Gravity, trade integration, and heterogeneity across industries

Natalie Chen [a,b,*], Dennis Novy [a,c,d]

[a] University of Warwick, Coventry CV4 7AL, United Kingdom
[b] Centre for Economic Policy Research, London, United Kingdom
[c] Centre for Economic Performance, London School of Economics, London, United Kingdom
[d] CESifo, Munich, Germany

ARTICLE INFO

ABSTRACT

We derive a micro-founded measure of bilateral trade integration that is consistent with a broad range of leading gravity models. This measure accounts for cross-industry heterogeneity by incorporating substitution elasticities estimated at the industry level. We then use it to provide a theory-based ranking of trade integration across manufacturing industries in European Union countries. In addition, we explore the determinants of trade integration, finding that substantial Technical Barriers to Trade in certain industries as well as high transportation costs associated with heavy-weight goods are the most notable trade barriers.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Trade costs are a staple ingredient in today's trade literature. They feature prominently in the vast majority of theoretical papers. Broadly defined, trade costs include any cost of engaging in international trade such as transportation costs, tariffs, non-tariff barriers, informational costs, time costs and different product standards, among others. 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 important 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), 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, 2010; Hummels, 2001b).[1] 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. One way of doing this is to use the "phi-ness" of trade to estimate "border effects," which mostly reflect the extent of border-related costs (Head and Ries, 2001; Baldwin et al., 2003; Head and Mayer, 2004).[2] This indirect approach has the obvious advantage of extending the analysis to more countries, years and more finely disaggregated data.

This paper is part of the research effort that attempts to indirectly infer trade impediments from trade flows. Following the lead of Head and Ries (2001) and Head and Mayer (2004), the first contribution of

---

* Corresponding author at: University of Warwick, Department of Economics, Coventry CV4 7AL, United Kingdom. Tel.: +44 2476 528419; fax: +44 2476 523032.
E-mail addresses: N.A.Chen@warwick.ac.uk (N. Chen), D.Novy@warwick.ac.uk (D. Novy).

---

[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. Kee et al. (2009) estimate theory-based trade restrictiveness indices based on tariff and non-tariff barriers.

[2] See also Anderson and van Wincoop (2003), Baldwin et al. (2003), Chen (2004), Eaton and Kortum (2002), Evans (2003), Head and Mayer (2000), Head and Ries (2001), McCallum (1995), Nitsch (2000) and Wei (1996).

the paper is to develop a micro-founded measure of bilateral trade integration that can be applied to disaggregated panel data and that can be computed from observable trade and output 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 to be heterogeneous across industries. In contrast to the *phi*-ness measure, our measure of bilateral trade integration accounts for heterogeneity across industries by incorporating industry-specific substitution elasticities.

Arguably, the Anderson and van Wincoop monopolistic competition 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 extend the micro-foundations of the trade integration measure by showing that an isomorphic measure can also be derived from other models. These include the Ricardian trade model by Eaton and Kortum (2002), Chaney's (2008) extension of the Melitz (2003) heterogeneous firms model as well as 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 approach is therefore consistent with a broad range of the recent theoretical trade literature.

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 163 manufacturing industries in 11 European Union (EU) countries over the period 1999-2003. The case of the EU is appealing since trade integration is expected to be strong among its member states for two reasons. 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.

As they are required for the trade integration measure, we first estimate the substitution elasticities across the 163 manufacturing industries using the estimation approach pioneered by Feenstra (1994) and adapted by Broda and Weinstein (2006) and Imbs and Méjean (2009). As expected, we find that the elasticities vary substantially across industries. We then construct our trade integration measure and obtain a theory-based ranking of industries with intuitive results. For example, trade integration appears particularly low for "Bricks," "Plaster" and "Cement" as these industries are characterized by high transportation costs. Trade integration also tends to be low for perishable goods such as "Bread, fresh pastry goods and cakes." On the contrary, industries that are well integrated include a number of high-tech industries such as "Aircraft and spacecraft," "Engines and turbines" and "Computers." Our ranking is thus potentially useful to policymakers who wish to identify industries with poor trade integration.

As the next step, we attempt to explain the variation of trade integration both across countries and industries. Consistent with the standard gravity literature, the variation of trade integration *across countries* can to a large extent be captured by typical gravity variables such as distance and adjacency but also by policy-related variables such as participation in the Schengen Agreement. But our focus lies on the substantial degree of heterogeneity in trade integration *across industries*. Our results confirm that modeling trade costs as a "one-size-fits-all" impediment is clearly at odds with empirical evidence.

In particular, we investigate the role of several industry characteristics in explaining trade integration across industries, with an 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 industries where TBTs are strong, suggesting that there is room left for policy action and that further increases in market integration 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.

Finally, we contrast our methodology with alternative approaches. We believe this yields important insights. First, it is well-known that in standard gravity regressions of bilateral trade flows, the estimated coefficient on a trade cost proxy such as bilateral distance is a combination of the distance elasticity of trade costs and the elasticity of substitution (Anderson and van Wincoop, 2004; Hummels, 2001a). The distance elasticity of trade costs can only be derived once the substitution elasticity is known. However, in contrast to our measure, gravity equations cannot deliver a *ranking* that captures the extent of trade barriers across industries.

Second, although the *phi*-ness approach to proxy for trade barriers is conceptually closer to our methodology, it also struggles to deliver a meaningful ranking. The reason is that the *phi*-ness approach compares industries only by considering simple trade ratios, thus neglecting other features that might vary across industries such as the substitutability of goods and the degree of competition. In contrast, the trade integration measure that we employ embodies the substitution elasticities estimated at an earlier stage. It is thus able to separate differences in trade barriers from other forms of heterogeneity across industries, making it useful to quantify the extent of trade barriers across industries. Finally and perhaps most importantly, we find that the different approaches lead to conflicting policy conclusions. For example, we are able to identify TBTs as a major trade impediment, while the gravity and *phi*-ness frameworks do not.

Closest to our work are recent papers modeling and measuring trade barriers at the disaggregate level of industries. Anderson and van Wincoop (2004) model disaggregated trade flows and explicitly allow trade costs to vary at the industry level. Head and Ries (2001) and Head and Mayer (2004) rely on the *phi*-ness approach to measure and explain trade barriers across industries.[4] We differ from this literature in that the measure of trade impediments we propose incorporates industry-specific substitution elasticities which, as we argue, are important in capturing cross-industry variation that is distinct from variation in trade costs. This allows us to provide a ranking of trade integration across industries which the gravity and *phi*-ness approaches cannot deliver.

The paper is organized as follows. In Section 2 we use the monopolistic competition model by Anderson and van Wincoop (2004) to derive the trade integration measure. We also show that the measure is consistent with other types of leading trade models. In addition, we conceptually and empirically contrast our methodology with the standard gravity and *phi*-ness approaches. In Section 3 we present our data set, estimate the elasticities of substitution and present the ranking of trade integration across industries. In Section 4 we explain the variation of trade integration by relating it to observable trade cost proxies, draw policy implications and provide robustness checks. Section 5 concludes.

2. A model with industry-specific trade costs

In their seminal paper, Anderson and van Wincoop (2003) set up a general equilibrium model of trade that results in a micro-founded

---

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

[4] See also Harrigan (1996) who examines the openness to trade of OECD manufacturing industries. However, his approach relies on trade to output ratios and is therefore less grounded in theory. For example, trade to output ratios capture not only trade barriers but also multilateral resistance effects.

gravity equation. We closely follow Anderson and van Wincoop's (2004) generalization in which they model bilateral trade for individual industries characterized by industry-specific bilateral trade costs and industry-specific elasticities of substitution.[5]

Our interest lies on the industry-specific trade cost variables. Following the approach by Head and Ries (2001), we derive an analytical solution for them that depends on observable variables only. This solution gives rise to a micro-founded measure of industry-specific bilateral trade integration. In addition we demonstrate the generality of this trade integration measure. In particular, we show that it is consistent with leading trade theories including the Ricardian trade model by Eaton and Kortum (2002), Chaney's (2008) extension of the Melitz (2003) heterogeneous firms model as well as the heterogeneous firms model by Melitz and Ottaviano (2008).

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. 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 $p_{ij}^k = t_{ij}^k p_i^k$.

Imposing market-clearing and solving for general equilibrium, Anderson and van Wincoop (2004) derive the following gravity equation for industry k:

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

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 relative to world output y^k in industry k. 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, or outward multilateral resistance) and large average inward trade barriers of country j (i.e., large P_j^k, or inward multilateral resistance) lead to more bilateral trade.

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 Eq. (1). 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, 2011). 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, we use gravity Eq. (1) and consider domestic trade flows for industry-k goods

$$x_{ii}^k = \frac{y_i^k x_i^k}{y^k} \left(\frac{t_{ii}^k}{\Pi_i^k P_i^k} \right)^{1-\sigma_k}, \tag{2}$$

where t_{ii}^k are domestic trade costs for industry-k goods, for example domestic transportation costs. Eq. (2) can be solved for the product of outward and inward multilateral resistance as

$$\Pi_i^k P_i^k = \left(\frac{y_i^k x_i^k}{x_{ii}^k y^k} \right)^{\frac{1}{1-\sigma_k}} t_{ii}^k. \tag{3}$$

Note that we do not impose zero domestic trade costs since the trade cost factor t_{ii}^k may exceed unity.

