, 2000).

[9] The standard CES markup is given by $\sigma_k/(\sigma_k - 1)$ and thus inversely related to σ_k.

[10] We show in Section 5 that regressions based on the *phi*-ness measure yield a considerably lower R^2 than regressions of θ_{ij}^k.

[11] Hanson and Xiang (2004) also use the elasticities from Hummels (2001a).

[12] Due to data limitations not all possible trade flow combinations across industries and countries are available. Although it would in principle be possible to span the longer period from 1997 to 2003, this would come at the cost of losing all observations for Germany, a core EU country, because German sectoral output data are missing prior to 1999.

Our dataset comprises a total of 15,040 domestic and bilateral trade flow observations. Only five bilateral trade flows are equal to zero, which is not surprising given the huge volume of intra-industry trade in the EU.[13] Those five cases would normally feature as missing values of θ_{ij}^{k}. But zero trade flows may contain valuable information as to why such low levels of trade are observed. It therefore seems more appropriate to associate these cases with large trade frictions. The approach we adopt is to replace the zeros with a value of one Euro, thus associating them with large values of θ_{ij}^{k}.[14]

Table 1 contains summary statistics for individual industries. Due to space constraints we list industries at the more aggregated 3-digit level, reporting averages of θ_{ij}^{k} over the lower-level 4-digit classifications as well as their maximum and minimum values. Industries are ordered by decreasing value of θ^{k}, which ranges from 146.91 to 1.15. We also report the average elasticities of substitution σ_k and the average weight-to-value ratios, measured as kilograms per Euro exported.

A first cursory glance at Table 1 attests the intuitive nature of the trade integration measure θ_{ij}^{k}. Trade integration is lowest for Bricks, followed by Cement, lime, plaster. The latter is also the industry with the largest average weight-to-value ratio (over 13 kg/Euro), indicating a very low transportability of the goods. Another related industry is Articles of concrete. Note that the geographic market for cement and concrete is very local since the perishable nature of such "wet" products constrains the distance over which they can be delivered.

Printing and Publishing are traded very little, too, which is hardly surprising given the reliance of such products on specific languages. This finding is consistent with earlier studies showing that trade in such sectors is subdued. Harrigan (1996) shows that the volume of trade relative to output in the OECD in 1985 is the lowest in Printing and Publishing. Finally, some of the sectors with high values of θ^{k} belong to the food industry, for which the perishability of many goods is most likely an important deterrent to trade. For example, Fruit and vegetables are ranked in fourth place, and Other food products display the biggest individual value of θ^{k} equal to 1,056.

Table 2 reports the average level of θ_i for the 11 individual countries in our sample. Ireland appears as the least integrated country, followed by Spain, Denmark, Finland and Italy, whereas France, the Netherlands and Germany are the most trade integrated.

Finally, Figure 1 plots the time series evolution of θ_t, averaged across countries and industries. It is interesting to observe that despite fluctuations from one year to the next, trade frictions display a downward trend, suggesting that the countries and industries we consider have on average become more bilaterally integrated over time.

4 Empirical Results

In the first part of this section, we analyze the determinants of trade integration across country pairs, industries and years. In the second part, we explain how our approach differs from standard

---

[13] The three industries with zero trade flows are Publishing of newspapers, Builders' carpentry and joinery and Mattresses. We consider those observations as "true zeros" because the corresponding output values are positive but exports are zero. We exclude the cases where output is zero but exports are positive as well as the cases where both output and exports are zero.

[14] Our treatment of the zero trade flows does not distort the overall results, see the robustness check in Section 5.

gravity. In the third part, we focus on the time series patterns of trade integration. In the fourth part, we decompose the growth in trade over the sample period into two components – the growth in manufacturing output and improvements in trade integration.

4.1 The Determinants of Trade Integration

We first analyze the determinants of bilateral EU trade integration. We estimate

$$\ln \theta_{ij,t}^{k} = \psi_t + \lambda_K + \beta \; Geo_{ij,t}^{k} + \zeta \; Policy_{ij,t}^{k} + \alpha \; Costs_{ij,t}^{k} + \gamma \; Controls_{ij,t}^{k} + \epsilon_{ij,t}^{k} \tag{16}$$

where $Geo_{ij,t}^{k}$ is a set of variables related to geography and transportation costs, $Policy_{ij,t}^{k}$ is a set of policy-related factors, $Costs_{ij,t}^{k}$ includes other types of costs such as fixed costs of exporting and productivity, and $Controls_{ij,t}^{k}$ includes variables controlling for measurement issues. Given that some of the explanatory variables vary across sectors only, we include industry fixed effects λ_K at the more aggregated level of 3-digit industries, assuming that the 4-digit groupings k are different varieties of the corresponding, more aggregated 3-digit sector K (Hummels, 2001a). The inclusion of other explanatory variables that only vary across country pairs precludes us from controlling for country pair fixed effects.[15] ψ_t denotes year intercepts. β, ζ, α and γ are vectors of coefficients to be estimated and $\epsilon_{ij,t}^{k}$ is an error term. A higher value of θ_{ij}^{k} should be interpreted as a lower degree of trade integration. We estimate our baseline results with OLS. To control for possible autocorrelation in each of the individual series of the panel, robust standard errors are adjusted for clustering at the 4-digit Nace rev.1 level in each country pair (15,040/5 = 3,008 clusters). As a robustness check, we also report results obtained by Poisson Pseudo-Maximum Likelihood (PPML) estimation, see Section 5.

Geography/Transportation Costs Variables Table 3 reports our main results. Column (1) only includes variables that are related to geography and transportation costs. Across countries, standard gravity variables perform well in explaining the variation in trade integration. Trade integration decreases with international distances D_{ij} and increases with domestic distances $D_{ii} \times D_{jj}$. It is higher between countries that share a common border Adj_{ij} and a common language $Lang_{ij}$.

We also consider proxies of transportation costs that vary across industries. First, we consider the ratio of *c.i.f.* and *f.o.b.* trade values for each industry and country pair, averaged in each year across all EU partners in order to minimize measurement error, $cfob_t^k$ (Harrigan, 1993).[16] Second, we consider the weight-to-value ratio of exports for each industry and country pair, also averaged in each year across all partners, wv_t^k. We do not consider *bilateral* weight-to-value ratios as Hillberry and Hummels (2000) show that bilateral weight-to-value significantly falls with distance. This suggests that the commodity composition of trade is sensitive to bilateral costs and that weight-to-value is endogenous.[17] Overall, since the freight component of trade costs is higher for bulky, high weight-to-value raw materials than for manufactures, weight-to-value should be associated with higher values for θ_{ij}^{k}. As expected, higher transportation costs as proxied by $cfob_t^k$ and wv_t^k decrease trade integration.

---

[15] As a robustness check, we show that our results hold up with country pair fixed effects, see Section 5.

[16] Apart from insurance and other possible discrepancies between partners, the ratio of *c.i.f.* and *f.o.b.* for intra-EU trade values mostly captures transportation costs by *road.* In 1998, 57.8 percent of total intra-EU15 trade went by road, 22.8 percent by sea, 3.9 percent by rail, 3.9 percent by air, 0.9 percent by inland waterways and 0.8 percent by pipeline (European Commission, 2000).

[17] Harrigan (2005) and Baldwin and Harrigan (2007) also find that export unit values are positively related to distance.

Policy Variables In Column (2) we consider the role of several policy-related factors in affecting trade integration. Across countries, we include a dummy variable for Finland and Austria, FI, AT_{ij}, as these two countries were the last in our sample to join the EU (in 1995), and they appear less integrated compared to the other countries. We also explore the effect of not adopting the Euro by including a dummy variable, $noEURO_{ij,t}$, which is equal to one for Denmark and the United Kingdom from 2002 onwards when the common currency was introduced.[18] As expected, non-adoption of the Euro has lowered the extent of trade integration for these two countries since 2002.

We also add a variable to capture the effects of the Schengen Agreement which abolishes physical border controls among participating EU countries. The date of the first implementation of the agreement differs across countries. Ireland and the United Kingdom have not started implementation yet. The coefficient on this variable, $Schengen_{ij,t}$, is negative and highly significant in Column (2), suggesting that the abolition of border controls among the participating countries has helped to foster trade integration, most probably through the elimination of time delays and administrative burdens that were previously experienced at borders.

We also consider policy variables that vary across industries. First, we address the role of Technical Barriers to Trade (TBTs). TBTs result from norms (regulations and standards) that affect the sale of goods in some markets by requiring specific product characteristics or production processes. Baldwin (2000) stresses the importance of TBTs in shaping trade flows between countries and industries. He argues that in the case of Europe, such barriers have become more and more visible over time, especially since tariff barriers were completely eliminated by 1968.[19] In addition, TBTs are a predominant concern in today's global trade negotiations and for the WTO in particular. Data on TBTs are hard to find, however, so our approach to measuring TBTs uses two different sources of information: a cross-country survey of EU managers who reveal whether they consider TBTs a problem for trade, and a ranking of industries according to the relevance of TBTs. The interacted variable, TBT^k_{ij}, increases with the extent of TBTs across countries and industries. The data appendix provides details. In Column (2) TBTs are found to lower trade integration, suggesting that some room is left for policy action and that the removal of such barriers might promote trade integration. This result is consistent with Chen (2004) who finds that TBTs in Europe are associated with larger border effects in international trade.

As is often the case, results based on survey data should be interpreted with caution. In particular, the cross-country measurement of TBTs might be problematic. For instance, if a manager states that TBTs are not a problem, this does not necessarily mean that TBTs are absent or unimportant. Instead, if that manager's company or industry is protected by TBTs initiated by his country, he might actually prefer them. This possibility might bias our cross-country measurement of TBTs.[20] In Column (3) we therefore report estimates when only including our industry-specific variable on TBTs,

---

[18] The Euro exchange rates were fixed in 1999 but Euro notes and coins and thus a greater degree of transparency only followed in 2002. The huge literature on the trade-creating effects of currency unions raises the issue of endogeneity of the common currency dummy (Rose, 2000, Baldwin, 2006). In this paper we do not attempt to address this issue.

[19] As explained by Baldwin (2000, p.255), "Europe's first liberalization efforts focused on the 'easy' barriers, tariffs and quotas. With these eliminated by 1968, liberalization attention turned to TBTs."

[20] If this is indeed the case, we will underestimate the impact of TBTs on trade integration so that our coefficients can be considered lower-bound estimates.

TBT^k. The results remain consistent, i.e., when TBTs are high, trade integration is low.[21]

Furthermore, it is well-established that national governments often favor domestic over foreign firms for some of their purchases, even if foreign suppliers could actually offer them a better deal (Davies and Lyons, 1996). Firms in such "public procurement" markets are hence protected from foreign imports, sometimes to such an extent that trade may be completely suppressed. In the last few years, competition has been steadily increasing in those markets, with the proportion of EU15 public procurement contracts openly advertised in the *Official Journal* of the EU steadily increasing from 8.4 percent in 1995 to 16.8 percent in 2005 (Eurostat). To investigate whether this opening of the markets has helped to lower trade frictions, we rely on time-varying cross-country data on the share of public procurement contracts advertised in the *Official Journal* and compute the average across partners. As similar data across industries are not always available, we compute an indicator variable for high public procurement sectors.[22] We interact this indicator variable with the time-varying data for each country pair, denoting the variable so obtained by $Proc^k_{ij,t}$. Column (2) reveals that the opening of public procurement markets to foreign competition has indeed been successful in fostering trade integration across countries and industries and over time.

Next, Value-Added Taxes (VATs), which differ extensively across EU countries and goods, are characterized by the so-called "destination principle": VAT for a good is paid in the country where it is sold, not where it is manufactured, implying that VATs uniformly affect domestic trade and imports. In our model, the imposition of VATs by country i thus increases both domestic costs t_{ii} and importing costs t_{ji} by the same proportion, leaving bilateral trade integration θ^k_{ij} unchanged. In Column (2) we include the log of one plus the average VAT rate across partners at the sectoral level, VAT^k_{ij}, and interestingly find that higher VATs are associated with significantly lower trade frictions.

One possible interpretation relates to VAT fraud. Baldwin (2006) documents that VAT fraud is a very serious problem in the EU. Since the removal of Europe's internal borders in 1993, trade statistics are now collected by VAT authorities, creating a direct correlation between trade statistics and VAT fraud: EU firms have an incentive to over-report their exports (to get the VAT rebate) and to under-report their imports (which are subject to VAT). This inflates export statistics relative to import statistics. As a result, in the countries and industries where VATs are high, the value of θ^k_{ij} might be lower simply because intra-EU export statistics are inflated due to VAT fraud.[23]

Other Costs The new trade literature on heterogeneous firms rationalizes why some firms export and others do not. In particular, the models by Melitz (2003), Melitz and Ottaviano (2008) and others show that only the most productive firms participate in foreign markets. Only those firms are productive enough to cover the fixed costs of exporting or to withstand tough competition on foreign markets. Due to this selection effect we should observe that ceteris paribus, more productive industries trade more on foreign markets than less productive industries, i.e., we should observe a negative

---

[21] The use of this sectoral variable on TBTs might be criticized on the grounds that it captures changes in TBTs rather than levels (see the data appendix). As in Chen (2004), we experimented using a dummy variable only, which is equal to one when TBTs are present. The results remain unaffected.

[22] Examples of high public procurement industries are Shipbuilding, Rail stock, Pharmaceuticals, Aerospace, Telecoms.

[23] If the intra-EU exports of country i are inflated because of VAT fraud, then x^k_{ij} will be high relative to x^k_{ii}, leading to a lower value of θ^k_{ij}.

relationship between productivity and trade frictions θ_{ij}^k. In Column (4) of Table 3 we therefore include average real labor productivity across industries and countries. Moreover, a growing empirical literature shows that as trade costs fall, less efficient firms exit from the market and average industry productivity increases (for example, see Bernard, Jensen and Schott, 2006, for the U.S. and Greenaway and Kneller, 2008, for the UK). To mitigate potential endogeneity problems we lag productivity by one year, $Prod_{ij,t-1}^k$, reducing the number of observations by one cross section. The negative relationship between productivity and θ_{ij}^k confirms the prediction that more productive industries trade more on foreign markets.[24]

We also include a dummy variable in Column (4) that controls for the five zero trade observations in our sample, $Zeros_{ij,t}^k$. We find that trade frictions between two countries in a given industry are stronger whenever either of them does not export. As models of heterogeneous firms typically predict that firms only export if they are able to cover the fixed costs of exporting, the result might reflect the role of these fixed costs.

Controls Finally, we control for some measurement issues and the possibility that the magnitude of θ_{ij}^k may be affected by the nature of trade. In particular, the value of θ_{ij}^k computed under intra-industry trade will be lower than under comparative advantage driven by technology or factor endowment differences across countries, even if the actual friction is the same.[25] To control for this potential bias we compute the (absolute) difference in capital shares between countries and industries in order to proxy for differences in factor endowments. For a larger differential in capital shares, trade is more likely based on comparative advantage, leading to an overestimated θ_{ij}^k. This variable, $KS_{ij,t}^k$, displays the expected positive (but insignificant) coefficient in Column (5) but its inclusion leaves the coefficients on the other variables virtually unchanged.

The previous reasoning presumes that differences in factor endowments can be fully captured by differences in capital shares, the latter being computed at the *industry* level. One obvious limitation stems from the use of data disaggregated at the level of industries and not at the level of products. Industry classifications inevitably aggregate many different types of products into one single category so that the volume of intra-industry trade can appear more important than it actually is. For instance, many different types of steel are produced, from flat-rolled to specialty steels, and it may be that the production of some types of steel requires resources or technologies in which one country has a comparative advantage. However, since all these types of steel are aggregated into one category, it appears as if the countries export and import almost identical products while in reality they export one type of steel and import another type.

Thus, the larger the number of varieties in each industry category, the more likely the industry aggregates trade flows with different comparative advantages in different varieties, but in the data this will show up as intra-industry trade, i.e., more balanced trade flows between the two countries.

---

[24] We also tried to use data on productivity for a country that is not included in our sample, i.e., the US. But those data are aggregated at the 2-digit level and proved insignificant in explaining trade integration across EU countries, possibly due to the higher level of aggregation.

[25] With intra-industry trade, the denominator of (14) will be given by the product of two balanced trade flows. But with comparative advantage, the denominator of (14) will be given by the product of two unbalanced trade flows, leading to a relatively small magnitude in the denominator and thus overestimated trade frictions.

The resulting θ_{ij}^k is therefore expected to be smaller because the denominator in equation (14) is the product of two more or less balanced (intra-industry) trade flows at the *industry* level, masking imbalances of (inter-industry) trade flows at the *product* level.[26] To control for this aggregation bias, we include the number of product categories within each industry, $Goods^k$. In Column (5) it is indeed the case that a larger number of varieties is associated with a lower value for θ_{ij}^k. But the inclusion of this control hardly affects other estimates.

In Column (6) we replicate the specification of Column (5) but replace the country-varying variable on TBTs by the sectoral variable, TBT^k. It is reassuring that all results remain virtually unchanged, providing some evidence that they are not biased by the use of cross-country survey data on TBTs.

Finally, Table 4 reports a variance decomposition to illustrate the quantitative contribution of each factor in explaining the variation of θ_{ij}^k. Following Fields (2003), the contribution of each explanatory variable x_m to the total variance of $\theta_{ij,t}^k$ is calculated as $c_m = \beta_m cov\left(x_m, \theta_{ij,t}^k\right)/var\left(\theta_{ij,t}^k\right)$ where β_m is the partial regression coefficient of $\theta_{ij,t}^k$ on the explanatory variable x_m (holding all other explanatory variables constant). Columns (1) and (2) correspond to regressions (5) and (6) in Table 3, respectively.

Column (1) shows that regression (5) in Table 3 explains 80.8 percent of the variation in $\theta_{ij,t}^k$. 42.8 percent is attributable to the 3-digit industry fixed effects and 0.01 percent to the year fixed effects. Of the 38 percent that the remaining regressors explain, geography and transport costs alone are responsible for about 25 percent of the variation in θ_{ij}^k, the most important factor being weight-to-value (16.7 percent), followed by bilateral distance (5.4 percent). Policy variables explain 7.6 percent of the variation in θ_{ij}^k, which is far from negligible. TBTs are the most important factor (4.5 percent), while public procurement, Schengen and opting out of the Euro only play minor roles. The number of product categories within each industry explains up to 5.4 percent of the variation in θ_{ij}^k. Finally, the decomposition reported in Column (2), which corresponds to regression (6) in Table 3, gives similar results.

4.2 How Our Approach Differs from Standard Gravity

We now compare our specification to the standard gravity approach and illustrate the empirical pitfalls of not taking the theory seriously. We believe this is an important and necessary exercise.

Although specification (16) is related to standard gravity estimation, there are three important differences. First, standard gravity uses output $y_i^k y_j^k$ as the size variable. In contrast, the trade integration measure θ_{ij}^k includes domestic trade $x_{ii}^k x_{jj}^k$ because this nets out the multilateral resistance variables. Second, correctly specified gravity equations have to include time-varying industry fixed effects for each country pair to control for multilateral resistance variables (Baldwin and Taglioni, 2006). But a practical problem is that these fixed effects are perfectly collinear with many explanatory variables of interest such as $cfob_t^k$ and wv_t^k in Table 3. We avoid this problem since θ_{ij}^k nets out the time-varying multilateral resistance variables. Third, our theory calls for trade cost elasticities that vary across industries. Those are embodied in θ_{ij}^k.

---

[26] The product of two balanced trade flows yields a relatively large magnitude in the denominator and thus a relatively low θ_{ij}^k.

We now evaluate the bias that arises in standard gravity equations due to the omission of multilateral resistance and the assumption of coefficient homogeneity across industries. In its simplest form (henceforth the "naive" model), the traditional gravity model relates bilateral trade to output and bilateral trade costs where the trade cost elasticity α_2 is assumed to be the same across industries:

$$\ln\left(x_{ij}^k x_{ji}^k\right) = \alpha_1\ \ln\left(y_i^k y_j^k\right) + \alpha_2\ \ln\left(t_{ij}^k t_{ji}^k\right) + \epsilon_{ij}^k \tag{17}$$

We drop time subscripts for simplicity. Compare the naive specification in equation (17) to the log-linearized version of micro-founded gravity equation (12):

$$\ln\left(x_{ij}^k x_{ji}^k\right) = \beta_1\ \ln\left(x_{ii}^k x_{jj}^k\right) + \beta_2\left(1-\sigma_k\right)\ln\left(t_{ij}^k t_{ji}^k\right) - \beta_3\left(1-\sigma_k\right)\ln\left(t_{ii}^k t_{jj}^k\right) + \epsilon_{ij}^k \tag{18}$$

The size variable is domestic trade instead of output. Moreover, equation (18) includes domestic trade costs and the trade cost coefficients now vary across industries.[27] An alternative micro-founded gravity specification can be obtained by log-linearizing equation (11):

$$\ln\left(x_{ij}^k x_{ji}^k\right) = \ln\left(y_i^k y_j^k\right) + \ln\left(\frac{x_i^k x_j^k}{y^k y^k}\right) + \left(1-\sigma_k\right)\ln\left(t_{ij}^k t_{ji}^k\right) + \left(\sigma_k - 1\right)\left(\Pi_i^k P_i^k \Pi_j^k P_j^k\right)$$

To control for the unobservable multilateral resistance variables $\Pi_i^k P_i^k \Pi_j^k P_j^k$, we include fixed effects for each industry (at the 4-digit level) in each year and separately for countries i and j, denoted by γ_i^k and γ_j^k. Such fixed effects also wipe out the domestic demand terms (if included separately for each country) and world output in industry k. To ease the comparison with the naive specification, we keep the domestic output terms so that we estimate

$$\ln\left(x_{ij}^k x_{ji}^k\right) = \delta_1\ \ln\left(y_i^k y_j^k\right) + \delta_2\left(1-\sigma_k\right)\ln\left(t_{ij}^k t_{ji}^k\right) + \gamma_i^k + \gamma_j^k + \epsilon_{ij}^k \tag{19}$$

To keep our exposition as simple as possible, we assume that trade costs (domestic and international) are proxied by the log of distance only. We then estimate the naive specification (17) and compare its distance elasticity to the ones estimated in the micro-founded specifications (18) and (19). Table 5 reports our results. The upper panel refers to naive gravity while the two lower panels refer to the micro-founded models. For all panels, Column (1) assumes homogeneity on the trade cost coefficient across industries (bilateral and domestic). We relax this assumption in Column (2) where we report the average of the elasticities estimated for individual industries. In Columns (3) and (4), we replicate the specifications in Columns (1) and (2) but we allow the output variables, domestic distances and domestic trade to enter separately for each country.

As can be seen in Column (1) of the upper panel, the naive gravity equation yields an absolute distance elasticity of 3.589, which increases slightly to 3.741 once we allow for coefficient heterogeneity in Column (2).[28] The sectoral output regressor displays a coefficient slightly larger than the unity value predicted by theory. All elasticities remain similar in Columns (3) and (4).

---

[27] Combes, Lafourcade and Mayer (2005, equation 13) estimate a form of equation (18) where they constrain β_1 to unity. But they use aggregate trade data so that by construction the trade cost elasticity cannot vary across industries.

[28] The magnitude of this coefficient is larger than that typically obtained in the literature because the dependent variable is the product of bilateral exports and not unidirectional bilateral exports.

The middle panel turns to the micro-founded specification in (18). The coefficients on domestic distances and domestic trade flows are significant with the expected signs. But the coefficients on bilateral distance are markedly different. In Column (1), the absolute magnitude of the distance elasticity is large and equal to 4.514. It further increases to 5.131 once we allow for coefficient heterogeneity, as required by theory.[29] As highlighted by Anderson and van Wincoop (2003) and Baldwin and Taglioni (2006), the omission of multilateral resistance biases the magnitude of trade cost variables. The magnitude of the distance elasticity in the naive specification is underestimated by 30 percent (3.589/5.131-1 = -0.30) and by 28 percent if we allow the other controls to enter separately for each country (3.629/5.045-1 = -0.28). The omitted variable bias goes in the direction predicted by theory.[30]

The lower panel focuses on the micro-founded specification with fixed effects. The fixed effects do a good job in eliminating the bias on the distance coefficient. The distance elasticity is 4.204 in Column (1) and increases to 4.653 when we allow for coefficient heterogeneity in Column (2). In absolute terms, those elasticities are not as large as in the other micro-founded specification but the difference is relatively small (between 7 and 9 percent).

In summary, we show that omitting multilateral resistance (and assuming coefficient homogeneity across industries) yields biased distance elasticities, and the bias can be large (up to 30 percent in our sample). Our original specification (16) is consistent with the estimation of the micro-founded gravity equation (18), which differs from standard gravity regressions by including domestic trade as the size variable. In principle, fixed effect estimation in equation (19) is a viable micro-founded alternative but it obviates the identification of many explanatory variables of interest such as $cfob_t^k$ and wv_t^k because they are linear combinations of the fixed effects.