2.2. A micro-founded measure of industry-specific trade integration

The solution for multilateral resistance can be exploited to solve the model. Gravity Eq. (1) contains the product of outward multilateral resistance of country i and inward multilateral resistance of country j, $\Pi_i^k P_j^k$, whereas Eq. (3) provides a solution for $\Pi_i^k P_i^k$. It is therefore useful to multiply gravity Eq. (1) 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

$$x_{ij}^k x_{ji}^k = \frac{y_i^k y_j^k x_i^k x_j^k}{y^k y^k} \left(\frac{t_{ij}^k t_{ji}^k}{\Pi_i^k P_i^k \Pi_j^k P_j^k} \right)^{1-\sigma_k}. \tag{4}$$

We substitute the solution for multilateral resistance given in Eq. (3) to obtain

$$x_{ij}^k x_{ji}^k = x_{ii}^k x_{jj}^k \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{1-\sigma_k}. \tag{5}$$

Following Head and Ries (2001) and Head and Mayer (2004), it is easy to solve for the trade cost factors as

$$\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} = \left(\frac{x_{ii}^k x_{jj}^k}{x_{ij}^k x_{ji}^k} \right)^{\frac{1}{\sigma_k-1}}. \tag{6}$$

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 $t_{ii}^k t_{jj}^k$.[6] We do not impose trade cost symmetry so that t_{ij}^k and t_{ji}^k on the left-hand side of Eq. (6) may be asymmetric ($t_{ij}^k \neq t_{ji}^k$). However, 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. Given this caveat from theory, we are unable to identify trade cost asymmetries. Instead, we take the square root to get an expression for the *average* bilateral relative trade barrier. It can be expressed as

$$\theta_{ij}^k \equiv \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{\frac{1}{2}} = \left(\frac{x_{ii}^k x_{jj}^k}{x_{ij}^k x_{ji}^k} \right)^{\frac{1}{2(\sigma_k-1)}}. \tag{7}$$

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

---

[5] Bergstrand (1989, 1990) also derives gravity equations for industry-level trade flows but in contrast to Anderson and van Wincoop (2004), he does not focus on multilateral resistance.

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

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. In the case approaching a closed economy, bilateral 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.

Importantly, the definition of θ_{ij}^k combines the ratio of domestic to bilateral trade with an exponent that involves the industry-specific elasticity of substitution σ_k. Intuitively, in an industry with a high elasticity σ_k, consumers are so price-sensitive that all else being equal, a fairly small price difference induced by bilateral trade costs can lead to a high ratio of domestic over bilateral trade. Therefore this ratio does not only reflect high bilateral trade barriers but also a low degree of product differentiation. θ_{ij}^k is precisely able to separate this competition, or *heterogeneity* effect, from the trade barrier effect. As Eq. (7) shows, a higher elasticity of substitution implies *lower* trade frictions θ_{ij}^k.

Finally, we stress that non-unitary income elasticities in gravity equations, as found by Santos Silva and Tenreyro (2006), do not pose a problem for our methodology. It is straightforward to show that if the elasticities of y_i^k and x_j^k in Eq. (1) differed from unity, the same trade integration measure θ_{ij}^k as in Eq. (7) would follow.[7]

2.3. Deriving the trade integration measure from alternative models

We have derived θ_{ij}^k in Eq. (7) from a model based on a CES demand system in combination with the Armington assumption that goods are differentiated by country of origin. In fact, our trade integration measure is valid more generally beyond this particular framework. We now show that it can also be derived from a range of other leading trade theories. As before, we apply these theories at the industry level, denoted by the index k.

2.3.1. 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^k and ζ_k. T_i^k determines the location of the productivity distribution for country i in industry k with a high T_i^k denoting high overall productivity. $\zeta_k > 1$ denotes the variation of productivity across goods and is treated as common across countries, with a high ζ_k denoting little variation. The model yields a gravity equation whose structure is similar to Eq. (1). It is given by

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

where c_i^k denotes the input cost in country i and industry k. We are interested in the trade cost parameters. T_i^k and c_i^k are unobservable but cancel out once the ratio of domestic over bilateral trade flows is formed as in Eq. (7). This yields

$$\theta_{ij}^{EK} = \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{\frac{1}{2}} = \left(\frac{x_{ii}^k x_{jj}^k}{x_{ij}^k x_{ji}^k} \right)^{\frac{1}{2\zeta_k}}. \tag{8}$$

Comparing Eqs. (7) and (8), it is obvious that $\theta_{ij}^{EK} = \theta_{ij}^k$ if $\zeta_k = \sigma_k - 1$.[8] As before, a low degree of heterogeneity (i.e., a high ζ_k) implies low trade barriers.

2.3.2. 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. 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:

$$x_{ij}^k = \mu^k \frac{y_i y_j}{y} \left(\frac{w_i t_{ij}^k}{\lambda_j^k} \right)^{-\gamma_k} \left(f_{ij}^k \right)^{-\left(\frac{\gamma_k}{\sigma_k - 1} - 1 \right)},$$

where μ^k is the weight of industry k 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^k is a remoteness variable akin to multilateral resistance and f_{ij}^k are the fixed costs of exporting from country i to j. γ_k is the shape parameter of the Pareto distribution, with a high γ_k denoting a low degree of heterogeneity and $\gamma_k > \sigma_k - 1$. Forming the ratio of domestic over bilateral trade flows yields

$$\theta_{ij}^{Ch} = \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{\frac{1}{2}} \left(\frac{f_{ij}^k f_{ji}^k}{f_{ii}^k f_{jj}^k} \right)^{\frac{1}{2} \left(\frac{1}{\sigma_k - 1} - \frac{1}{\gamma_k} \right)} = \left(\frac{x_{ii}^k x_{jj}^k}{x_{ij}^k x_{ji}^k} \right)^{\frac{1}{2\gamma_k}}. \tag{9}$$

Thus, the trade integration measure θ_{ij}^{Ch} is a function of both variable and fixed trade costs.[9] A low degree of heterogeneity (i.e., a high γ_k) again implies low trade barriers.

2.3.3. Melitz and Ottaviano (2008)

Melitz and Ottaviano (2008) also model heterogeneous firms. But in contrast to 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:

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

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

$$\theta_{ij}^{MO} = \left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k} \right)^{\frac{1}{2}} = \left(\frac{x_{ii}^k x_{jj}^k}{x_{ij}^k x_{ji}^k} \right)^{\frac{1}{2\gamma_k}}. \tag{10}$$

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.

---

[7] Suppose these elasticities were $\varkappa \neq 1$ instead of unity with $\varkappa > 0$. When forming the ratio of bilateral over domestic trade flows, one would obtain the same trade integration measure as in Eq. (7).

[8] 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).

[9] For non-zero trade flows (as is generally the case in our sample) and particular parameter restrictions in the productivity distribution, θ_{ij}^k is also consistent with the model by Helpman et al. (2008).

Variable trade costs are sufficient to induce selection into export markets because of bounded non-CES marginal utility.[10]

2.3.4. Summary

The four measures θ_{ij}^k, 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_k in Eq. (7) indicates a high degree of differentiation across products; a low ζ_k in Eq. (8) indicates a high variation of productivity across products; and a low γ_k in Eqs. (9) and (10) indicates a high degree of firm heterogeneity. For given trade flows, higher heterogeneity therefore 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 should have a larger incentive to trade. The fact that consumers do not trade more implies that trade frictions must be higher in that industry.

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 structure similar to Eq. (1). 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 for understanding the role of trade frictions.[11]

More generally, the method of considering ratios of domestic and bilateral trade flows is expected to work in models whose trade equations link origin-specific variables, destination-specific variables and country pair specific trade costs in a multiplicative way. These also include the models by Anderson (1979) and Krugman (1980). By taking the appropriate trade ratio as, for instance, in Eq. (7) one can ensure that all exporter-specific and importer-specific terms cancel and only bilateral trade cost variables remain.

2.4. How we differ from other approaches

Finally we explain how our method differs empirically from other approaches such as the gravity equation and the *phi*-ness of trade which is commonly used to evaluate the extent of trade frictions (Head and Ries, 2001; Baldwin et al., 2003; Head and Mayer, 2004). We describe how our measure differs *conceptually* from the other two approaches. In particular, we show that in contrast to gravity and *phi*-ness, θ_{ij}^k is able to measure and rank trade frictions across countries and industries as it distinguishes between the effect of trade barriers and the effect of heterogeneity. We believe this is an important feature.

2.4.1. The gravity equation

Dropping time subscripts for simplicity, we start with a micro-founded gravity specification that can be obtained by log-linearizing Eq. (4):

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

As is now commonly done in the literature (Baldwin and Taglioni, 2006), to control for the unobservable multilateral resistance variables $\Pi_i^k P_i^k \Pi_j^k P_j^k$ we need to include fixed effects for each industry in each year and separately for countries i and j, denoted by γ_i^k and γ_j^k. Thus, Eq. (4) becomes:

$$\ln\left(x_{ij}^k x_{ji}^k\right) = (1-\sigma_k)\ln\left(t_{ij}^k t_{ji}^k\right) + \gamma_i^k + \gamma_j^k. \tag{11}$$

To keep the exposition as simple as possible, assume that bilateral trade costs depend on distance only, i.e., $\ln(t_{ij}^k t_{ji}^k) = \beta_k \ln D_{ij}$ so that Eq. (11) can be estimated as:

$$\ln\left(x_{ij}^k x_{ji}^k\right) = \underbrace{(1-\sigma_k)\beta_k}_{\kappa_k} \ln D_{ij} + \gamma_i^k + \gamma_j^k + \varepsilon_{ij}^k, \tag{12}$$

where ε_{ij}^k denotes an error term. This micro-founded gravity equation only allows us to estimate the industry-specific elasticities of trade with respect to distance, κ_k. These are the products of the elasticities of substitution $(1-\sigma_k)$ and the elasticities of the "true" trade costs with respect to distance β_k (Anderson and van Wincoop, 2004; Hummels, 2001a).

Only once the elasticities of substitution σ_k are observed can the elasticities of trade costs with respect to distance be retrieved as $\beta_k = \kappa_k / (1-\sigma_k)$. But even in that case, the variation in trade costs $t_{ij}^k t_{ji}^k$ is still unknown since the β_k's do not provide any ranking of trade integration across industries.[12] As the gravity equation is unable to disentangle the heterogeneity effect from the trade barrier effect, it is impossible to quantify and compare trade frictions across countries and industries.[13]

---

[10] The fact that our approach works in this context is due to several particular features of the Melitz and Ottaviano (2008) model. First, the productivity cut-off for the marginal exporter to a certain destination market is a multiplicative function of bilateral trade costs and the domestic cut-off in that destination market. Second, due to the linear demand system an individual exporter's revenue is also a multiplicative function of (the square of) bilateral trade costs. Third, the assumption of a Pareto distribution for productivity ensures that aggregate revenue from exporting to the destination market is a simple multiplicative function of bilateral trade costs raised to the power of the Pareto shape parameter. We stress that we do not make any additional assumptions to make our approach work with this model.

[11] 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.)

---

[12] It is impossible to determine whether, for instance, "Concrete products" are more or less integrated than "Bricks."