4.3 The Evolution of Trade Integration over Time

To investigate the evolution of trade integration over time, we run the following regression

$$\ln \theta_{ij,t}^k = \gamma_{ij}^k + \psi_t + \alpha \ \ln wv_{ij,t}^k + \varepsilon_{ij,t}^k \tag{20}$$

where γ_{ij}^k are country pair dummies interacted with industry intercepts (at the 4-digit level) to control for systematic differences between trade partners in each industry. The *bilateral* weight-to-value $wv_{ij,t}^k$ controls for compositional change over time (see Hummels, 2007).[31] The evolution of θ_{ij}^k is given by the estimated coefficients on the year dummies ψ_t. The annualized growth rate of these coefficients obtained from estimating equation (20) is equal to -0.94 percent over our pooled sample with a t-statistic of 11.48.[32] This result is highly significant and we conclude that trade integration has on average improved between 1999 and 2003.

---

[29] Chen (2004) also finds that the inclusion of fixed effects to control for multilateral resistance leads to larger magnitudes of the distance coefficients.

[30] The naive gravity equation yields a distance elasticity $\widehat{\alpha_2}$ such that $E(\widehat{\alpha_2}) = \alpha_2 + ab$ where $\alpha_2 < 0$ is the true elasticity, $a = \partial \ln (\Pi_i P_i \Pi_j P_j) / \partial \ln (t_{ij} t_{ji}) > 0$ and $b = (\sigma_k - 1) > 0$. The bias, equal to $ab > 0$, therefore leads to an underestimation of the absolute magnitude of the distance elasticity.

[31] In contrast to Section 4.1, we now consider the *bilateral* weight-to-value as distance is not included in the regression.

[32] The significance is calculated using the delta method.

We also estimate equation (20) separately for each industry. Figure 2 plots the annualized growth rates for the individual industries, distinguishing between significant (at the ten percent level) and insignificant growth rates.[33] About half of the industries did not experience any significant change over time (82 industries out of 160). However, among the industries that did, the vast majority (69 out of 78 significant) experienced an increase in trade integration. Games and toys display the strongest improvement, while trade integration worsened for Articles of concrete followed by Paper and Explosives. Note that according to equation (14), in the few cases where θ_{ij}^k increases over time, the worsening in bilateral trade integration could reflect a fall in domestic trade costs, $t_{ii}^k t_{jj}^k$, rather than an increase in bilateral trade costs, $t_{ij}^k t_{ji}^k$.

Lacking any product or firm-level data, we are not able to investigate whether the deepening of trade integration over time is due to the extensive or intensive margin of trade. On the one hand, Hummels and Klenow (2005) report evidence in favor of the extensive margin, while Helpman, Melitz and Rubinstein (2008) and Besedeš and Prusa (2007) argue that the intensive margin is quantitatively more important. To investigate this issue we would ideally need to observe more years and more highly disaggregated data. We leave this question for future research.

4.4 Explaining the Growth in Trade

Baier and Bergstrand (2001) try to explain the growth of world trade by tracing two fundamental channels through which trade can grow – (1) the growth of income and (2) reductions in trade barriers. For that purpose, they estimate a gravity equation in first differences and use the gravity coefficients to extrapolate the contributions of income growth and trade barrier reductions to the average growth of trade in their sample.

We use our model to perform a similar exercise of decomposing the growth of trade into these two contributions. But our approach differs on three counts. First, our approach is considerably easier to implement because our solution for time-varying multilateral resistance variables allows us to eliminate the unobservable and highly nonlinear price indices that appear in Baier and Bergstrand's (2001) gravity equation. Second, we use industry-level trade data instead of aggregate trade data. Third, our decomposition is not restricted to the sample average because we are able to provide decompositions for specific bilateral observations in each industry. These finer decompositions allow us to uncover a substantial degree of heterogeneity across various cuts of our dataset.

Our decomposition is based on gravity equation (11), which can be rearranged as

$$\begin{aligned} x_{ij}^k x_{ji}^k &= y_i^k y_j^k \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{1-\sigma_k} \frac{x_i^k x_j^k}{y^k y^k} \left(\frac{t_{ii}^k t_{jj}^k}{\Pi_i^k P_i^k \Pi_j^k P_j^k} \right)^{1-\sigma_k} \\ &= y_i^k y_j^k \left(\theta_{ij}^k \right)^{2(1-\sigma_k)} \frac{x_i^k x_j^k}{y^k y^k} \left(\frac{t_{ii}^k t_{jj}^k}{\Pi_i^k P_i^k \Pi_j^k P_j^k} \right)^{1-\sigma_k} \qquad (21) \end{aligned}$$

---

[33] We can only estimate equation (20) separately for 160 industries because of too few degrees of freedom for the other 6 industries. To conserve space we do not report the annualized growth rates for each industry but they are available upon request.

Using the solution for the multilateral resistance terms in (10), we can express the last two fractions on the right-hand side of equation (21) in terms of observable variables, yielding

$$x_{ij}^k x_{ji}^k = y_i^k y_j^k \left(\theta_{ij}^k\right)^{2(1-\sigma_k)} \frac{x_{ii}^k}{y_i^k} \frac{x_{jj}^k}{y_j^k} \tag{22}$$

We now take logs and first differences of equation (22) and divide by the left-hand side to obtain our decomposition equation

$$100\% = \underbrace{\frac{\Delta \ln \left(y_i^k y_j^k\right)}{\Delta \ln \left(x_{ij}^k x_{ji}^k\right)}}_{\text{(First factor)}} + \underbrace{\frac{2(1-\sigma_k)\Delta \ln \left(\theta_{ij}^k\right) + \Delta \ln \left(\frac{x_{ii}^k}{y_i^k} \frac{x_{jj}^k}{y_j^k}\right)}{\Delta \ln \left(x_{ij}^k x_{ji}^k\right)}}_{\text{(Second factor)}} \tag{23}$$

The denominator on the right-hand side of equation (23), $\Delta \ln \left(x_{ij}^k x_{ji}^k\right)$, is the growth of bilateral trade between countries i and j in industry k. The first factor thus represents the contribution of output growth, $\Delta \ln \left(y_i^k y_j^k\right)$, to the growth of bilateral trade.

The second factor is the contribution of trade integration in that industry. If bilateral trade frictions decline relative to domestic trade frictions, then $\Delta \ln \left(\theta_{ij}^k\right) < 0$. As $\sigma_k > 1$, it follows that the contribution of declining bilateral trade frictions is positive (given positive trade growth $\Delta \ln \left(x_{ij}^k x_{ji}^k\right) > 0$). The change in the domestic trade to output ratios, $\Delta \ln \left(\frac{x_{ii}^k}{y_i^k} \frac{x_{jj}^k}{y_j^k}\right)$, reflects changes in multilateral factors. In particular, from equation (21) it can be seen that this captures changes in world demand shares and changes in multilateral resistance relative to domestic trade costs.

Based on annual changes, we first calculate the two factors in equation (23) for each country pair-industry observation. To avoid giving undue weight to quantitatively unimportant observations that only represent a small fraction of total trade, we then compute a weighted average where the weights are the products of sectoral bilateral exports in the initial year 1999.[34]

Table 6 reports the average growth in trade and the contributions of the two factors. In the full sample, trade grew on average by 7.96 percent over the period.[35] 42 percent of this trade expansion can be accounted for by changes in the trading partners' output growth, whereas changes in trade integration account for 58 percent of the growth in trade. These proportions are different from those reported by Baier and Bergstrand (2001) who argue that two-thirds of the growth in trade amongst OECD countries between 1958 and 1988 can be explained by the growth of output. Our results therefore suggest that trade integration amongst EU countries over our sample period has progressed relatively faster. Likewise, Jacks, Meissner and Novy (2008) find that more than half of the global trade boom between 1870 and 1913 is attributable to reductions in trade costs.

---

[34] Unweighted decompositions are very similar to the weighted decompositions reported in the paper and are available upon request.

[35] To be precise, this is the average annual growth rate of the product of bilateral trade, $x_{ij}^k x_{ji}^k$, weighted as described above.

The next rows of the table focus on various subsamples.[36] These finer decompositions show that the extent to which trade expansion is affected by trade integration is strongly limited by geography and transportation costs, both across countries and industries. The contribution of trade integration is larger for countries that are separated by short distances (79 percent) or adjacent to each other (72 percent), and for goods that are easier and cheaper to transport, i.e., with a low *c.i.f./f.o.b.* value (74 percent) and a low weight-to-value ratio (63 percent).[37] The latter two findings are consistent with some recent evidence showing that the relative price of air to non-air shipping has declined over time (Hummels, 2007, Choate, 2008). Such a decrease in relative costs should encourage trade in air transported goods, which are indeed characterized by a low weight-to-value (or low *c.i.f./f.o.b.*). Finally, Table 6 shows that speaking a common language, which is of course correlated with short distances and adjacency, is also associated with a higher contribution of trade integration (76 percent).

The effectiveness of trade integration in promoting the growth in trade is also influenced by policy factors. On average, the contribution of trade integration is slightly stronger for Finland and Austria than for the other countries in the sample (64 percent against 56 percent).[38] This indicates that Finland and Austria, who only joined the EU in 1995, experienced a somewhat faster rate of integration. Interestingly, non-Eurozone countries only experienced a very small increase (equal to 0.32 percent) in trade with Eurozone countries since 2002, which is consistent with our earlier finding that the introduction of the Euro has hampered the integration of those countries. Moreover, changes in trade integration only had a minor effect on the trade growth of Euro outsiders (15 percent versus 79 percent for Eurozone countries). This indicates that in contrast to the Eurozone, the trade growth of Euro outsiders was predominantly driven by secular output growth, not by improvements in trade integration.[39]

We then repeat the exercise for subsamples in which none, one or both of the partners have implemented the Schengen Agreement. Consistent with our previous findings, the elimination of physical borders seems to have greatly fostered trade. Trade integration accounts for 76 percent of the growth in trade when both partners have implemented the Agreement but for only 21 percent when just one has joined. Trade integration has actually declined for the pairs in which none of the partners has implemented the Agreement so that the contribution is negative (-44 percent).

The pattern of the results is somewhat different once we focus on industry-specific characteristics such as TBTs and the transparency of public procurement. In all three cases, changes in trade integration have been most effective where trade barriers were initially higher, i.e., in industries affected by high TBTs (83 percent) and markets with a low transparency in public procurement (76 percent). This indicates that the process of EU integration has been to some degree successful in overcoming such barriers. Finally, trade expansion in highly productive industries is predominantly driven by changes in trade integration (88 percent), whereas changes in output are still the main driver of trade growth in low productivity industries.

---

[36] We do not report decompositions for VAT, capital shares or the number of goods in each sector as those are only included in our earlier regressions to control for possible measurement error.

[37] The sample is split at the mean value of the variable in question in 1999, denoted by † in Table 6.

[38] Given that those relative contributions are calculated and not estimated, we are not able to test whether the difference between these two numbers is significant.

[39] As we only focus on the post-2002 period, note that the sample size is much smaller than in the full sample.

5 Robustness

In this section we report a number of alternative specifications that we implement to ensure the robustness of our conclusions. First, we estimate our main specification by Poisson Pseudo-Maximum Likelihood (PPML) as recommended by Santos Silva and Tenreyro (2006). They show that under heteroskedasticity, coefficient estimates from log-linear regressions estimated by OLS can be strongly biased. Martin and Pham (2008) confirm that the PPML estimator solves the heteroskedasticity bias problem, but instead show that this estimator yields severely biased estimates when the dependent variable takes on many zero values. As our dependent variable has virtually no zero values but heteroskedasticity might be present, the PPML estimator appears adequate as a robustness check. But as Column (1) of Table A1 shows, there are no qualitative changes and generally only small quantitative changes in the estimation results compared to those obtained with OLS.

Second, we verify in Columns (2) and (3) that the results are robust to the inclusion of country pair fixed effects interacted with year dummies and to the exclusion of the series that contain zero trade observations.[40] We also want to make sure that our results are not driven by observations that are economically small. For instance, in Table 3 we find that the effect of remaining outside the Eurozone is significant. But the dummy that captures the effect is equal to unity in only 1,796 cases out of a total of 15,040 observations (i.e., 12 percent of the sample), and roughly half of those cases involve Denmark, one of the smallest economies in our sample. We therefore run a regression where the observations are weighted by the product of sectoral bilateral exports in the initial year 1999. The results, reported in Column (4), reveal that a significant Eurozone effect is still present.