[13] Also, as correctly specified gravity equations like Eq. (12) have to include time-varying industry fixed effects γ_i^k and γ_j^k to control for multilateral resistance, a practical problem is that these fixed effects are perfectly collinear with many explanatory variables of interest that only vary across sectors and over time. With our approach we avoid this problem because θ_{ij}^k nets out the time-varying multilateral resistance variables.

2.4.2. The phi-ness of trade

The (inverse) *phi*-ness of trade, denoted by ϕ_{ij}^k, is defined as the simple ratio of domestic to bilateral trade flows. Eq. (5) allows us to directly derive an expression for (inverse) *phi*-ness:

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

Assuming that domestic trade costs also only depend on distance, i.e., $\ln(t_{ii}^k t_{jj}^k) = \alpha_k \ln(D_{ii} \times D_{jj})^{1/2}$, Eq. (13) can be estimated as:

$$\ln \phi_{ij}^k = \underbrace{(\sigma_k - 1)\beta_k}_{\tilde{\kappa}_k} \ln D_{ij} + \underbrace{(1-\sigma_k)\alpha_k}_{\hat{\kappa}_k} \ln\left(D_{ii} \times D_{jj}\right)^{1/2} + \eta_{ij}^k, \tag{14}$$

where η_{ij}^k denotes an error term. The elasticities of trade costs with respect to distance can be computed as $\beta_k = \tilde{\kappa}_k / (\sigma_k - 1)$ and $\alpha_k = \hat{\kappa}_k / (1-\sigma_k)$ once the σ_k's are observed. But similar to gravity Eq. (12), the variation and ranking in trade costs are still unknown. The *phi*-ness estimation approach is thus inappropriate to capture the variation in trade frictions across countries and industries (and therefore to determine which countries or industries are more or less integrated).[14]

2.4.3. Bilateral trade integration

In contrast to gravity and *phi*-ness, θ_{ij}^k controls for heterogeneity across industries as the σ_k's are embodied in θ_{ij}^k. Regressing θ_{ij}^k on trade costs thus takes the form:

$$\ln \theta_{ij}^k = \ln\left(\frac{x_{ii}^k x_{jj}^k}{x_{ij}^k x_{ji}^k}\right)^{\frac{1}{2(\sigma_k - 1)}} = \ln\left(\frac{t_{ij}^k t_{ji}^k}{t_{ii}^k t_{jj}^k}\right)^{\frac{1}{2}} = (\beta_k / 2)\ln D_{ij} \tag{15}$$

$$+ (\alpha_k / 2)\ln\left(D_{ii} \times D_{jj}\right)^{1/2} + \xi_{ij}^k,$$

where ξ_{ij}^k denotes an error term. While the β_k and α_k coefficients can be indirectly derived from the gravity and *phi*-ness approaches, with θ_{ij}^k they can be estimated directly.[15] But most importantly, it is now possible to evaluate trade frictions across countries and industries as θ_{ij}^k directly captures relative trade costs $t_{ij}^k t_{ji}^k / t_{ii}^k t_{jj}^k$.

3. Data and descriptive statistics

3.1. Trade and output data

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. Our sample includes trade flows for 163 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.[16,17,18]

Our data set comprises a total of 15,145 domestic and bilateral trade flow observations. Only three bilateral trade flows are equal to zero, which is not surprising given the huge volume of intra-industry trade in the EU.[19] Those three 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, leading to large values of θ_{ij}^k.

3.2. Elasticities of substitution

Feenstra (1994) develops a methodology to estimate disaggregated substitution elasticities using data on imported quantities and prices (unit values). He estimates a demand schedule for US sectors where the σ_k's are derived by Instrumental Variables (IV) using country-sector fixed effects as instruments. Identification thus depends on the cross-section of exporters and is achieved in deviation from a reference country. This methodology can sometimes generate estimates that do not match any of the theoretically plausible values for σ_k. To solve this issue, Broda and Weinstein (2006) propose a grid search algorithm that minimizes the sum of squared residuals over the intervals of plausible values.

Imbs and Méjean (2009) implement Feenstra's (1994) methodology as well as Broda and Weinstein's (2006) grid search algorithm in which case the standard errors are obtained by bootstrapping with 1,000 repetitions. However, they do not estimate the σ_k's at the finest level of disaggregation of the data. Using disaggregated trade data at the 6-digit HS level they instead estimate the elasticities at the 3-digit ISIC Rev.3 level, assuming that all HS products are equally substitutable within an ISIC industry but not between. They also employ a correction suggested by Pesaran (2006) to control for aggregate shocks in exporting countries that affect all sectors simultaneously.

We closely follow Imbs and Méjean (2009) and estimate the elasticities σ_k at the 4-digit NACE Rev.1 level using highly disaggregated trade data at the 8-digit Combined Nomenclature (CN) level, assuming that all CN products are equally substitutable within a NACE industry. The trade data, taken from the same source as the data used to compute the ratio between domestic and bilateral trade flows, are the bilateral imports of the 11 EU countries from 46 developed and developing countries between 1999 and 2003 at the CN level.[20] But in contrast to Imbs and Méjean (2009) who estimate σ_k's that

---

[14] But in contrast to gravity Eq. (12), the *phi*-ness estimation approach nets out multilateral resistance variables so that time-varying country/industry fixed effects do not have to be included.

[15] Here, β_k and α_k are divided by 2 because θ_{ij}^k is the *average* trade barrier, i.e., we take the square root of $(t_{ij}^k t_{ji}^k)/(t_{ii}^k t_{jj}^k)$.

[16] 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.

[17] Manufacturing is split into 217 industries at the 4-digit NACE level while our sample only includes 163. Missing data for gross output values (required to compute domestic trade flows) in a single or several years between 1999 and 2003 is the reason why some industries are excluded from the analysis.

[18] The data appendix is available online at www.warwick.ac.uk/go/nachen.

[19] The zero trade flows are observed in "Publishing of newspapers" and twice in "Mattresses." We consider those observations as "true zeros" because the corresponding output values are positive whereas 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.

[20] The reason why our main analysis is conducted at the industry and not at the product level is the unavailability of product level output data that are required to compute θ_{ij}^k.

Table 1
Descriptive statistics for individual industries.

4-digit NACE Rev.1 industry (code)	σ_k	θ^k (rank)	ϕ^k (rank)
Bricks (2640)	1.82*	32,295.10 (1)	1.33×10^8 (9)
Plaster (2653)	2.17	798.67 (2)	2.38×10^7 (21)
Plate-making (2224)	2.24*	484.42 (3)	1.93×10^7 (24)
Striking of coins, medals (3621)	2.57*	368.46 (4)	4.02×10^8 (8)
Cold forming or folding (2733)	2.18*	119.90 (5)	7.54×10^5 (64)
Sugar (1583)	2.63*	84.93 (6)	1.87×10^7 (25)
Boilers (2822)	2.41*	54.98 (7)	9.86×10^6 (34)
Bread, fresh pastry goods, cakes (1581)	3.12	43.46 (8)	4.79×10^8 (7)
Cement (2651)	2.89	42.94 (9)	2.37×10^7 (22)
Publishing of newspapers (2212)	4.07	31.52 (10)	3.58×10^{12} (1)
Ceramic sanitary fixtures (2622)	2.33*	21.13 (11)	6,862.3 (121)
Other articles of paper (2125)	3.06	18.43 (12)	2.71×10^6 (51)
Imitation jewellery (3661)	3.28*	12.75 (13)	1.44×10^6 (56)
Concrete products (2661)	4.10	11.90 (14)	2.11×10^{10} (2)
Sawmilling, planing of wood (2010)	3.55*	11.78 (15)	4.05×10^7 (16)
Tanks, reservoirs (2821)	3.59*	11.31 (16)	1.16×10^8 (10)
Leather clothes (1810)	3.70	11.14 (17)	1.07×10^6 (60)
Mortars (2664)	3.79*	10.61 (18)	7.46×10^6 (38)
Jewellery (3622)	3.34*	9.19 (19)	6.07×10^5 (66)
Cutting, shaping of stone (2670)	4.11	8.14 (20)	7.11×10^7 (12)
Prepared feeds for farm animals (1571)	4.24	7.51 (21)	6.34×10^6 (42)
Fibre cement (2665)	3.86*	7.48 (22)	1.05×10^7 (32)
Mattresses (3615)	4.28	7.46 (23)	4.67×10^9 (3)
Builders' carpentry, joinery (2812)	4.49	7.44 (24)	7.99×10^7 (11)
Condiments, seasonings (1587)	3.82	7.20 (25)	4.68×10^6 (45)
Publishing of sound recordings (2214)	4.04*	7.09 (26)	6.36×10^6 (41)
Hollow glass (2613)	3.54	6.87 (27)	3.34×10^5 (74)
Builders' carpentry, joinery of metal (2030)	4.58	6.05 (28)	6.03×10^7 (13)
Ice cream (1552)	3.43	5.97 (29)	24,328.2 (110)
Food preparations, dietetic food (1588)	4.29	5.95 (30)	1.30×10^6 (59)
Cordage, rope, twine, netting (1752)	3.53	5.76 (31)	2.03×10^5 (80)
Processing of potatoes (1531)	4.75	4.86 (32)	4.19×10^6 (47)
Wire products (2873)	4.11	4.84 (33)	3.11×10^7 (19)
Processing of tea and coffee (1586)	4.18	4.80 (34)	3.42×10^6 (49)
Steam generators (2830)	5.14	4.73 (35)	5.66×10^7 (14)
Ceramic tiles, flags (2630)	3.87	4.63 (36)	1.43×10^6 (57)
Knitted and crocheted hosiery (1771)	4.42	4.57 (37)	1.47×10^7 (28)
Other non-metallic mineral products (2682)	3.62	4.47 (38)	41,098.2 (103)
Other kitchen furniture (3613)	5.58*	4.26 (39)	7.68×10^8 (6)
Wooden containers (2040)	5.55	4.25 (40)	2.69×10^9 (4)
Publishing of books (2211)	5.02	4.07 (41)	1.91×10^6 (54)
Publishing of journals, periodicals (2213)	5.91	4.02 (42)	5.24×10^7 (15)
Plaster products (2662)	5.47*	3.84 (43)	9.44×10^6 (36)
Building, repairing of ships (3511)	5.95	3.75 (44)	1.71×10^7 (26)
Printing (2222)	6.27	3.66 (45)	2.70×10^7 (20)
Builders' ware of plastic (2523)	5.61	3.39 (46)	4.31×10^6 (46)
Metal structures, parts (2811)	6.05	3.32 (47)	3.25×10^5 (50)
Grain mill products (1561)	5.17	3.25 (48)	1.36×10^6 (58)
Flat glass (2611)	4.41	3.23 (49)	9,684.8 (117)
Other furniture (3614)	5.32	3.17 (50)	9.27×10^5 (62)
Light metal packaging (2872)	5.24*	3.11 (51)	2.37×10^5 (78)
Other rubber products (2513)	4.73	3.05 (52)	56,565.2 (99)
Synthetic rubber (2417)	6.37	2.99 (53)	1.96×10^7 (23)
Other products of wood (2051)	6.24	2.95 (54)	2.25×10^6 (53)
Shaping, processing of flat glass (2612)	5.39	2.92 (55)	5.96×10^6 (43)
Abrasive products (2681)	3.64	2.90 (56)	692.3 (149)
Other plastic products (2524)	5.37	2.84 (57)	1.89×10^5 (81)
Other food products (1589)	5.92	2.74 (58)	7.83×10^5 (63)
Ceramic household articles (2621)	5.48	2.69 (59)	6.84×10^5 (65)
Brooms, brushes (3662)	5.20	2.68 (60)	87,378.1 (90)
Other ceramic products (2625)	4.47	2.65 (61)	1.58×10^5 (85)
Other office, shop furniture (3612)	7.03	2.63 (62)	3.66×10^5 (18)
Soaps, detergents (2451)	5.60	2.60 (63)	3.74×10^7 (17)
Wire drawing (2734)	5.28	2.58 (64)	3.77×10^5 (70)
Optical, photographic equipment (3340)	4.05*	2.58 (65)	1,517.0 (141)