With panel data, one issue relates to the computation of θ_{ij}^k in *real* terms. Ideally, for deflating we would need domestic price indices for the numerator and export price indices for the denominator in equation (14). Export price indices are not available, so common practice it to use domestic deflators instead (Baldwin and Taglioni, 2006). In that case, the deflators in (14) cancel out and using nominal or real variables yields the same values for θ_{ij}^k (which is the approach we have followed in the paper). Since there are reasons to believe that domestic and export price indices differ from each other, for instance in the case of pricing-to-market practices, we need to tackle this empirical problem up-front.

One way of doing this is to include a set of time-varying sector (3-digit) fixed effects, which leaves our results unaffected (Column 5). A second way is to run cross-sectional regressions for each year, as we do in Table A2. This approach comes at the cost of strongly reducing the number of observations for each regression so that not all variables are significant in all years, but overall the results largely hold up.[41] It is interesting that the magnitude of the coefficient on the non-Eurozone dummy substantially increases between 2002 and 2003, suggesting that the negative effect of not adopting the Euro might have strengthened over time. There is also some evidence that the importance of TBTs increased in the last three years. Finally, the coefficients on distance remain stable.[42]

---

[40] In the case of zero trade observations, we drop the entire series to keep a balanced sample over time. This results in 16 fewer observations in Column (3) as the five zero trade observations correspond to four different series only.

[41] Lagged productivity is omitted for 1999.

[42] The literature that examines the evolution of distance effects over time yields mixed results. Coe, Subramanian and Tamirisa (2007) find declining distance effects once they estimate a non-linear version of the gravity equation with an additive error term. In contrast, Berthelon and Freund (2008) find an increasing impact of distance as the effect of distance remains unchanged in 75 percent of their industries but increases in the remaining industries. For an overview, see Disdier and Head (2008).

In order to compare how θ_{ij}^k performs relative to alternative measures of trade integration, we use the log inverse *phi*-ness of trade as a dependent variable in Column (6) of Table A1. This implicitly assumes that the elasticity of substitution is equal to 1.5 for all industries so that the exponent in equation (14) becomes unity. Assuming a homogeneous elasticity of substitution across industries reduces the adjusted R^2 considerably by about twelve percentage points to 0.693. Moreover, although their signs remain mostly unchanged, the size of most estimated coefficients is radically different. In addition, not being a member of the Eurozone, TBTs and sharing a common border are no longer significantly associated with trade integration.[43]

6 Concluding Remarks

This paper explores the sources and size of trade impediments. Measuring these impediments directly is often difficult because of data limitations. Instead, we indirectly infer trade impediments from observable trade flows. For this purpose we use the gravity framework pioneered by Anderson and van Wincoop (2003) and apply it to industries with heterogeneous trade costs and heterogeneous elasticities of substitution. The model yields a micro-founded measure of bilateral trade integration that controls for time-varying multilateral resistance and that can be applied to disaggregated panel data. We show that this trade integration measure is also consistent with the Ricardian model by Eaton and Kortum (2002) and recent heterogeneous firms models.

From an empirical perspective, we use this micro-founded measure to explore the key determinants of trade integration across countries and industries in the European Union. We show that cross-country trade integration is lower for those countries in our sample that joined the EU most recently and that opted out of the Euro common currency, but it is higher for countries that have implemented the Schengen Agreement. We also document a considerable variation in trade integration across industries. Consistent with the literature on firm heterogeneity, we find that industries with higher productivity trade more on international markets. Trade integration is hampered by transportation costs as captured by industry-specific weight-to-value and *c.i.f./f.o.b.* ratios. Trade integration is also severely hampered by policy factors, in particular Technical Barriers to Trade and intransparent public procurement procedures. Our findings suggest for public policy that gains from improved international trade integration are possible through the elimination of TBTs and through more transparent advertising of public procurement contracts.

Moreover, we find that trade integration has improved on average between 1999 and 2003. We find that roughly 58 percent of the growth in trade can be accounted for by improvements in trade integration through a decrease in bilateral and multilateral trade barriers, while the rest can be attributed to increases in manufacturing output.

To conclude, our measure yields sensible systematic differences in trade integration across industries, suggesting that modeling trade costs as a "one-fits-all" impediment is clearly at odds with empirical evidence. Instead, when dealing with industry-level data it is important to allow for trade cost heterogeneity across industries.

---

[43] We also tried using the elasticities from Broda and Weinstein (2006) for the US in computing θ_{ij}^k. However, this led to large outliers in θ_{ij}^k and the overall results were weaker. The most likely reason is that those elasticities are much more finely disaggregated and the classifications are thus harder to match. The results are available upon request.

Ultimately, of course even highly disaggregated industry-level data “can never be as fine as reality, so some degree of aggregation bias is inevitable” (Anderson and van Wincoop, 2004, p.729). Thus, the use of industry-level data precludes us from exploring the determinants and change in trade impediments at the firm or product level. In our view, analyzing trade barriers at these very fine levels of disaggregation is a natural and important next step.

References

Anderson J.E. and E. van Wincoop, 2002, Borders, Trade, and Welfare, *Brookings Trade Forum.* In: Collins and Rodrik, Eds., Washington: Brookings Institution, 207-244.

Anderson J.E. and E. van Wincoop, 2003, Gravity with Gravitas: A Solution to the Border Puzzle, *American Economic Review* 93(1), 170-192.

Anderson J.E. and E. van Wincoop, 2004, Trade Costs, *Journal of Economic Literature* 42(3), 691-751.

Arkolakis K., Demidova S., Klenow P.J. and A. Rodríguez-Clare, 2008, Endogenous Variety and the Gains from Trade, *American Economic Review* 98(2), Papers & Proceedings, 444-450.

Baier S.L. and J.H. Bergstrand, 2001, The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity, *Journal of International Economics* 53(1), 1-27.

Baldwin R.E., 2000, Regulatory Protectionism, Developing Nations, and a Two-Tier World Trade System, *Brookings Trade Forum.* In: Collins and Rodrik, Eds., Washington: Brookings Institution, 237-293.

Baldwin R.E., 2006, The Euro's Trade Effects, *European Central Bank Working Paper* 594.

Baldwin R.E., Forslid R., Martin P., Ottaviano G.I.P. and F. Robert-Nicoud, 2003, The Core-Periphery Model: Key Features and Effects. In: Heijdra and Brakman, Eds., The Monopolistic Competition Revolution in Retrospect, Cambridge University Press, 213-235.

Baldwin R.E. and J. Harrigan, 2007, Zeros, Quality and Space: Trade Theory and Trade Evidence, *NBER Working Paper* 13214.

Baldwin R.E. and D. Taglioni, 2006, Gravity for Dummies and Dummies for Gravity Equations, *NBER Working Paper* 12516.

Bergstrand J., 1989, The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade, *Review of Economics and Statistics* 71(1), 143-153.

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

Bernard A.B., Jensen J.B. and P.K. Schott, 2006, Trade Costs, Firms and Productivity, *Journal of Monetary Economics* 53(5), 917-937.

Berthelon M. and C. Freund, 2008, On the Conservation of Distance in International Trade, *Journal of International Economics* 75(2), 310-320.

Besedeš T. and T.J. Prusa, 2007, The Role of Extensive and Intensive Margins and Export Growth, *NBER Working Paper* 13628.

Broda C. and D.E. Weinstein, 2006, Globalization and the Gains from Variety, *Quarterly Journal of Economics* 121(2), 541-585.

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

Chen N.A., 2004, Intra-National Versus International Trade in the European Union: Why Do National Borders Matter?, *Journal of International Economics* 63(1), 93-118.

Choate D.A., 2008, Relative Transport Price Movements and U.S. Imports, Columbia University, job market paper.

Coe D.T., Subramanian A. and N.T. Tamirisa, 2007, The Missing Globalization Puzzle: Evidence of the Declining Importance of Distance, *IMF Staff Papers* 54(1), 34-58.

Combes P.-P. and M. Lafourcade, 2005, Transport Costs: Measures, Determinants, and Regional Policy Implications for France, *Journal of Economic Geography* 5(3), 319-349.

Combes P.-P., Lafourcade M. and T. Mayer, 2005, The Trade-Creating Effects of Business and Social Networks: Evidence from France, *Journal of International Economics* 66(1), 1-29.

Costinot A. and I. Komunjer, 2007, What Goods Do Countries Trade? New Ricardian Predictions, *NBER Working Paper* 13691.

Davies S. and B. Lyons, 1996, Industrial Organization in the European Union – Structure, Strategy, and the Competitive Mechanism, Oxford University Press.

Deardorff A.V., 1998, Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?. In: The Regionalization of the World Economy, Frankel, Ed., Chicago: University of Chicago Press, 7-22.

Disdier A.C. and K. Head, 2008, The Puzzling Persistence of the Distance Effect on Bilateral Trade, *Review of Economics and Statistics* 90(1), 37-48.

Eaton J. and S. Kortum, 2002, Technology, Geography and Trade, *Econometrica* 70(5), 1741-1779.

European Commission, 1998, Technical Barriers to Trade. The Single Market Review, Subseries III: Dismantling of Barriers 1.

European Commission, 2000, EU Transport in Figures – Statistical Pocket Book, Directorate General for Energy and Transport.

European Commission, 2006, Internal Market: Opinions and Experiences of Businesses in EU-15, *Flash Eurobarometer 180*, TNS Sofres/EOS Gallup Europe.

European Commission, 2008, VAT Rates Applied in the Member States of the European Community, DOC/2412/2008 – EN.

Evans C.L., 2003, The Economic Significance of National Border Effects, *American Economic Review* 93(4), 1291-1312.

Evans C.L. and J. Harrigan, 2005, Distance, Time, and Specialization: Lean Retailing in General Equilibrium, *American Economic Review* 95(1), 292-313.

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

Feenstra R.C., Markusen J.R. and A.K. Rose, 2001, Using the Gravity Equation to Differentiate among Alternative Theories of Trade, *Canadian Journal of Economics* 34(2), 430-447.

Fields G.S., 2003, Accounting for Income Inequality and Its Change: A New Method, With Application to the Distribution of Earnings in the United States, *Research in Labor Economics* 22, 1-38.

Greenaway D. and R. Kneller, 2008, Exporting, Productivity and Agglomeration, *European Economic Review* 52(5), 919-939.

Grossman, G.M., 1998, Comment. In: The Regionalization of the World Economy, Frankel, Ed., Chicago: University of Chicago Press, 29-31.

Hanson G.H. and C. Xiang, 2004, The Home-Market Effect and Bilateral Trade Patterns, *American Economic Review* 94(4), 1108-1129.

Harrigan J., 1993, OECD Imports and Trade Barriers in 1983, *Journal of International Economics* 35(1/2), 91-111.

Harrigan J., 1996, Openness to Trade in Manufactures in the OECD, *Journal of International Economics* 40(1/2), 23-39.

Harrigan J., 2005, Airplanes and Comparative Advantage, *NBER Working Paper* 11688.

Head K. and T. Mayer, 2000, Non-Europe: The Magnitude and Causes of Market Fragmentation in the EU, *Weltwirtschaftliches Archiv* 136 (2), 285-314.

Head K. and T. Mayer, 2004, The Empirics of Agglomeration and Trade, *Handbook of Regional and Urban Economics*, Chapter 59, North-Holland.

Head K. and J. Ries, 2001, Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade, *American Economic Review* 91(4), 858-876.

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

Hillberry R. and D. Hummels, 2000, Explaining Home Bias in Consumption: Production Location, Commodity Composition and Magnification, Purdue University, mimeo.

Hummels D., 2001a, Toward a Geography of Trade Costs, Purdue University, mimeo.

Hummels D., 2001b, Time as a Trade Barrier, Purdue University, mimeo.

Hummels D., 2007, Transportation Costs and International Trade in the Second Era of Globalization, *Journal of Economic Perspectives* 21(3), 131-154.

Hummels D. and P.T. Klenow, 2005, The Variety and Quality of a Nation's Exports, *American Economic Review* 95(3), 704-723.

Jacks D.S., C.M. Meissner and D. Novy, 2008, Trade Costs, 1870-2000, *American Economic Review* 98(2), Papers & Proceedings, 529-534.

Limão N. and A.J. Venables, 2001, Infrastructure, Geographical Disadvantage and Transport Costs, World Bank Economic Review 15, 451-479.

Martin W. and C.S. Pham, 2008, Estimating the Gravity Model when Zero Trade Flows Are Frequent, mimeo.

McCallum J., 1995, National Borders Matter: Canada-US Regional Trade Patterns, *American Economic Review* 85(3), 615-623.

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

Melitz M.J. and G.I.P. Ottaviano, 2008, Market Size, Trade, and Productivity, *Review of Economic Studies* 75(1), 295-316.