Table 1 (*continued*)

4-digit NACE Rev.1 industry (code)	σ_k	θ^k (rank)	ϕ^k (rank)
Fasteners, screw machine products (2874)	4.86	2.55 (66)	46,013.0 (101)
Knitted, crocheted pullovers, cardigans (1772)	4.93	2.51 (67)	1.38×10^5 (86)
Macaroni, noodles, couscous (1585)	6.15	2.49 (68)	78,656.1 (93)
Other fabricated metal products (2875)	5.31	2.46 (69)	51,190.0 (100)
Paints, varnishes (2430)	5.95	2.44 (70)	3.66×10^5 (73)
Processing of fruit, vegetables (1533)	6.23	2.41 (71)	3.15×10^5 (75)
Locks, hinges (2863)	5.31	2.35 (72)	84,507.9 (91)
Carpets, rugs (1751)	5.65	2.32 (73)	40,405.0 (104)
Pulp (2111)	6.86	2.30 (74)	6.01×10^5 (67)
Plastic packing goods (2522)	6.76	2.26 (75)	1.08×10^7 (31)
Chairs, seats (3611)	6.88	2.25 (76)	6.47×10^6 (40)
Electrical equipment for engines, vehicles (3161)	5.66	2.25 (77)	60,711.0 (97)
Dressing, dyeing of fur (1830)	4.81*	2.23 (78)	1,968.3 (137)
Knitted, crocheted fabrics (1760)	6.39	2.22 (79)	67,149.0 (95)
Paper stationery (2123)	5.88	2.21 (80)	7,480.4 (118)
Machinery for metallurgy (2951)	5.95	2.20 (81)	3,879.3 (126)
Steel tubes (2722)	5.27	2.18 (82)	1.86×10^5 (82)
Medical, surgical equipment (3310)	4.95*	2.17 (83)	1,478.7 (142)
Man-made fibres (2470)	5.58	2.15 (84)	4,229.4 (124)
Invalid carriages (3543)	5.43	2.12 (85)	999.1 (146)
Furnaces, furnace burners (2921)	5.66	2.10 (86)	4,978.4 (123)
Plastics in primary forms (2416)	6.06	2.05 (87)	78,927.1 (92)
Wallpaper (2124)	4.48	2.04 (88)	467.0 (153)
Other publishing (2215)	8.30	2.02 (89)	1.46×10^7 (29)
Tobacco products (1600)	7.77*	1.99 (90)	15,670.2 (114)
Household, sanitary goods (2122)	6.23	1.98 (91)	13,157.9 (116)
Cocoa, chocolate, sugar confectionery (1584)	7.67	1.98 (92)	74,385.0 (94)
Plastic plates, sheets, tubes (2521)	6.22	1.94 (93)	17,725.5 (112)
Other agricultural machinery (2932)	8.46	1.93 (94)	4.43×10^5 (69)
Starches (1562)	6.19	1.91 (95)	2,034.5 (135)
Other outerwear (1822)	7.97	1.91 (96)	2.68×10^5 (77)
Tanning, dressing of leather (1910)	7.08	1.90 (97)	2.37×10^6 (52)
Perfumes, toilet preparations (2452)	5.75	1.90 (98)	3,498.6 (128)
Non-electric domestic appliances (2972)	7.09	1.88 (99)	32,472.9 (105)
Tools (2862)	7.58	1.85 (100)	56,662.1 (98)
Railway, tramway locomotives (3520)	9.08	1.80 (101)	4.05×10^6 (48)
Other electrical equipment (3162)	7.38	1.78 (102)	27,844.8 (108)
Other wearing apparel (1824)	9.24	1.78 (103)	1.69×10^5 (83)
Other manufacturing (3663)	7.73	1.73 (104)	15,938.2 (113)
Insulated wire, cable (3130)	7.60	1.72 (105)	7,350.1 (120)
Articles of cork, straw (2052)	6.44	1.72 (106)	662.5 (150)
Processing, preserving of fish (1520)	9.38	1.72 (107)	4.87×10^6 (44)
Steel drums (2871)	10.97	1.71 (108)	1.58×10^7 (27)
Explosives (2461)	11.67	1.69 (109)	1.03×10^7 (33)
Underwear (1823)	8.26	1.69 (110)	3.12×10^5 (76)
Processing of other glass (2615)	7.80	1.69 (111)	32,367.8 (106)
Non-wovens, except apparel (1753)	6.25	1.68 (112)	1,818.7 (138)
Machinery for food, beverage processing (2953)	9.66	1.68 (113)	3.68×10^5 (72)
Taps, valves (2913)	6.91	1.66 (114)	4,102.0 (125)
Other textiles (1754)	6.98	1.65 (115)	3,824.2 (127)
Made-up textile articles (1740)	9.19	1.65 (116)	43,914.9 (102)
Luggage, handbags (1920)	10.38	1.62 (117)	1.65×10^5 (84)
Glues, gelatines (2462)	8.60	1.62 (118)	13,266.0 (115)
Other transport equipment (3550)	11.43	1.61 (119)	1.09×10^7 (30)
Glass fibres (2614)	6.38	1.61 (120)	809.7 (147)
Weapons, ammunition (2960)	12.50	1.59 (121)	5.91×10^5 (68)
Lighting equipment (3150)	7.61	1.59 (122)	5,104.4 (122)
Machinery for mining, quarrying (2952)	7.61	1.59 (123)	2,229.4 (134)
Rubber tyres, tubes (2511)	6.64	1.58 (124)	1,149.2 (143)
Non-domestic cooling equipment (2923)	9.10	1.58 (125)	1.01×10^5 (88)
Parts, accessories for motor vehicles (3430)	8.56	1.58 (126)	1.82×10^6 (55)
Pharmaceutical preparations (2442)	9.07	1.57 (127)	23,565.3 (111)
Cutlery (2861)	7.20	1.57 (128)	583.3 (151)
Musical instruments (3630)	9.37	1.57 (129)	26,090.9 (109)

Table 1 (*continued*)

4-digit NACE Rev.1 industry (code)	σ_k	θ^k (rank)	ϕ^k (rank)
Workwear (1821)	13.35	1.57 (130)	9.61×10^8 (5)
Refractory ceramic products (2626)	10.02	1.55 (131)	3.71×10^5 (71)
Dairies, cheese (1551)	14.01	1.54 (132)	8.30×10^6 (37)
Pesticides (2420)	7.74	1.53 (133)	3,029.5 (131)
Bicycles (3542)	6.66	1.52 (134)	156.3 (156)
Footwear (1930)	10.16	1.51 (135)	1.27×10^5 (87)
Lifting, handling equipment (2922)	10.26	1.49 (136)	2.08×10^5 (79)
Other general purpose machinery (2924)	9.72	1.49 (137)	96,756.2 (89)
Paper, paperboard (2112)	8.00	1.47 (138)	1,110.2 (144)
Bearings, gears (2914)	7.07	1.45 (139)	518.2 (152)
Aircraft, spacecraft (3530)	6.72*	1.44 (140)	118.0 (157)
Electronic valves (3210)	8.28*	1.44 (141)	2,577.0 (132)
Electric domestic appliances (2971)	7.74	1.42 (142)	192.0 (155)
Fruit, vegetable juice (1532)	15.91	1.41 (143)	6.59×10^6 (39)
Electricity distribution (3120)	11.02	1.40 (144)	2,033.5 (136)
Sports goods (3640)	8.96	1.39 (145)	1,531.8 (140)
Prepared pet foods (1572)	13.19	1.37 (146)	9.67×10^5 (61)
Engines, turbines (2911)	12.99	1.37 (147)	66,696.6 (96)
Accumulators (3140)	7.25	1.36 (148)	110.3 (158)
Agricultural tractors (2931)	11.52*	1.35 (149)	1,756.5 (139)
Other chemical products (2466)	8.72	1.34 (150)	241.6 (154)
Electric motors (3110)	10.68	1.32 (151)	720.1 (148)
Building, repairing of boats (3512)	13.68	1.32 (152)	2,314.5 (133)
Basic pharmaceutical products (2441)	17.13	1.30 (153)	9.82×10^6 (35)
Machinery for paper production (2955)	14.13	1.29 (154)	7,364.1 (119)
Pumps, compressors (2912)	11.46	1.24 (155)	1,068.2 (145)
Essential oils (2463)	9.42	1.23 (156)	101.5 (159)
Computers (3002)	7.66	1.21 (157)	17.6 (164)
Motor vehicles (3410)	11.85*	1.21 (158)	95.0 (160)
Motorcycles (3541)	10.98*	1.21 (159)	94.1 (161)
Games, toys (3650)	9.67	1.16 (160)	19.6 (163)
Machinery for textile production (2954)	26.30	1.11 (161)	3,308.6 (129)
Office machinery (3001)	18.47	1.09 (162)	27.7 (162)
Other special purpose machinery (2956)	41.22	1.08 (163)	3,071.6 (130)

Notes: Authors' calculations. For θ^k and ϕ^k, the numbers reported are the averages across country pairs and years for each 4-digit industry. ϕ^k denotes the inverse *phi*-ness of trade. The σ_k's are estimated by Instrumental Variables (IV), while an asterisk (*) indicates that the elasticity is obtained by a grid search algorithm. "rank" provides the ranking of each 4-digit industry from the largest to the smallest value of θ^k and ϕ^k, respectively.

Fig. 1. Distribution of substitution elasticities σ_k across industries.

are specific to a destination market (i.e., the US), we constrain the elasticities to be the same for all 11 EU importing countries.[21] As a result, identification depends on the cross-section of exporters, importers and products.

The first column in Table 1 reports the elasticities σ_k estimated for each industry k. On average, σ_k is equal to 7.1, with values ranging from 1.8 to 41.2 and a standard deviation of 4.3. The median value is equal to 6.2, suggesting that the distribution of the elasticities is skewed to the right (5 out of 163 estimates are above 15). The distribution of those elasticities is depicted in Fig. 1. Those results confirm the well-known fact that elasticities are heterogeneous across industries, and are substantially larger the more disaggregated the data, reflecting that disaggregated quantities are more responsive to import prices than aggregate quantities.

The σ_k's are crucial to determine the level of inferred trade frictions θ_{ij}^k as well as their ranking across industries. It is therefore important to show that our elasticities are comparable to other estimates reported in the literature. Given that we use the methodology implemented by Imbs and Méjean (2009), we should expect our estimates to display some similarities with theirs, and they do. The
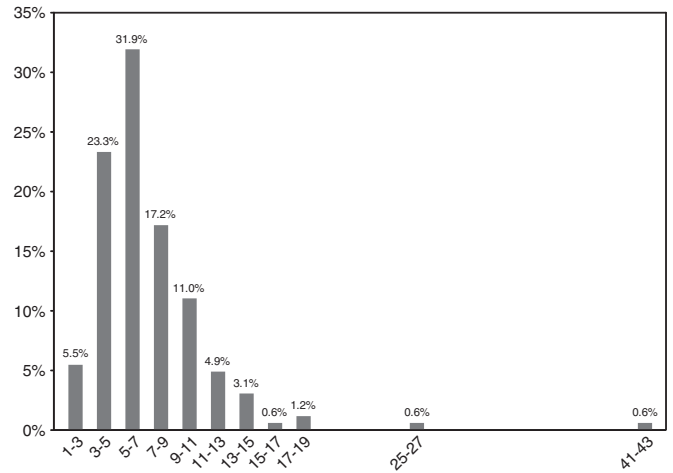
σ_k's by Imbs and Méjean (2009) that match with 159 of our 163 industries (i.e., 44 out of the 56 elasticities they estimate) range from 3.1 to 24.3 with an average of 4.9, a standard deviation of 1.9 and a median of 4.5. Their industries are more aggregated than ours, which explains why their elasticities are on average slightly smaller. But most importantly, notable similarities with our estimates can be identified. For instance, for "Knitted and crocheted fabrics and articles" (ISIC 1730) they estimate an elasticity of 5.10 against our elasticity of 4.93 for "Knitted and crocheted pullovers and cardigans" (NACE 1772). Also, for "Man-made fibres" (ISIC 2430 and NACE 2470) their elasticity of 5.36 directly compares to our estimate of 5.58. Divergences can however also be found. For instance, for "Aircraft and spacecraft" (ISIC 3530) their estimate of 24.30 contrasts with the value of 6.72 we find for the corresponding EU industry (NACE 3530).

A comparison with the σ_k's estimated by Broda and Weinstein (2006) at the 5-digit SITC Rev.3 level can be performed, too.[22] Among the SITC elasticities that match with our EU industries (i.e., 1,555 out of the 2,716 elasticities they estimate), their σ_k's range from 1.1 to 16,049 with an average of 16.4, a standard deviation of 407 and a median of 2.6. Here, the more disaggregated nature of their estimates explains why they are on average larger, but similarities with our σ_k's can still be found. For instance, for "Butadiene rubber" (SITC 232.12), a variety of synthetic rubber, they estimate an elasticity of 6.36 against our elasticity of 6.37 for "Synthetic rubber" (NACE 2417). Also, the elasticities for "Cocoa paste, not defatted" (SITC 72.31) and "Cocoa, chocolate and sugar confectionery" (NACE 1584) are equal to 7.69 and 7.67, respectively. Conversely, leaving aside an outlier value of 16,049 among their estimates, discrepancies between the two sets of elasticities can also be found, such as in "Motorcycles" (SITC 785.16) where their elasticity of 119.28 contrasts with our elasticity of 10.98 (NACE 3541).

Overall, despite the similarities between our estimates and those reported by Imbs and Méjean (2009) or Broda and Weinstein (2006), we believe that our elasticities are more appropriate in order to measure θ_{ij}^k as they are estimated for the same sample of EU countries, the same time period and at the same level of disaggregation as the trade data we use to construct the bilateral trade integration measure.

3.3. Trade frictions

The second column in Table 1 reports, for each industry, the average value of θ_{ij}^k across country pairs and over time, where

---

[21] For the sake of comparison, we also estimated elasticities that vary across importing countries, σ_j^k, as in Imbs and Méjean (2009). The correlations between those elasticities and the constrained elasticities used to compute θ_{ij}^k are all positive but not perfect.

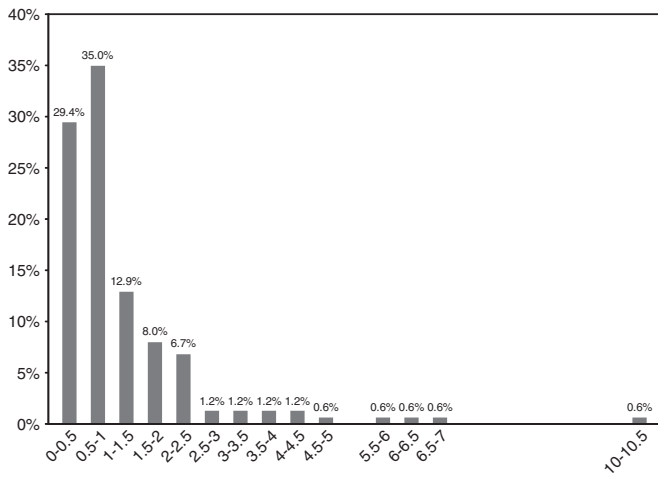[22] We do not compare our results to Feenstra (1994) as he reports estimates for six manufacturing products only.

Fig. 2. Distribution of average (log) trade friction measure θ^k across industries.

Fig. 3. Values of σ_k (left vertical axis and black bars) with corresponding values of (log) θ^k (right vertical axis and grey bars), ordered by decreasing values of σ_k.

industries are ordered by decreasing value of θ^k. The values range from 32,295 to 1.08. For example, the value 1.08 means that bilateral trade barriers in that industry are on average 8% higher than domestic trade barriers. The median value of θ^k is around 2.1, implying bilateral trade barriers that are on average 110% higher than domestic trade barriers. This value compares to an average tariff equivalent of 74% for international trade costs as reported by Anderson and van Wincoop (2004) for aggregate trade data.

Fig. 2 depicts the distribution of the (log) average trade frictions across industries. It illustrates the enormous heterogeneity across industries. Consistent with Eq. (7), which shows that a higher elasticity σ_k is associated with lower implied trade frictions, the correlation between (log) θ^k and σ_k is equal to -49%, suggesting that our ranking of trade integration across industries is substantially driven by the elasticities of substitution. Another way to evaluate the importance of σ_k for the trade integration measure is to perform a variance decomposition. After regressing (log) θ^k on its elasticity component $1/2(\sigma_k - 1)$ and its trade component given by the average (log inverse) phi-ness of trade ϕ^k for each industry, the decomposition shows that the contribution to the variance of (log) θ^k is 82% for the elasticity component and only 11% for the phi-ness component.[23],[24] The negative relationship between (log) θ^k and σ_k is represented in Fig. 3.

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 "Plaster" and "Plate-making." "Bricks" and "Plaster" are characterized by large weight-to-value ratios, indicating a very low transportability of the goods. Other related industries are "Cement" and "Concrete products." Note that the geographic market for cement and concrete is locally confined since the perishable nature of such "wet" products constrains the distance over which they can be delivered. Other industries with low trade integration include "Bread, fresh pastry goods and cakes" due to their perishable nature. "Publishing of newspapers" is traded very little, too, which is hardly surprising given
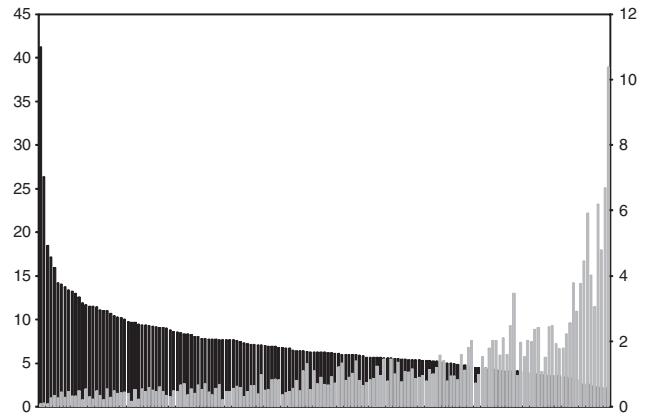
the reliance of such products on specific languages. This finding is consistent with earlier studies showing that trade in such sectors is subdued. For instance, Harrigan (1996) shows that the volume of trade relative to output in the OECD in 1985 is the lowest in "Printing and Publishing." At the other end, industries that are well integrated include a number of high-tech industries such as "Aircraft and spacecraft," "Engines and turbines," "Accumulators," "Agricultural tractors," "Electric motors," "Computers," "Motor vehicles," "Machinery for textile production" and the like.