Nitsch V., 2000, National Borders and International Trade: Evidence from the European Union, *Canadian Journal of Economics* 33(4), 1091–1105.

Novy D., 2008, Gravity Redux: Measuring International Trade Costs with Panel Data, University of Warwick, mimeo.

Rauch J.E., 1999, Networks Versus Markets in International Trade, *Journal of International Economics* 48 (1), 7– 35.

Rose A.K., 2000, One Money, One Market: Estimating the Effect of Common Currencies on Trade, *Economic Policy* 30, 7-45.

Santos Silva J.M.C. and S. Tenreyro, 2006, The Log of Gravity, *Review of Economics and Statistics* 88(4), 641–658.

Wei S.-J., 1996, Intra-National versus International Trade: How Stubborn Are Nations in Global Integration? *NBER Working Paper* 5531.

Wolf H.C., 2000, Intranational Home Bias in Trade, *Review of Economics and Statistics* 82(4), 555-563.

Table 1: Descriptive Statistics for Individual Industries

3-digit Nace rev.1 Industry	θ^k	min θ^k	max θ^k	σ_k	wv^k
Bricks	146.91	5.52	994.94	2.65	4.92
Cement, lime, plaster	109.88	23.21	480.73	2.52	13.36
Stone	54.67	8.29	552.31	2.65	2.14
Fruit, vegetables	41.58	6.11	387.65	2.46	1.03
Articles of concrete, plaster, cement	25.12	2.24	581.78	3.24	3.11
Ceramic tiles, flags	18.15	2.00	266.85	2.65	1.75
Sawmilling, planing of wood	14.75	2.33	106.51	3.36	2.28
Jewellery	14.00	1.95	72.17	3.48	0.01
Other food products	11.31	1.53	1,056.31	4.81	0.58
Wooden containers	10.27	2.46	87.53	3.99	2.14
Furniture	9.98	1.81	157.52	3.64	0.25
Ceramic goods	9.66	0.86	157.07	3.27	0.71
Prepared animal feeds	9.57	1.04	42.25	3.61	1.75
Glass	8.93	0.80	122.86	3.09	0.84
Builders' carpentry, joinery	8.76	2.74	48.50	3.99	0.51
Processing of iron, steel	6.90	1.71	30.51	3.53	1.10
Publishing	6.66	1.02	78.18	4.88	0.17
Printing	6.39	1.60	21.38	5.58	0.29
Other non-metallic mineral products	5.88	1.23	34.85	2.96	0.57
Structural metal products	5.84	2.27	37.71	4.79	0.36
Other products of wood	5.57	1.79	19.47	4.13	0.38
Tanks, reservoirs	4.48	1.40	22.47	4.76	0.29
Tubes	4.30	1.73	21.12	3.53	0.94
Weapons, ammunition	4.09	2.35	8.80	4.88	0.13
Paper, paperboard	4.00	1.37	12.18	4.64	0.37
Leather clothes	4.00	2.97	5.20	5.66	0.02
Rubber products	3.90	0.80	10.21	3.88	0.26
Fish	3.73	1.35	12.61	4.71	0.33
Pulp, paper, paperboard	3.50	1.36	12.60	4.28	1.71
Other fabricated metal products	3.30	1.24	17.52	4.89	0.39
Wearing apparel	3.02	0.70	15.42	5.66	0.11
Plastic products	2.95	1.29	11.41	5.36	0.28
Miscellaneous manufacturing	2.81	1.41	8.01	4.99	0.13
Grain mill products	2.79	1.65	8.76	5.05	1.69
Musical instruments	2.75	1.47	5.09	4.88	0.05
Cutlery	2.65	1.17	7.08	4.85	0.10
Steam generators	2.59	1.60	6.08	7.87	0.20
Ships, boats	2.57	1.27	4.70	7.40	0.27
Dairy products	2.47	1.57	5.60	6.77	0.52
Luggage, handbags	2.45	1.41	4.20	6.18	0.07

(continued on next page)

Table 1 (continued)

3-digit Nace rev.1 Industry	θ^k	min θ^k	max θ^k	σ_k	wv^k
Soap, detergents	2.30	1.43	10.55	5.74	0.60
Other transport equipment	2.30	1.39	5.05	7.11	0.27
Tobacco	2.29	2.02	2.54	6.62	0.26
Paints, varnishes	2.27	1.41	4.93	6.37	0.44
Basic chemicals	2.22	1.07	6.40	6.09	0.87
Knitted, crocheted articles	2.20	1.12	5.77	6.90	0.04
Electrical equipment	2.14	1.13	4.70	5.94	0.06
Lighting equipment	2.11	1.19	4.26	5.17	0.09
Railway locomotives	2.10	1.31	4.74	7.40	0.13
Insulated wire, cable	2.10	1.26	3.27	5.88	0.19
Knitted, crocheted fabrics	2.07	1.51	3.35	6.90	0.10
Other chemical products	2.04	1.00	6.01	6.46	0.30
Electricity distribution apparatus	2.01	1.33	2.80	5.88	0.04
Sports goods	2.00	0.68	3.79	4.88	0.12
Domestic appliances	1.94	1.04	4.05	5.75	0.18
Made-up textile articles	1.89	1.12	2.93	7.46	0.11
Man-made fibres	1.87	1.41	2.51	6.59	0.42
Footwear	1.86	1.12	3.25	7.22	0.04
Medical equipment	1.84	1.43	2.63	6.00	0.03
Other general purpose machinery	1.84	1.25	4.28	7.00	0.13
Agricultural, forestry machinery	1.83	1.19	3.03	8.36	0.17
Parts for motor vehicles	1.75	1.06	4.50	7.28	0.17
Electronic valves, tubes	1.74	1.04	2.84	5.88	0.02
Other textiles	1.68	0.96	3.13	7.82	0.17
Other special purpose machinery	1.66	1.09	3.25	8.17	0.08
Pesticides	1.66	1.22	2.40	6.75	0.28
Tanning, dressing of leather	1.62	1.14	3.35	8.92	0.08
Pharmaceuticals	1.62	1.11	3.39	9.05	0.05
Machinery	1.57	0.88	2.86	7.21	0.09
Electronic motors	1.56	1.24	2.06	7.02	0.09
Dressing, dyeing of fur	1.53	1.19	1.97	8.09	0.02
Motorcycles, bicycles	1.52	1.08	1.93	7.11	0.10
Optical instruments	1.52	1.21	2.10	7.70	0.01
Accumulators	1.48	1.12	1.87	5.88	0.32
Games, toys	1.41	1.11	1.72	4.88	0.10
Motor vehicles	1.40	1.21	1.61	7.25	0.12
Aircraft, spacecraft	1.37	1.21	1.60	7.55	0.00
Office machinery, computers	1.15	1.04	1.24	10.94	0.02

Notes: Authors' calculations. For θ^k, σ_k and wv^k, the numbers reported are averages for each 3-digit industry. wv^k is in units of kg/Euro.

Table 2: Descriptive Statistics for Individual Countries

	θ_i	min θ_i	max θ_i
Ireland	9.47	1.208	261.49
Spain	9.02	0.677	1056.30
Denmark	8.80	1.138	782.81
Finland	8.61	1.195	552.31
Italy	7.75	0.701	1,056.30
United Kingdom	7.73	0.999	994.94
Austria	6.69	0.927	154.63
Portugal	6.50	0.677	414.15
Germany	5.67	0.879	480.73
Netherlands	5.48	0.970	123.36
France	5.28	0.701	414.15

Notes: Authors' calculations. The numbers reported for θ_i are averages for each country across industries.

Table 3: The Determinants of EU Trade Integration

	(1)	(2)	(3)	(4)	(5)	(6)
Geography/Transport Costs						
$\ln D_{ij}$	0.583^{a}	0.515^{a}	0.530^{a}	0.483^{a}	0.495^{a}	0.506^{a}
	(28.378)	(23.296)	(24.799)	(21.420)	(22.074)	(22.806)
$\ln(D_{ii} \times D_{jj})$	-0.930^{a}	-0.945^{a}	-0.990^{a}	-0.923^{a}	-0.928^{a}	-0.969^{a}
	(-11.973)	(-11.381)	(-12.247)	(-11.323)	(-11.672)	(-12.354)
Adj_{ij}	-0.063^{a}	-0.051^{b}	-0.044^{c}	-0.068^{a}	-0.061^{a}	-0.055^{b}
	(-2.810)	(-2.218)	(-1.939)	(-2.917)	(-2.625)	(-2.428)
$Lang_{ij}$	-0.201^{a}	-0.355^{a}	-0.353^{a}	-0.363^{a}	-0.369^{a}	-0.368^{a}
	(-4.854)	(-8.109)	(-7.995)	(-8.186)	(-8.432)	(-8.331)
$\ln cfob_t^k$	0.030^{a}	0.030^{a}	0.031^{a}	0.030^{a}	0.015^{a}	0.016^{a}
	(5.464)	(6.152)	(6.294)	(6.146)	(3.291)	(3.400)
$\ln wv_t^k$	0.719^{a}	0.631^{a}	0.642^{a}	0.648^{a}	0.664^{a}	0.673^{a}
	(5.973)	(6.414)	(6.459)	(6.698)	(6.833)	(6.854)
Policy Variables						
FI, AT_{ij}	–	0.199^{a}	0.187^{a}	0.233^{a}	0.228^{a}	0.220^{a}
		(11.215)	(10.666)	(12.376)	(12.331)	(11.851)
$noEURO_{ij,t}$	–	0.042^{a}	0.032^{b}	0.040^{a}	0.036^{a}	0.027^{b}
		(3.108)	(2.388)	(2.906)	(2.704)	(2.072)
$Schengen_{ij,t}$	–	-0.137^{a}	-0.118^{a}	-0.135^{a}	-0.132^{a}	-0.113^{a}
		(-5.443)	(-4.774)	(-4.914)	(-4.954)	(-4.302)
$\ln TBT_{ij}^k$	–	0.142^{a}	–	0.146^{a}	0.131^{b}	–
		(2.822)		(2.721)	(2.434)	
TBT^k	–	–	0.057^{b}	–	–	0.056^{b}
			(2.156)			(1.967)
$\ln Proc_{ij,t}^k$	–	-0.573^{c}	-0.528^{c}	-0.963^{a}	-0.917^{b}	-0.892^{b}
		(-1.886)	(-1.738)	(-2.729)	(-2.564)	(-2.495)
$\ln VAT_{ij}^k$	–	-2.282^{a}	-2.290^{a}	-2.259^{a}	-1.910^{a}	-1.931^{a}
		(-5.741)	(-5.744)	(-5.693)	(-5.122)	(-5.138)
Other Costs						
$\ln Prod_{ij,t-1}^k$	–	–	–	-0.143^{a}	-0.144^{a}	-0.156^{a}
				(-4.453)	(-4.581)	(-5.120)
$Zeros_{ij,t}^k$	–	–	–	2.277^{a}	2.210^{a}	2.193^{a}
				(9.913)	(9.849)	(9.511)
Controls						
$KS_{ij,t}^k$	–	–	–	–	0.054	0.063
					(1.465)	(1.637)
$\ln Goods^k$	–	–	–	–	-0.112^{a}	-0.112^{a}
					(-9.653)	(-9.633)
N	15,040	15,040	15,040	12,032	12,032	12,032
Adj-R^2	0.791	0.802	0.801	0.804	0.810	0.810

Notes: The dependent variable is $\ln \theta_{ij,t}^k$. Robust standard errors are adjusted for clustering at the 4-digit Nace rev.1 level in each country pair (3,008 clusters). Year and 3-digit industry fixed effects are included in all regressions. The sample period is 1999-2003. t-statistics in parentheses. Constant terms are included but not reported. ^{a}, ^{b} and ^{c} indicate significance at 1, 5 and 10 percent levels, respectively.