A crucial advantage of the θ^k measure is to provide a theory-based ranking of trade integration across industries. For example, "Plaster," "Cement" and "Mortars" are all similar, poorly integrated industries with a high weight-to-value. But it is hard to determine, a priori, which is the most integrated. The θ^k measure precisely allows us to answer this question and reveals that the least integrated of the three is "Plaster" while the most integrated is "Mortars." Also, although "Imitation jewellery" and "Leather clothes" produce low weight-to-value products, they are less integrated than the high weight-to-value industry "Mortars." This suggests they are both hampered by significant trade barriers other than transportation costs.

Finally, as (inverse) phi-ness is a component of θ_{ij}^k, the two measures are naturally expected to be correlated with each other. Column (3) in Table 1 reports the inverse phi-ness of trade for each industry ϕ^k. The correlation between θ^k and ϕ^k (in logs) is equal to 51% (as depicted in Fig. 4) while their rank correlation reaches

---

[23] By definition, $\ln\theta_{ij}^k = 1/(2(\sigma_k - 1))\ln\phi_{ij}^k$.
[24] We repeated the same exercise using as a dependent variable (log) θ_{ij}^k. Depending on the specifications (no fixed effects or with 3-digit, exporter, importer or pair fixed effects, interacted or not with year dummies), the contributions from the variance in the elasticity component vary between 60 and 63% while the contributions from the variance in (inverse) phi-ness vary between 29 and 32%. The contribution of the variance of each component x_m to the total variance of the dependent variable Y is calculated as $c_m = \beta_m cov(x_m, Y)/var(Y)$, where β_m is the partial regression coefficient of Y on the explanatory variable x_m, holding all other explanatory variables constant (Fields, 2003).
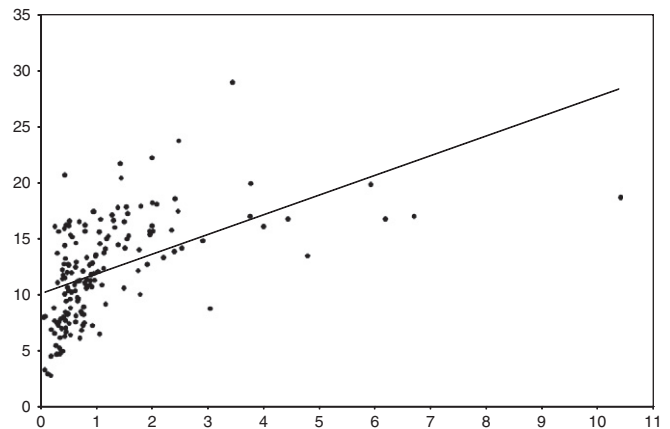

Fig. 4. Average (log) inverse phi-ness of trade ϕ^k (vertical axis) versus average (log) trade friction measure θ^k (horizontal axis) across industries.

66%.[25] It therefore follows that the two measures yield roughly similar insights regarding trade integration across industries. For example, "Bricks" and "Striking of coins" are consistently among the less integrated industries while "Office machinery" and "Other purpose machinery" are among the most integrated.

Our trade integration measure would only coincide with the (inverse) *phi*-ness of trade if we assumed that the elasticity of substitution were equal to 1.5 for all industries such that $1/2(\sigma_k - 1) = 1$ for all k in Eq. (7).[26] But this assumption would clearly be at odds with the findings of the recent literature (Imbs and Méjean, 2009; Broda and Weinstein, 2006; Feenstra, 1994) as well as with our own estimates. Most importantly, this assumption affects the conclusions regarding trade integration across industries. As can be seen from Table 1, "Workwear" and "Basic pharmaceutical products" appear to be highly integrated according to θ^k but poorly integrated according to ϕ^k. The opposite is true for "Ceramic sanitary fixtures." In terms of rankings, θ^k indicates that "Bricks" is less integrated than "Publishing of newspapers" while ϕ^k suggests the opposite. The same observation applies to "Cement" and "Concrete products," among other industries.

To conclude, even if θ^k and ϕ^k are correlated with each other, they are different because θ^k has a feature (the σ_k's) that ϕ^k misses, and this difference matters not only in theory but also in the data. In contrast to θ_{ij}^k, *phi*-ness is unable to distinguish between the trade barrier effect and the heterogeneity effect, as captured by the elasticity of substitution σ_k.

4. The determinants of bilateral trade integration

4.1. Our trade integration measure

We now examine formally how the trade integration measure performs in the data. To analyze the factors that correlate with bilateral EU trade integration across countries and industries we estimate

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

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 as well as the elasticity of substitution embedded in $\theta_{ij,t}^k$ vary across 4-digit industries 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. ψ_t denotes year intercepts that absorb common trade cost shocks. 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. 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 (15,145/5 = 3,029 clusters).[27]

4.1.1. Geography/transportation costs variables

Table 2 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 frictions. Trade frictions increase with international distances D_{ij} and decrease with domestic distances $D_{ii} \times D_{jj}$, the elasticities being equal to 0.471 and −0.821, respectively. Trade frictions between two contiguous countries (Adj_{ij}) are 7.4% lower than between countries that do not share a border and 14% lower between two countries that speak the same language $Lang_{ij}$.[28]

We also consider proxies of transportation costs that vary across industries. First, we consider the weight-to-value ratio of exports for each industry and year, and averaged across all partners, wv_t^k. We do not consider *bilateral* weight-to-value ratios because they are endogenous (they include trade flows that are part of the dependent variable), and they cannot be computed when trade is zero. 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. Second, we consider the ratio of *c.i.f.* and *f.o.b.* trade values for each industry, $cfob_t^k$, averaged in each year across all EU partners in order to minimize measurement error (Harrigan, 1993) as well as endogeneity issues, but also because bilateral *c.i.f.* over *f.o.b.* trade values cannot be computed when trade is zero.[29] As expected, higher transportation costs as proxied by $cfob_t^k$ and wv_t^k increase trade frictions with elasticities equal to 0.062 and 0.461, respectively.

4.1.2. 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 as θ_{ij}^k is 23% higher compared to the other countries in the sample. 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.[30] Nonadoption of the Euro had no effect on the extent of trade integration for these two countries since 2002. This result is perhaps explained by the short span of data (2002–2003) used to investigate the hypothesis, or by anticipation effects.

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. Trade frictions between two participating countries are 10% lower than between two non-participating countries. And overall, if all countries participated in the agreement trade frictions would on average be lower by 1.9%.[31]

---

[25] The ranking of each industry according to the two measures is reported in parentheses in the last two columns of Table 1.

[26] The *phi*-ness of trade is sometimes also defined as the square root of bilateral to domestic trade flows (Head and Mayer, 2004). In that case, the relevant elasticity of substitution would be equal to 2.

[27] Our set of determinants is not meant to be exhaustive. There are other possible determinants that we do not consider here such as business and immigration networks (Dunlevy, 2006; Rauch, 2001).

[28] The effects are computed as $(exp(\beta) - 1) \times 100\%$ where β is the estimated coefficient.

[29] 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% of total intra-EU15 trade went by road, 22.8% by sea, 3.9% by rail, 3.9% by air, 0.9% by inland waterways and 0.8% by pipeline (European Commission, 2000).

[30] 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). This issue is beyond the scope of this paper.

[31] The sample mean of $Schengen_{ij,t}$ is equal to 0.81. For all countries to participate in Schengen ($Schengen_{ij,t} = 1$), the variable would need to increase by 0.19 units which would translate into a $0.19 \times 10 = 1.9\%$ decrease in θ_{ij}^k.

Table 2
The determinants of EU trade integration.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	$\ln\theta_{ij,t}^k$	$\ln\theta_{ij,t}^k$	$\ln\theta_{ij,t}^k$	$\ln\theta_{ij,t}^k$	$\ln\theta_{ij,t}^k$	$\ln\theta_{ij,t}^k$	$\ln\theta_{ij,t}^{Ch}$	$\ln\theta_{ij,t}^{Ch}$	$\ln\phi_{ij,t}^k$	$\ln(x_{ij,t}^k x_{ji,t}^k)$
Outputs										
$\ln y_{i,t}^k$	–	–	–	–	–	–	–	–	–	1.172a
										(31.146)
$\ln y_{j,t}^k$	–	–	–	–	–	–	–	–	–	1.088a
										(30.590)
Geography/transport costs										
$\ln D_{ij}$	0.471a	0.411a	0.447a	0.340a	0.382a	0.408a	0.649a	0.332a	3.883a	−2.872a
	(15.294)	(11.720)	(13.165)	(9.547)	(11.405)	(12.422)	(19.213)	(10.635)	(28.808)	(−24.044)
$\ln(D_{ii}\times D_{jj})$	−0.821a	−1.030a	−1.134a	−0.917a	−0.928a	−1.015a	−1.464a	−0.835a	−7.658a	–
	(−6.253)	(−7.033)	(−7.748)	(−6.320)	(−6.910)	(−7.541)	(−12.940)	(−6.692)	(−15.316)	
Adj_{ij}	−0.077b	−0.070b	−0.053	−0.118a	−0.094a	−0.081b	−0.049	−0.101a	−0.052	0.952a
	(−2.357)	(−2.038)	(−1.570)	(−3.398)	(−2.851)	(−2.511)	(−1.509)	(−3.189)	(−0.430)	(8.949)
$Lang_{ij}$	−0.149a	−0.296a	−0.294a	−0.305a	−0.327a	−0.325a	−0.483a	−0.286a	−2.620a	2.092a
	(−2.630)	(−4.868)	(−4.826)	(−5.084)	(−5.688)	(−5.644)	(−7.717)	(−5.218)	(−12.225)	(9.846)
$\ln wv_t^k$	0.461a	0.319a	0.329a	0.379a	0.385a	0.396a	0.579a	0.470a	2.400a	−1.963a
	(3.734)	(3.014)	(3.249)	(3.442)	(3.565)	(3.875)	(11.528)	(4.327)	(7.676)	(−6.742)
$\ln cfob_t^k$	0.062a	0.064a	0.065a	0.062a	0.019b	0.019b	0.026a	0.017b	0.126a	−0.043c
	(6.487)	(7.080)	(7.038)	(6.760)	(2.449)	(2.472)	(3.983)	(2.276)	(4.636)	(−1.744)
Policy variables										
FI, AT_{ij}	–	0.205a	0.177a	0.250a	0.230a	0.207a	0.332a	0.209a	1.715a	−1.714a
		(7.491)	(6.543)	(8.817)	(8.743)	(7.902)	(12.694)	(8.526)	(16.203)	(−16.302)
$noEURO_{ij,t}$	–	0.010	−0.014	0.021	0.007	−0.013	0.011	0.010	0.063	−0.210a
		(0.506)	(−0.700)	(0.998)	(0.362)	(−0.669)	(0.540)	(0.514)	(0.795)	(−2.678)
$Schengen_{ij,t}$	–	−0.106a	−0.058	−0.133a	−0.130a	−0.091b	−0.162a	−0.117a	−0.776a	1.022a
		(−2.738)	(−1.567)	(−3.173)	(−3.382)	(−2.447)	(−4.168)	(−3.268)	(−4.988)	(6.906)
$\ln TBT_{ij}^k$	–	0.454a	–	0.477a	0.382a	–	0.196a	0.365a	0.174	−0.888a
		(3.768)		(3.727)	(3.035)		(4.903)	(2.833)	(0.542)	(−2.809)
TBT^k	–	–	0.100b	–	–	0.084c	–	–	–	–
			(2.351)			(1.855)				
$\ln Proc_{ij,t}^k$	–	−2.990a	−2.671b	−4.002a	−3.868a	−3.621a	1.541	−3.471a	−9.915	14.732b
		(−2.812)	(−2.506)	(−3.332)	(−3.225)	(−3.010)	(0.935)	(−3.314)	(−1.416)	(2.246)
$\ln VAT_{ij}^k$	–	−16.834a	−16.626a	−16.644a	−13.244a	−13.222a	−8.112a	−11.065a	−49.638a	−7.094
		(−6.472)	(−6.369)	(−6.432)	(−5.959)	(−5.923)	(−4.058)	(−5.269)	(−5.527)	(−0.811)
Other costs										
$\ln Prod_{ij,t-1}^k$	–	–	–	−0.367a	−0.353a	−0.367a	−0.326a	−0.387a	−1.403a	1.243a
				(−6.185)	(−6.970)	(−7.458)	(−6.964)	(−7.549)	(−7.887)	(6.961)
$Zeros_{ij,t}^k$	–	–	–	3.038a	2.789a	2.775a	2.705a	2.055a	14.668a	−14.296a
				(10.852)	(16.705)	(15.593)	(6.799)	(6.682)	(10.080)	(−7.202)
Controls										
$KS_{ij,t}^k$	–	–	–	–	0.026	0.044	−0.042	0.006	0.383b	0.014
					(0.563)	(0.909)	(−0.894)	(0.131)	(2.185)	(0.085)
$\ln Goods^k$	–	–	–	–	−0.340a	−0.340a	−0.217a	−0.317a	−1.183a	0.915a
					(−14.330)	(−14.204)	(−14.877)	(−14.310)	(−15.331)	(10.710)
Fixed effects	t, K	t, K	t, K	t, K	t, K	t, K	$t, \tilde K$	t, K	t, K	t, K
N	15,145	15,145	15,145	12,116	12,116	12,116	9,856	12,112	12,116	12,116
Adj−R^2	0.665	0.677	0.675	0.683	0.722	0.721	0.714	0.741	0.690	0.812

Notes: The dependent variable is $\ln\theta_{ij,t}^k$ in (1) to (6), $\ln\theta_{ij,t}^{Ch}$ in (7) and (8), $\ln\phi_{ij,t}^k$ in (9) and $\ln(x_{ij,t}^k x_{ji,t}^k)$ in (10). In (7) and (8), $\theta_{ij,t}^{Ch}$ is computed using the Pareto shape parameters γ_k from Crozet and Koenig (2010) and di Giovanni et al. (2011), respectively. The fixed effects t, K and $\tilde K$ denote year, 3-digit and 2-digit fixed effects, respectively. Robust standard errors are adjusted for clustering at the 4-digit NACE Rev.1 level in each country pair. 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% levels, respectively.

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.[32] 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_{ij}^k, increases with the extent of TBTs across countries and industries. The data appendix provides details and descriptions across countries and industries. 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

[32] 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."

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.[33] In column (3) we therefore report estimates when only including our industry-specific variable on TBTs, TBT^k. The results remain consistent, i.e., when TBTs are high, trade integration is low.[34]

The results in column (3) indicate that a unit increase in TBT^k increases θ_{ij}^k by 10% while the elasticity evaluated at sample mean is equal to 0.191.[35] Most importantly, the complete removal of remaining TBTs would lower θ_{ij}^k by 9.1% (as the sample mean of TBT^k is equal to 1.91, for all TBTs to be completely removed TBT^k would need to be reduced by 0.91 units to reach a value of 1. This would lower trade frictions by $0.91 \times 10 = 9.1\%$). This magnitude is far from negligible and suggests that by removing those barriers policymakers could significantly promote trade integration across industries.

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 increasing in those markets, with the proportion of EU15 public procurement contracts openly advertized in the *Official Journal* of the EU steadily increasing from 8.4% in 1995 to 16.8% in 2005 (Eurostat). To investigate whether this opening of markets has helped to lower trade frictions, we rely on time-varying cross-country data on the share of public procurement contracts advertized in the *Official Journal*. As similar data across industries are not always available, we compute an indicator variable for high public procurement sectors.[36] We interact this indicator variable with the time-varying data for each country, denoting the variable so obtained by $Proc_{ij,t}^k$. Column (2) reveals that the opening of public procurement markets to foreign competition has indeed been successful in fostering trade integration across countries, industries and over time. However, the complete liberalization of public procurement markets would lower trade frictions by a modest 0.041% only.[37]

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_{ij}^k unchanged. In column (2) we include the product of the log of one plus the VAT rate of each partner at the sectoral level, VAT_{ij}^k, 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_{ij}^k might be lower simply because intra-EU export statistics are inflated due to VAT fraud.[38]

4.1.3. 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 relationship between productivity and trade frictions θ_{ij}^k. In column (4) of Table 2 we therefore include for each industry and year the average of real labor productivity between trading partners. 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 et al., 2006, for the US 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 (with an elasticity of -0.367).

We also include a dummy variable in column (4) that controls for the three 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.

4.1.4. 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.[39] 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

---

[33] If this were indeed the case, we would underestimate the impact of TBTs so that our coefficients could be considered lower-bound estimates.

[34] 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.

[35] This can be calculated as $\partial \ln\theta_{ij}^k / \partial \ln TBT^k = \partial \ln\theta_{ij}^k / \partial TBT^k \times \partial TBT^k / \partial \ln TBT^k = 0.100 \times TBT^k = 0.191$ if calculated at the sample mean of TBT^k, which is equal to 1.91.

[36] Examples of high public procurement industries are "Shipbuilding," "Rail stock," "Pharmaceuticals" or "Aerospace."

[37] We calculate the elasticities evaluated at sample means of θ_{ij}^k with respect to the proportion of public procurement contracts advertized in each of the two partner countries and then compute the total effect on θ_{ij}^k, assuming that the proportions of public procurement contracts advertized increase to 100% in each country (see the data appendix for details on these data).

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

[39] With intra-industry trade, the denominator of Eq. (7) will be given by the product of two balanced trade flows. But with comparative advantage, the denominator of Eq. (7) 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.

Table 3
Variance decompositions.

Dependent variable	(1) $\ln\theta_{ij,t}^k$	(2) $\ln\theta_{ij,t}^k$	(3) $\ln\theta_{ij,t}^{Ch}$	(4) $\ln\theta_{ij,t}^{Ch}$	(5) $\ln\theta_{ij,t}^k$	(6) $\ln(x_{ij}^k x_{ji,t}^k)$
Outputs						**40.14%**
$\ln y_{i,t}^k$	–	–	–	–	–	17.93%
$\ln y_{j,t}^k$	–	–	–	–	–	22.21%
Geography/transport costs	**9.43%**	**9.71%**	**22.90%**	**11.03%**	**31.78%**	**19.91%**
$\ln D_{ij}$	1.69%	1.81%	11.37%	1.30%	19.66%	13.10%
$\ln(D_{ii} \times D_{jj})$	0.73%	0.79%	2.38%	0.58%	3.56%	–
Adj_{ij}	0.43%	0.37%	0.63%	0.42%	0.20%	2.98%
$Lang_{ij}$	0.28%	0.27%	1.07%	0.24%	1.70%	0.87%
$\ln wv_t^k$	5.69%	5.85%	6.90%	7.95%	5.68%	2.67%
$\ln cfob_t^k$	0.61%	0.62%	0.55%	0.54%	0.98%	0.29%
Policy variables	**4.93%**	**4.28%**	**7.08%**	**4.46%**	**7.69%**	**8.70%**
FI, AT_{ij}	0.78%	0.70%	3.75%	0.56%	5.60%	6.13%
$noEURO_{ij,t}$	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
$Schengen_{ij,t}$	0.04%	0.03%	0.41%	0.02%	0.59%	0.65%
$\ln TBT_{ij}^k$	3.25%	–	2.35%	2.61%	0.43%	1.55%
TBT^k	–	2.70%	–	–	–	–
$\ln Proc_{ij,t}^k$	0.07%	0.06%	0.00%	0.08%	0.11%	0.19%
$\ln VAT_{ij}^k$	0.79%	0.79%	0.57%	1.19%	0.96%	0.18%
Other costs	**0.24%**	**0.24%**	**1.45%**	**0.13%**	**1.50%**	**1.99%**
$\ln Prod_{ij,t-1}^k$	0.01%	0.01%	1.09%	0.00%	0.91%	1.58%
$Zeros_{ij,t}^k$	0.23%	0.23%	0.36%	0.13%	0.59%	0.41%
Controls	**13.94%**	**13.96%**	**8.51%**	**12.78%**	**11.83%**	**8.32%**
$KS_{ij,t}^k$	0.04%	0.06%	0.00%	0.01%	0.01%	0.00%
$\ln Goods^k$	13.90%	13.90%	8.51%	12.77%	11.82%	8.32%
3-digit industry fixed effects	43.67%	43.91%	31.49%	45.73%	16.20%	2.12%
Year fixed effects	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
Variation explained	72.22%	72.11%	71.44%	74.14%	69.01%	81.19%
Residual	27.78%	27.89%	28.56%	25.86%	30.99%	18.81%
Sum	100%	100%	100%	100%	100%	100%
N	12,116	12,116	9,856	12,112	12,116	12,116

Notes: The variance decompositions are calculated according to Fields (2003). The contribution of each explanatory variable x_m to the total variance of a dependent variable Y is given by $c_m = \beta_m cov(x_m, Y)/var(Y)$ where β_m is the partial regression coefficient of Y on the explanatory variable x_m (holding all other explanatory variables constant). The decompositions in columns (1) to (6) correspond to regressions (5) to (10) in Table 2. The contributions sum to 100%.