Table 4: Variance Decomposition

	(1)	(2)
Geography/Transport Costs	**24.6%**	**25.0%**
$\ln D_{ij}$	5.4%	5.6%
$\ln\left(D_{ii} \times D_{jj}\right)$	0.7%	0.8%
Adj_{ij}	0.6%	0.5%
$Lang_{ij}$	0.5%	0.5%
$\ln cfob_t^k$	0.7%	0.7%
$\ln wv_t^k$	16.7%	16.9%
Policy Variables	**7.6%**	**6.8%**
FI, AT_{ij}	1.4%	1.4%
$noEURO_{ij,t}$	0.01%	0.01%
$Schengen_{ij,t}$	0.2%	0.2%
$\ln TBT_{ij}^k$	4.5%	–
TBT^k	–	3.7%
$\ln Proc_{ij,t}^k$	0.6%	0.6%
$\ln VAT_{ij}^k$	0.8%	0.8%
Other Costs	**0.3%**	**0.4%**
$\ln Prod_{ij,t-1}^k$	0.1%	0.1%
$Zeros_{ij,t}^k$	0.3%	0.3%
Controls	**5.6%**	**5.5%**
$KS_{ij,t}^k$	0.1%	0.1%
$\ln Goods^k$	5.4%	5.4%
3-digit industry fixed effects	42.8%	43.2%
Year fixed effects	0.01%	0.01%
Variation explained	80.8%	80.8%
Residual	19.2%	19.2%
Sum	100%	100%
N	12,032	12,032

Notes: The variance decompositions are calculated according to Fields (2003). The contribution of each explanatory variable x_m to the total variance of $\theta_{ij,t}^k$ is given by $c_m = \beta_m cov\left(x_m, \theta_{ij,t}^k\right)/var\left(\theta_{ij,t}^k\right)$ where β_m is the partial regression coefficient of $\theta_{ij,t}^k$ on the explanatory variable x_m (holding all other explanatory variables constant). The decompositions in Columns (1) and (2) correspond to regressions (5) and (6) in Table 3. The contributions sum to 100 percent.

Table 5: The Gravity Equation

"Naive" Gravity	(1)	(2)	(3)	(4)
$\ln D_{ij}$	-3.589^a (-31.563)	-3.741^a (-46.554)	-3.629^a (-30.342)	-3.757^a (-45.862)
$\ln\left(y_i^k y_j^k\right)$	1.226^a (57.064)	1.525^a (68.313)	–	–
$\ln y_i^k$	–	–	1.271^a (30.217)	1.546^a (43.766)
$\ln y_j^k$	–	–	1.184^a (30.617)	1.507^a (53.293)
Distance coefficients	Same $\forall k$	k-specific	Same $\forall k$	k-specific
Adj-R^2	0.602	0.849	0.602	0.849

Micro-Founded Gravity	(1)	(2)	(3)	(4)
$\ln D_{ij}$	-4.514^a (-34.104)	-5.131^a (-52.663)	-4.541^a (-32.669)	-5.045^a (-53.067)
$\ln\left(D_{ii} \times D_{jj}\right)$	11.551^a (19.551)	7.584^a (17.071)	–	–
$\ln D_{ii}$	–	–	2.022^a (13.343)	1.113^a (6.612)
$\ln D_{jj}$	–	–	2.223^a (12.676)	1.659^a (12.387)
$\ln\left(x_{ii}^k x_{jj}^k\right)$	0.750^a (31.486)	1.010^a (41.879)	–	–
$\ln x_{ii}^k$	–	–	0.789^a (16.522)	1.048^a (27.501)
$\ln x_{jj}^k$	–	–	0.712^a (15.942)	0.967^a (28.263)
Distance coefficients	Same $\forall k$	k-specific	Same $\forall k$	k-specific
Adj-R^2	0.510	0.842	0.510	0.852

Micro-Founded Gravity with FE	(1)	(2)	(3)	(4)
$\ln D_{ij}$	-4.204^a (-66.739)	-4.653^a (-15.537)	–	–
$\ln\left(y_i^k y_j^k\right)$	1.388^a (39.887)	1.462^a (40.914)	–	–
Distance coefficients	Same $\forall k$	k-specific	–	–
Adj-R^2	0.581	0.639	–	–

Notes: The upper, middle and lower panels refer to the naive and the two micro-founded gravity equations, estimated according to (17), (18) and (19) in the text. In Columns (1) and (3), the coefficients on bilateral and domestic distances are constrained to be the same across industries. In Columns (2) and (4), they are allowed to vary across industries, in which case we report the average of the elasticities estimated across industries. In all cases the number of observations is equal to 15,040. The sample period is 1999-2003. Constant terms are included but not reported. t-statistics in parentheses. ^{a}, ^{b} and ^{c} indicate significance at 1, 5 and 10 percent levels, respectively.

Table 6: Decomposition of the Growth in Trade

	Trade growth 1999-2003, average, in %	**Factor (1) Contribution of output growth**	**Factor (2) Contribution of trade integration**	N
Full sample	7.96	**42%**	**58%**	12,032
Geography/Transport Costs				
†Long D_{ij}	9.13	**63%**	**37%**	6,452
†Short D_{ij}	6.78	**21%**	**79%**	5,580
$Adj_{ij} = 0$	7.65	**47%**	**53%**	9,300
$Adj_{ij} = 1$	8.91	**28%**	**72%**	2,732
†High $cfob^{k}_{1999}$	7.50	**70%**	**30%**	4,704
†Low $cfob^{k}_{1999}$	8.24	**26%**	**74%**	7,328
†High wv^{k}_{1999}	7.07	**54%**	**46%**	3,744
†Low wv^{k}_{1999}	8.36	**37%**	**63%**	8,288
$Lang_{ij} = 0$	8.06	**43%**	**57%**	11,704
$Lang_{ij} = 1$	4.83	**24%**	**76%**	328
Policy Variables				
FI, AT_{ij}	9.06	**36%**	**64%**	3,432
Rest of sample	7.58	**44%**	**56%**	8,600
$noEURO_{ij}$ (from 2002)	0.32	**85%**	**15%**	1,796
$EURO_{ij}$ (from 2002)	9.98	**21%**	**79%**	4,132
$Schengen_{ij,t} = 0$	6.56	**144%**	**-44%**	270
$Schengen_{ij,t} = 0.5$	6.89	**79%**	**21%**	3,455
$Schengen_{ij,t} = 1$	8.45	**24%**	**76%**	8,307
†High TBT^{k}_{ij}	7.41	**17%**	**83%**	4,232
†Low TBT^{k}_{ij}	8.25	**55%**	**45%**	7,800
†Low $Proc_{ij,1999}$	7.10	**24%**	**76%**	5,812
†High $Proc_{ij,1999}$	8.79	**60%**	**40%**	6,220
Productivity				
†Low $Prod^{k}_{ij,1999}$	6.88	**76%**	**24%**	5,848
†High $Prod^{k}_{ij,1999}$	8.93	**12%**	**88%**	6,184

Notes: Decompositions are obtained according to equation (23) in the text. The numbers reported for trade growth and for factors (1) and (2) are weighted averages of the individual observations for each country pair and industry where the weights are the products of sectoral bilateral exports in the initial year 1999. † denotes that the sample is split at the mean value of the variable in question in the initial year 1999. Mean values are equal to $D_{ij} = 1,161$ kilometers, $cfob^{k}_{1999} = 4.43$, $wv^{k}_{1999} = 1.41$ kg/Euro, $TBT^{k}_{ij} = 3.65$, $Proc_{ij,1999} = 1.12$ and $Prod^{k}_{ij,1999} = 41,060$ Euros/employee.

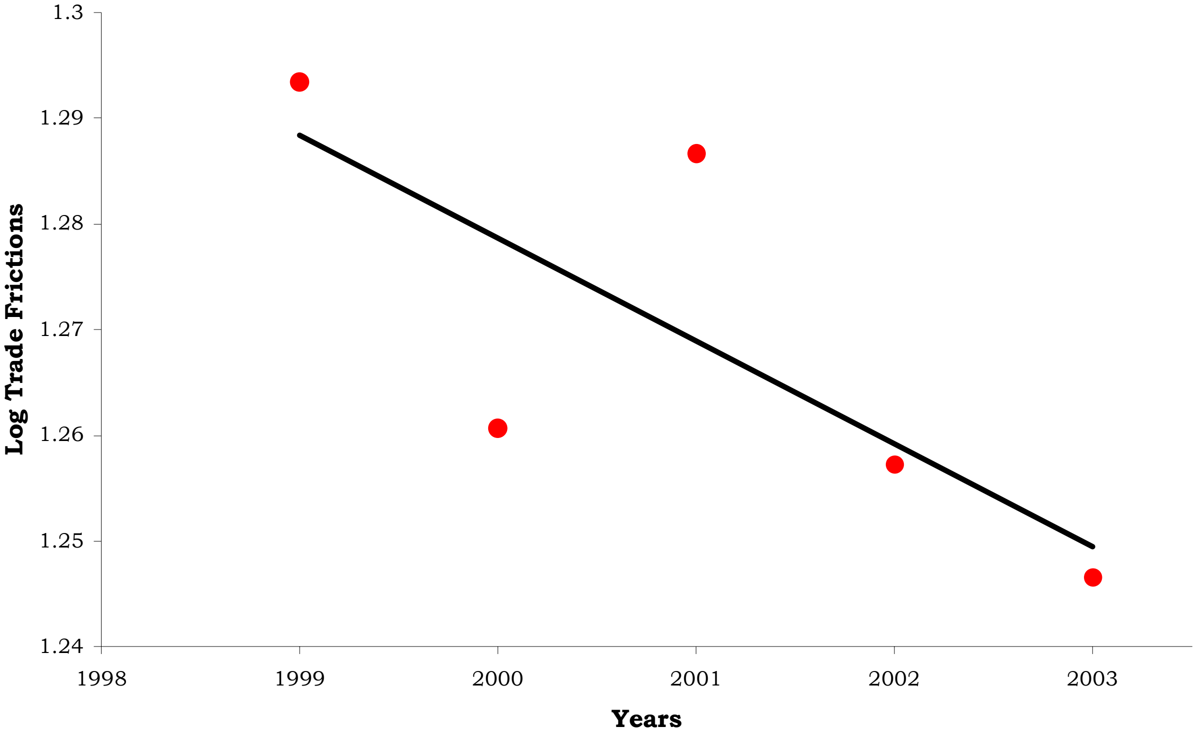

Figure 1: Trade frictions $\ln \theta_t$, 1999-2003, average across countries and industries.

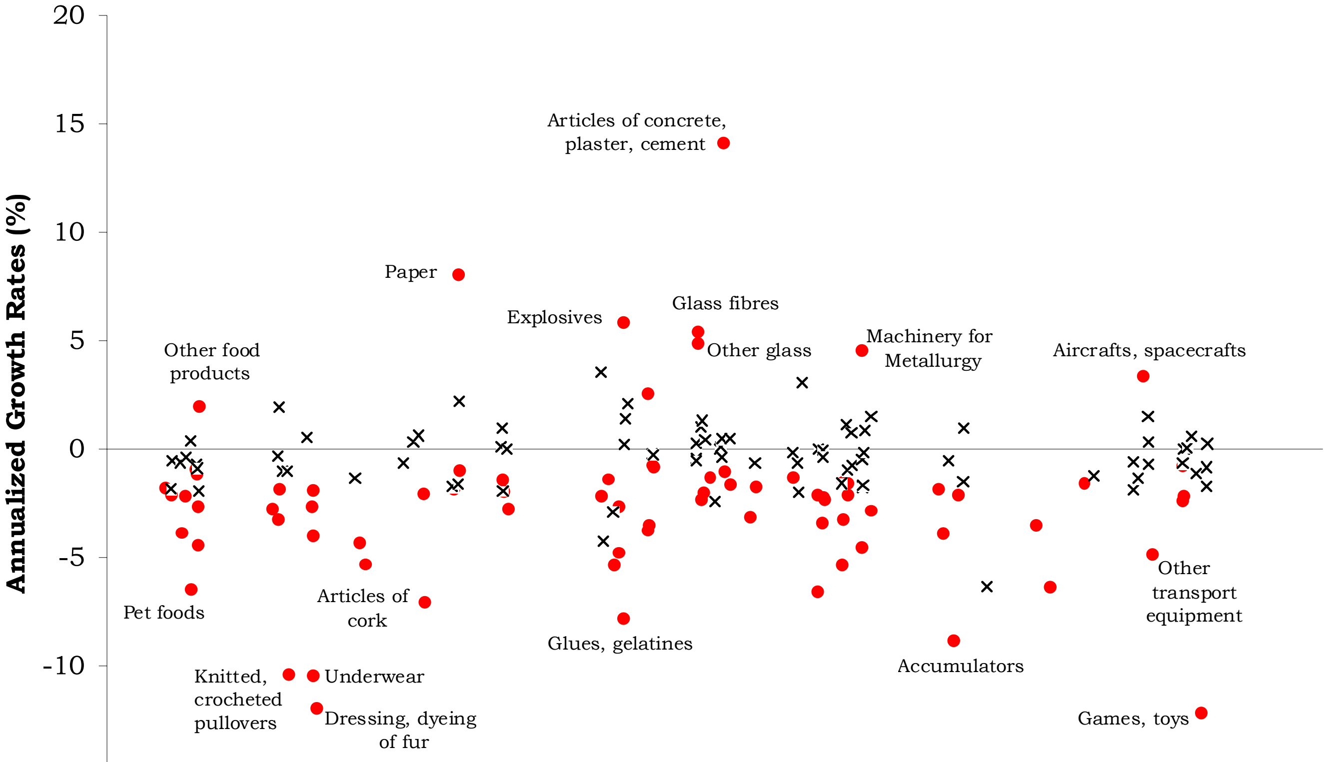

Figure 2: Annualized growth rates per industry, 1999-2003, see regression (20). ● indicates significance at the 10 percent level. × indicates insignificantly different from zero.

Appendix A

Table A1: The Determinants of EU Trade Integration – Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
Geography/Transport Costs						
$\ln D_{ij}$	0.786^a	–	0.492^a	0.465^a	0.495^a	3.785^a
	(7.221)		(21.932)	(22.586)	(21.828)	(28.313)
$\ln (D_{ii} \times D_{jj})$	-1.112^a	–	-0.925^a	-0.873^a	-0.927^a	-7.779^a
	(−3.151)		(−11.642)	(−11.970)	(−11.556)	(−16.284)
Adj_{ij}	-0.449^a	–	-0.063^a	-0.060^a	-0.061^a	−0.097
	(−3.629)		(−2.704)	(−2.852)	(−2.603)	(−0.811)
$Lang_{ij}$	−0.159	–	-0.368^a	-0.357^a	-0.369^a	-2.590^a
	(−1.145)		(−8.423)	(−8.612)	(−8.363)	(−11.979)
$\ln cfob_t^k$	0.067^a	0.017^a	0.015^a	0.013^a	0.018^a	0.150^a
	(3.223)	(3.596)	(3.174)	(3.106)	(2.615)	(5.350)
$\ln wv_t^k$	0.606^a	0.670^a	0.665^a	0.648^a	0.724^a	1.961^a
	(3.814)	(7.007)	(6.854)	(7.255)	(6.826)	(6.120)
Policy Variables						
FI, AT_{ij}	0.406^a	–	0.227^a	0.215^a	0.229^a	1.827^a
	(5.146)		(12.253)	(12.390)	(12.212)	(16.812)
$noEURO_{ij,t}$	0.248^a	–	0.034^b	0.032^a	0.037^a	0.077
	(3.212)		(2.549)	(2.675)	(2.628)	(0.950)
$Schengen_{ij,t}$	-0.449^a	–	-0.130^a	-0.119^a	-0.134^a	-0.577^a
	(−3.265)		(−4.928)	(−4.930)	(−4.937)	(−3.696)
$\ln TBT_{ij}^k$	1.110^a	0.105^c	0.133^b	0.127^a	0.125^b	−0.057
	(4.430)	(1.876)	(2.473)	(2.664)	(2.346)	(−0.430)
$\ln Proc_{ij,t}^k$	-3.320^a	-1.377^a	-0.899^b	-0.779^b	-1.279^a	-5.717^b
	(−3.245)	(−3.789)	(−2.521)	(−2.409)	(−2.752)	(−1.967)
$\ln VAT_{ij}^k$	-4.103^b	-2.520^a	-1.848^a	-1.502^a	-1.917^a	-13.274^a
	(−2.468)	(−5.573)	(−4.979)	(−4.279)	(−5.063)	(−5.579)
Other Costs						
$\ln Prod_{ij,t-1}^k$	-0.297^a	-0.109^a	-0.144^a	-0.134^a	-0.152^a	-1.638^a
	(−2.784)	(−2.615)	(−4.574)	(−4.501)	(−4.720)	(−9.872)
$Zeros_{ij,t}^k$	2.145^a	2.249^a	–	2.287^a	2.214^a	13.245^a
	(13.827)	(12.775)		(9.872)	(9.853)	(11.914)
Controls						
$KS_{ij,t}^k$	0.318^c	0.042	0.054	0.055	0.055	0.280
	(1.740)	(1.027)	(1.459)	(1.610)	(1.467)	(1.599)
$\ln Goods^k$	-0.181^a	-0.112^a	-0.110^a	-0.102^a	-0.111^a	-1.205^a
	(−4.381)	(−9.888)	(−9.487)	(−9.680)	(−9.465)	(−15.924)
Estimator	PPML	OLS	OLS	OLS	OLS	OLS
Fixed Effects	t, K	$t \times ij, K$	t, K	t, K	$t \times K$	t, K
Weighted	No	No	No	Yes	No	No
Sample	Full	Full	Excl. zeros	Full	Full	Full
N	12,032	12,032	12,016	12,032	12,032	12,032
Adj-R^2	–	0.821	0.810	0.812	0.812	0.693

Notes: In (1) the dependent variable is $\theta_{ij,t}^k$ and in (2) to (5) the dependent variable is $\ln \theta_{ij,t}^k$. In (6), the dependent variable is the (log) inverse *phi*-ness of trade. Poisson Pseudo-Maximum Likelihood (PPML) estimation in (1); fixed effects OLS regressions in (2) to (6). Robust standard errors are adjusted for clustering at the 4-digit Nace rev.1 level in each country pair (3,008 clusters). The sample period is 1999-2003. Weighted regression in (4) where the weights are the product of sectoral bilateral exports in the initial year 1999. t-statistics in parentheses. Constant terms are included but not reported. ^{a}, ^{b} and ^{c} indicate significance at 1, 5 and 10 percent levels, respectively.

Table A2: The Determinants of EU Trade Integration: Cross-sectional samples

	1999	2000	2001	2002	2003
Geography/Transport Costs					
$\ln D_{ij}$	0.515^{a}	0.475^{a}	0.515^{a}	0.493^{a}	0.498^{a}
	(22.108)	(19.175)	(20.834)	(20.453)	(20.306)
$\ln(D_{ii} \times D_{jj})$	-0.857^{a}	-0.955^{a}	-0.988^{a}	-0.906^{a}	-0.836^{a}
	(−10.064)	(−10.932)	(−10.942)	(−11.068)	(−9.928)
Adj_{ij}	-0.050^{b}	-0.068^{a}	-0.057^{b}	-0.054^{b}	-0.056^{b}
	(−2.146)	(−2.860)	(−2.225)	(−2.183)	(−2.180)
$Lang_{ij}$	-0.351^{a}	-0.357^{a}	-0.376^{a}	-0.383^{a}	-0.377^{a}
	(−7.436)	(−7.970)	(−8.145)	(−8.165)	(−7.701)
$\ln cfob_{t}^{k}$	0.031^{b}	0.015	0.004	0.002	0.040^{a}
	(2.109)	(1.347)	(0.301)	(0.204)	(3.292)
$\ln wv_{t}^{k}$	0.793^{a}	0.928^{a}	0.743^{a}	0.733^{a}	0.541^{a}
	(7.132)	(7.775)	(6.202)	(6.011)	(4.804)
Policy Variables					
FI, AT_{ij}	0.179^{a}	0.219^{a}	0.224^{a}	0.237^{a}	0.246^{a}
	(9.504)	(11.122)	(10.800)	(11.875)	(12.257)
$noEURO_{ij,t}$	–	–	–	0.029	0.061^{a}
				(1.353)	(2.690)
$Schengen_{ij,t}$	-0.113^{a}	-0.123^{a}	-0.161^{a}	-0.150^{a}	-0.134^{a}
	(−3.751)	(−4.136)	(−4.268)	(−3.361)	(−2.907)
$\ln TBT_{ij}^{k}$	0.098^{b}	0.059	0.125^{b}	0.142^{a}	0.170^{a}
	(2.393)	(1.572)	(2.162)	(2.595)	(2.589)
$\ln Proc_{ij,t}^{k}$	−0.113	−0.728	-1.243^{c}	−0.752	-1.866^{a}
	(−0.128)	(−1.342)	(−1.925)	(−1.034)	(−4.520)
$\ln VAT_{ij}^{k}$	-1.817^{a}	-2.269^{a}	-1.847^{a}	-1.407^{a}	-2.154^{a}
	(−4.604)	(−5.226)	(−4.778)	(−3.461)	(−4.503)
Other Costs					
$\ln Prod_{ij,t-1}^{k}$	–	-0.198^{a}	-0.144^{a}	-0.108^{a}	-0.163^{a}
		(−5.622)	(−4.102)	(−3.008)	(−4.249)
$Zeros_{ij,t}^{k}$	–	–	–	2.239^{a}	2.205^{a}
				(7.733)	(11.095)
Controls					
$KS_{ij,t}^{k}$	−0.060	−0.044	0.076^{c}	0.083^{c}	0.137^{b}
	(−1.269)	(−1.005)	(1.744)	(1.827)	(2.412)
$\ln Goods^{k}$	-0.096^{a}	-0.114^{a}	-0.121^{a}	-0.111^{a}	-0.098^{a}
	(−7.040)	(−8.656)	(−9.038)	(−9.161)	(−7.577)
N	3,008	3,008	3,008	3,008	3,008
Adj-R^2	0.818	0.814	0.811	0.817	0.813

Notes: The dependent variable is $\ln \theta_{ij,t}^{k}$. Robust standard errors are adjusted for clustering at the 4-digit Nace rev.1 level in each country pair (3,008 clusters). 3-digit industry fixed effects are included in all regressions. The sample period is 1999-2003. t-statistics in parentheses. Constant terms are included but not reported. ^{a}, ^{b} and ^{c} indicate significance at 1, 5 and 10 percent levels, respectively.

Appendix B: Technical Appendix

As discussed in Section 2.4, our approach is consistent with a number of leading trade theories. This appendix provides the details. For simplicity we drop the industry index k so that we can write the trade integration measure in equation (14) as

$$\theta_{ij} = \left(\frac{x_{ii}x_{jj}}{x_{ij}x_{ji}}\right)^{\frac{1}{2(\sigma-1)}} = \left(\frac{t_{ij}t_{ji}}{t_{ii}t_{jj}}\right)^{\frac{1}{2}} \tag{24}$$

We use the notation from our model when referring to expressions from the papers below.

Eaton and Kortum (2002) In the Ricardian model by Eaton and Kortum (2002), productivity in each country is drawn from a Fréchet distribution that has two parameters, T_i and ζ. T_i determines the location of the productivity distribution for country i, with a high T_i denoting high overall productivity. $\zeta > 1$ denotes the variation of productivity across goods and is treated as common across countries, with a high ζ denoting little variation. Costinot and Komunjer (2007) generalize the model to multiple industries. The model yields a gravity equation for an aggregate of homogeneous goods whose structure is similar to equation (7). It is given by

$$\frac{x_{ij}}{x_j} = \frac{T_i\left(c_i t_{ij}\right)^{-\zeta}}{\sum_{i=1}^{J} T_i\left(c_i t_{ij}\right)^{-\zeta}}$$

where c_i denotes the input cost in country i. We are interested in the trade cost parameters. T_i and c_i are unobservable but cancel out once the ratio of domestic over bilateral trade flows is formed as in equation (24). This yields

$$\theta_{ij}^{EK} = \left(\frac{x_{ii}x_{jj}}{x_{ij}x_{ji}}\right)^{\frac{1}{2\zeta}} = \left(\frac{t_{ij}t_{ji}}{t_{ii}t_{jj}}\right)^{\frac{1}{2}} \tag{25}$$

Comparing equations (24) and (25), it is obvious that $\theta_{ij}^{EK} = \theta_{ij}$ if $\zeta = \sigma - 1$. For more details on the comparison of Armington-type and Ricardian models, see Eaton and Kortum (2002, footnote 20) and Anderson and van Wincoop (2004, pp. 709-710).

Chaney (2008) Chaney (2008) builds on the seminal paper by Melitz (2003) and derives a gravity equation based on a model with heterogeneous productivity across firms and fixed costs of exporting. In contrast to standard trade models, the two assumptions of heterogeneous firms and fixed costs of exporting introduce an extensive margin of trade. Exporters do not only vary the size of shipments in response to trade cost changes (the intensive margin), but the set of exporters also changes (the extensive margin). Chaney derives the following industry-level gravity equation

$$x_{ij} = \mu\frac{y_i y_j}{y}\left(\frac{w_i t_{ij}}{\lambda_j}\right)^{-\gamma}\left(f_{ij}\right)^{-\left(\frac{\gamma}{\sigma-1}-1\right)}$$

where μ is the weight of the industry in the consumers' utility function, y_i is total income of country i, y is world income, w_i is workers' productivity in country i, λ_j is a remoteness variable akin to multilateral resistance and f_{ij} are the fixed costs of exporting from country i to j. γ is the shape parameter of the Pareto distribution, with a high γ denoting a low degree of heterogeneity and $\gamma > \sigma - 1$. Forming the ratio of domestic over bilateral trade flows yields

$$\theta_{ij}^{Ch} = \left(\frac{x_{ii}x_{jj}}{x_{ij}x_{ji}}\right)^{\frac{1}{2\gamma}} = \left(\frac{t_{ij}t_{ji}}{t_{ii}t_{jj}}\right)^{\frac{1}{2}}\left(\frac{f_{ij}f_{ji}}{f_{ii}f_{jj}}\right)^{\frac{1}{2}\left(\frac{1}{\sigma-1}-\frac{1}{\gamma}\right)} \tag{26}$$

The trade integration measure θ_{ij}^{Ch} is a function of both variable and fixed trade costs. For details on welfare implications of the above models, see Arkolakis, Demidova, Klenow and Rodríguez-Clare (2008).

Melitz and Ottaviano (2008) Melitz and Ottaviano (2008) also model heterogeneous firms. But in contrast to Melitz (2003) and Chaney (2008), firms face fixed costs of market entry f_E that can be interpreted as product development and production start-up costs. When exporting, the firms only face variable trade costs and no fixed costs of exporting. The model is based on non-CES preferences that give rise to endogenous markups. More specifically, markups tend to be low in large markets with many competitors. Their multiple country model leads to the following gravity equation

$$x_{ij} = \frac{1}{2\delta(\gamma+2)} N_E^i \psi^i L^j \left(c_d^j\right)^{\gamma+2} (t_{ij})^{-\gamma}$$

where δ is a parameter from the utility function that indicates the degree of product differentiation (a higher δ means a higher degree of differentiation). N_E^i is the number of entrants in country i. ψ^i is an index of comparative advantage in technology. A high ψ^i means that entrants in country i have a high chance of obtaining good productivity draws. L^j denotes the number of consumers in country j. c_d^j is the marginal cost cut-off above which domestic firms in country j do not produce. The intuition is that tougher competition in country j, reflected by a lower c_d^j, makes it harder for exporters from i to break into that market. Forming the ratio of domestic over bilateral trade flows yields

$$\theta_{ij}^{MO} = \left(\frac{x_{ii}x_{jj}}{x_{ij}x_{ji}}\right)^{\frac{1}{2\gamma}} = \left(\frac{t_{ij}t_{ji}}{t_{ii}t_{jj}}\right)^{\frac{1}{2}} \tag{27}$$

Fixed costs do not enter the trade integration measure θ_{ij}^{MO} because all firms face identical entry costs f_E and no fixed costs of exporting. Variable trade costs are sufficient to induce selection into export markets because of bounded non-CES marginal utility.

Summary The four measures θ_{ij}, θ_{ij}^{EK}, θ_{ij}^{Ch} and θ_{ij}^{MO} have in common that they scale the ratio of domestic over bilateral trade flows by parameters that indicate a particular form of heterogeneity. A low σ in equation (24) indicates a high degree of differentiation across products; a low ζ in equation (25) indicates a high variation of productivity across countries; and a low γ in equations (26) and (27) indicates a high degree of firm heterogeneity. For given trade flows, higher heterogeneity in equations (24)-(27) implies higher trade frictions. Intuitively, suppose we observe two industries with the same ratio of domestic over bilateral trade but one industry is characterized by a higher degree of heterogeneity. With higher heterogeneity countries produce goods that are more different from each other or reflect larger differences in comparative advantage. Thus, consumers have a larger incentive to trade. The fact that consumers do not trade more implies that trade frictions must be higher in that industry.

Appendix C: Data Appendix

Trade data, weight-to-value and c.i.f./f.o.b. Bilateral and total export and import trade flows (thousand Euros), and their corresponding weight (tons) are used for 11 EU countries between 1999 and 2003 at the 4-digit Nace rev.1 level of manufacturing industries. *Source: Eurostat.*

In Section 4.1, the bilateral weight-to-value ratio of exports (kilograms per Euro exported) is calculated separately for the two partners in each industry and in each year, and is then averaged across all partners. In Section 4.3, the bilateral weight-to-value ratio of exports is calculated for each country, industry and year. In both cases, we calculate the log of one plus weight-to-value as weight-to-value often takes on values much smaller than one.

The ratio between bilateral import ("Costs, Insurance and Freight," *c.i.f.*) and export ("Free On Board," *f.o.b.*) flows is calculated separately for the two partners in each industry and in each year, and is then averaged across all partners, dropping the few cases where the ratio is smaller than one. We then use the log of the ratio.

Output Gross value of output (million Euros) at the 4-digit Nace rev.1 level. *Source: Eurostat's New Cronos database.*

Elasticities of Substitution We draw on estimates by Hummels (2001a) who uses data on bilateral trade flows, import tariffs and transport costs to estimate elasticities of substitution at the 2-digit SITC rev.3 level of industries. We use his OLS estimates (he also reports IV estimates which yield higher values for the elasticities). The elasticities are converted using tables of correspondence from the SITC rev.3 to the ISIC rev.3, and then from the ISIC rev.3 to the Nace rev.1. In the few cases where the Nace rev.1 level industries have to be matched with several SITC industries, we calculate the average across SITC industries. Ideally, one would compute a weighted average where the weights are given by the share of each SITC rev.3 industry into each Nace rev.1 industry grouping, but this information is not available. Among the 62 elasticities, 4 are not significantly different from zero so we set them to missing. *Source: Hummels (2001a).*

Gravity Variables Dummies for sharing a common land border and for sharing a common (official) language. *Source: Centre d'Etudes Prospectives et d'Informations Internationales (CEPII).* International and domestic distances, which are weighted averages of the distances between regions using GDP shares as weights. *Source: Chen (2004).*

Schengen Agreement For each country, a dummy is equal to one in the years in which the provisions of the Schengen Agreement are fully implemented. We then take the average between the two partners. The resulting variable takes on values of zero, one half and one. The years of first implementation are chosen depending on the month in which the Agreement went into force: Austria (1998-2004), Denmark (2001-2004), Finland (2001-2004), France (1995-2004), Germany (1995-2004), Italy (1998-2004), Netherlands (1995-2004), Portugal (1995-2004), Spain (1995-2004) while Ireland and the United Kingdom have not yet fully implemented the Agreement.

Technical Barriers to Trade We use two sources. The European Commission's Eurobarometer reports opinions and experiences of European managers about the Single Market. A total of 4,900 managers at companies were interviewed by telephone in early 2006, the sample of companies being selected according to the size of countries and of companies, and the industry of activity. We use the answer to the question: "Could you tell me whether you consider that for your company it is very important, rather important, rather unimportant or not important at all that future Single Market Policy tackles the question of *removing remaining technical barriers to trade in goods?*" For each country, we group the answers from all managers who replied that TBTs are indeed an important issue, and use the percentage so obtained as a country-specific indication on the relevance of TBTs. *Source: European Commission (2006).*

To capture the sectoral relevance of TBTs, industries are classified on a five-point scale according to the effectiveness of different measures undertaken by the Single Market Programme to eliminate TBTs: measures are successful and all significant barriers are removed (value of 1), measures are implemented and function well but some barriers remain (value of 2), measures are adopted but with implementation or transitional problems still to be overcome (value of 3), measures are proposed or implemented but not effective or with operating problems (value of 4), and no solution has been adopted (value of 5). Some industries are also identified as being not affected by TBTs prior to European integration (and are given a value of 0). Our industry-specific qualitative variable takes on values between 0 and 5, with larger values indicating a lack of market integration due to persisting TBTs. *Source: European Commission (1998).*

We then interact the log of one plus the average across partners of the share of managers replying that TBTs are important with one plus the industry-specific variable on TBTs.

Public Procurement We use two sources. For each country, we use the proportion of public procurement contracts advertised each year in the *Official Journal* of the EU. *Source: Eurostat.* The industries strongly affected by public procurement in Europe are identified at the Nace 70 level, are then converted at the Nace rev.1 level and identified by a dummy variable. *Source: Davies and Lyons (1996).* We calculate the log of one plus the average across partners of the proportion of public procurement contracts advertised in the *Official Journal* of the EU, which we then interact with the sectoral dummy.

Value-Added Taxes For each country, we use the standard VAT rate and replace it by the reduced VAT rates that apply to some categories of goods, as of January 1^{st}, 2008. We do not have information on the evolution of reduced VAT rates across goods over time, but we do not expect those to have changed much as the standard VAT rates changed very little during the time period we consider. We then compute the log of one plus the average across partners of the VAT rates that apply in each country and industry. *Source: European Commission (2008).*

Productivity Real labor productivity is value added divided by the number of employees at the 4-digit Nace Rev.1 level, deflated by GDP deflators. We then use the log of the average across partners of real labor productivity. *Source: Eurostat's New Cronos database.*

Capital Shares Capital shares are value added minus personnel costs, divided by personnel costs. We then calculate the log of one plus the absolute difference in capital shares between countries for each industry and in each year. *Source: Eurostat's New Cronos database.*

Number of Product Categories *Source: Eurostat's Prodcoms.*

GDP Deflators Value of GDP (million Euros) and volume of GDP (millions of 1995 constant Euros). The deflator is computed as the ratio between the value and the volume of GDP. *Source: Eurostat's New Cronos database.*

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

907	Dongshu Ou	To Leave or Not to Leave? A Regression Discontinuity Analysis of the Impact of Failing the High School Exit Exam
906	Andrew B. Bernard J. Bradford Jensen Stephen J. Redding Peter K. Schott	The Margins of US Trade
905	Gianluca Benigno Bianca De Paoli	On the International Dimension of Fiscal Policy
904	Stephen J. Redding	Economic Geography: A Review of the Theoretical and Empirical Literature
903	Andreas Georgiadis Alan Manning	Change and Continuity Among Minority Communities in Britain
902	Maria Bas	Trade, Technology Adoption and Wage Inequalities: Theory and Evidence
901	Holger Breinlich Chiara Criscuolo	Service Traders in the UK
900	Emanuel Ornelas John L. Turner	Protection and International Sourcing
899	Kosuke Aoki Takeshi Kimura	Central Bank's Two-Way Communication with the Public and Inflation Dynamics
898	Alan Manning Farzad Saidi	Understanding the Gender Pay Gap: What’s Competition Got to Do with It?
897	David M. Clark Richard Layard Rachel Smithies	Improving Access to Psychological Therapy: Initial Evaluation of the Two Demonstration Sites
896	Giorgio Barba Navaretti Riccardo Faini Alessandra Tucci	Does Family Control Affect Trade Performance? Evidence for Italian Firms
895	Jang Ping Thia	Why Capital Does Not Migrate to the South: A New Economic Geography Perspective
894	Kristian Behrens Frédéric Robert-Nicoud	Survival of the Fittest in Cities: Agglomeration, Selection and Polarisation
893	Sharon Belenzon Mark Schankerman	Motivation and Sorting in Open Source Software Innovation

892	Guy Michaels Ferdinand Rauch Stephen J. Redding	Urbanization and Structural Transformation
891	Nicholas Bloom Christos Genakos Ralf Martin Raffaella Sadun	Modern Management: Good for the Environment or Just Hot Air?
890	Paul Dolan Robert Metcalfe	Comparing willingness-to-pay and subjective well- being in the context of non-market goods
889	Alberto Galasso Mark Schankerman	Patent Thickets and the Market for Innovation: Evidence from Settlement of Patent Disputes
888	Raffaella Sadun	Does Planning Regulation Protect Independent Retailers?
887	Bernardo Guimaraes Kevin Sheedy	Sales and Monetary Policy
886	Andrew E. Clark David Masclet Marie-Claire Villeval	Effort and Comparison Income Experimental and Survey Evidence
885	Alex Bryson Richard B. Freeman	How Does Shared Capitalism Affect Economic Performance in the UK?
884	Paul Willman Rafael Gomez Alex Bryson	Trading Places: Employers, Unions and the Manufacture of Voice
883	Jang Ping Thia	The Impact of Trade on Aggregate Productivity and Welfare with Heterogeneous Firms and Business Cycle Uncertainty
882	Richard B. Freeman	When Workers Share in Profits: Effort and Responses to Shirking
881	Alex Bryson Michael White	Organizational Commitment: Do Workplace Practices Matter?
880	Mariano Bosch Marco Manacorda	Minimum Wages and Earnings Inequality in Urban Mexico. Revisiting the Evidence
879	Alejandro Cuñat Christian Fons-Rosen	Relative Factor Endowments and International Portfolio Choice
878	Marco Manacorda	The Cost of Grade Retention

The Centre for Economic Performance Publications Unit
Tel 020 7955 7284 Fax 020 7955 7595 Email info@cep.lse.ac.uk
Web site http://cep.lse.ac.uk