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. The resulting θ_{ij}^k is therefore expected to be smaller because the denominator in Eq. (7) 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.[40] 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 are 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.

4.2. Policy implications

For both regressions in columns (5) and (6) of Table 2, columns (1) and (2) in Table 3 report variance decompositions to illustrate the quantitative contribution of each factor in explaining the variation in the dependent variable. Column (1) in Table 3 shows that regression (5) in Table 2 explains 72% of the variation in θ_{ij}^k. 44% can be attributed to the 3-digit industry fixed effects and 0.01% to the year fixed effects. Of the 28% that the remaining regressors explain, geography and transport costs alone are responsible for about 9% of the variation in θ_{ij}^k, the most important factor being weight-to-value (5.7%), followed by bilateral distance (1.7%). Policy variables explain 5% of the variation in θ_{ij}^k, which is far from negligible. TBTs are the most important factor (3.2%), while the other factors only play minor roles. These policy variables can arguably be influenced by policymakers. The number of product categories within each industry explains up to 13.9% of the variation in θ_{ij}^k. The decomposition reported in column (2), which corresponds to regression (6) in Table 2, gives similar results. From a policy perspective, those results suggest that in order to reduce the variance of trade integration across countries and industries, policymakers should give high priority to removing TBTs. Only then should they improve infrastructure policies (associated with the role of distance, assuming that policymakers cannot influence weight-to-value).

4.3. Our trade integration measure derived from alternative models

As we show in Eqs. (8)–(10), the exponent of the trade integration measure is not necessarily a function of the CES substitution elasticity. Instead, it could also be a function of technology parameters, i.e., the Fréchet or Pareto parameters.

To the best of our knowledge, we are not aware of Fréchet parameter estimates at a fine level of industry disaggregation that are explicitly based on an Eaton–Kortum type of model.[41] However, we check how our results are affected when using the Pareto shape parameter as a measure of heterogeneity. We obtain the Pareto parameters γ_k from two different sources and compute a measure of trade integration θ_{ij}^{Ch} as defined in Eq. (9).[42] The first source is Crozet and Koenig (2010) who estimate the Pareto shape parameter γ_k and the elasticity of substitution σ_k for 28 2-digit industries using French firm-level exports to countries that are contiguous to France. This choice of destination countries is attractive since it fits fairly well with the EU data that we use. The γ_k's vary between 1.65 and 7.31 with a mean of 3.33 while the σ_k's range from 1.15 to 6.01 with an average of 2.44. Interestingly, the correlation between the two sets of estimates is equal to 95%, indicating that the two sources of heterogeneity are strongly correlated.

When regressing the θ_{ij}^{Ch} measure so obtained, we only include 2-digit fixed effects $\tilde{K}$ because with 3-digit fixed effects the variation in θ_{ij}^{Ch} would become the same as the variation in (inverse) *phi*-ness within each 3-digit sector as there is no variation in γ_k within 3-digit industries. The sample is also reduced from 163 to 125 industries. The results, reported in column (7) of Table 2, show that both contiguity and public procurement lose significance. From a policy point of view, the variance decomposition reported in column (3) of Table 3 indicates that in addition to TBTs, joining the EU as well as distance

---

[40] The product of two balanced trade flows yields a relatively large magnitude in the denominator and thus a relatively low θ_{ij}^k. Also, in a model of heterogeneous firms such as Chaney (2008) a large number of tradable goods in a sector may indicate low average fixed costs of exporting, or a high value of the Pareto shape parameter implying a low degree of firm heterogeneity. Both these features would imply a small value of θ_{ij}^k.

[41] Although Caliendo and Parro (2011) provide industry-level estimates, their estimation is based on a gravity equation framework with tariffs as the trade cost regressor. Their coefficients therefore represent a combination of the Fréchet parameter and the tariff elasticity of trade costs.

[42] The estimation of those parameters remains outside the scope of this paper as it typically requires firm-level data. Note that the exponent of the trade flow ratio in Eq. (10) for the trade integration measure based on the Melitz and Ottaviano (2008) model, θ_{ij}^{MO}, is the same as for θ_{ij}^{Ch}.

are the most important factors in explaining the variance in trade integration. The use of the Pareto parameters estimated by Crozet and Koenig (2010) to measure trade integration therefore leads to results that are broadly similar, albeit with a few differences.

The second source is di Giovanni et al. (2011) who estimate $\nu_k = \gamma_k/(\sigma_k - 1)$ for 25 industries using French firm-level data. The mapping between those estimates and our NACE industries reduces our sample by one industry only. Using our σ_k estimates we indirectly retrieve the γ_k parameters as $\gamma_k = \nu_k(\sigma_k - 1)$ and use them to compute θ_{ij}^{Ch}. Note that the γ_k's vary across 4-digit NACE sectors thanks to the variation in the σ_k's, and they range from 0.8 to 53.8 with an average of 7.01.[43] The regression for θ_{ij}^{Ch}, reported in column (8) of Table 2, shows that our results are strongly robust to those γ_k parameters capturing heterogeneity. This consistency in the results also applies to the variance decomposition in column (4) of Table 3.

To conclude, there is evidence of a strong correlation between the two sources of heterogeneity as measured by γ_k or σ_k. It follows that measuring trade integration using either type of heterogeneity typically leads to similar results.

4.4. Comparison to phi-ness and gravity

We also illustrate the implications for empirical analysis of the conceptual difference between θ_{ij}^k on the one hand, and gravity and *phi*-ness on the other hand. We regress $\ln(x_{ij}^k x_{ji}^k)$ and $\ln\phi_{ij}^k$ on the full set of regressors we considered earlier, in which case the estimated trade cost coefficients mix the trade barrier and the heterogeneity effects.[44] Compared to θ_{ij}^k, the results remain similar but with a few exceptions (see columns 9 and 10 in Table 2). With ϕ_{ij}^k, sharing a common border, public procurement and TBTs become insignificant. With $x_{ij}^k x_{ji}^k$, not joining the Euro significantly decreases bilateral trade while VATs are insignificant.

Columns (5) and (6) in Table 3 report the variance decompositions corresponding to the regressions in columns (9) and (10) of Table 2. The policy implications are now different compared to the ones we derived from θ_{ij}^k. According to both decompositions, policymakers should not pay much attention to TBTs but should instead focus exclusively on the effects of geography and transportation costs. This finding suggests that ignoring the role of heterogeneity in measuring trade frictions may generate results that would induce policymakers not to focus on important policy areas.

4.5. Robustness

We now report a number of alternative specifications that we implement to ensure the robustness of our conclusions. The results can be found in Tables A1 and A2 which are available online.[45] First, we verify in columns (1) and (2) of Table A1 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.[46] We also want to make sure that our results are not driven by

observations that are economically small. We therefore run a regression in column (3) where the observations are weighted by the product of sectoral bilateral exports in the initial year 1999. Finally, given that Ireland and the United Kingdom are the only two countries that remain outside the Schengen Agreement but are also the only two that are separated from the other countries by a sea, we include in column (4) a control for being an island to make sure that the Schengen variable is not simply capturing the effect of the sea. The results in columns (1) to (4) remain virtually unchanged.

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 Eq. (7). Export price indices are not available, so common practice is to use domestic deflators instead (Baldwin and Taglioni, 2006). In that case, the deflators in Eq. (7) 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 considerably 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.

As our main specification includes 3-digit fixed effects, it is important to observe that the variation in θ_{ij}^k we want to explain is the same as the variation in (inverse) *phi*-ness if σ_k remains constant within a 3-digit sector.[47] In our sample, this happens to 36 industries that do not disaggregate as we move from the 3-digit to the 4-digit level (e.g., "Tobacco products" is the same at the 3-digit and 4-digit levels).[48] It also occurs in "Steel tubes" (NACE 2722) as it is the only 4-digit industry of the 3-digit NACE 272 included in the sample. To ensure that our results are not contaminated by explaining the variation in (inverse) *phi*-ness in those industries, column (6) excludes the 37 industries from the sample (i.e., 3,175 observations out of 15,145). This leaves our findings unaffected.

Besides, given that the elasticities of substitution σ_k are estimated and not observed, we also compute trade frictions θ_{ij}^k based on elasticities σ_k plus or minus one standard error. The results are reported in columns (7) and (8) and show that our findings remain unchanged.

We also regress a θ_{ij}^k measure computed using the Imbs and Méjean (2009) and the Broda and Weinstein (2006) elasticities. As their estimates display some similarities with our σ_k's, we expect the results to remain broadly similar. As the 3-digit ISIC Rev.3 groupings of Imbs and Méjean (2009) are aggregates of the 4-digit NACE Rev.1 sectors, the mapping of their elasticities from the ISIC to the NACE is pretty accurate (i.e., only 44 ISIC elasticities are required to match 159 of our 163 NACE industries, in which case the sample is reduced by 48 series). But this also means that 3-digit fixed effects cannot be included, otherwise the variation in θ_{ij}^k to be explained becomes the same as the variation in (inverse) *phi*-ness within each 3-digit sector. With 2-digit fixed effects $\tilde{K}$ instead, the regression in column (9) shows that the results remain largely unaffected.[49]

Computing θ_{ij}^k using the Broda and Weinstein (2006) elasticities is more problematic. The issue comes from the inaccurate aggregation of the 5-digit SITC Rev.3 level elasticities at the 4-digit NACE Rev.1

---

[43] We use their estimates for non-exporting firms (see their Section 4.3 for an explanation as to why non-exporting firms should be used). The $\nu_k > 1$ condition ensures that the size distribution of firms has a finite mean in equilibrium but only nine out of the 15 ν_k estimates that match with our NACE industries satisfy this condition. Restricting the sample to $\nu_k > 1$ substantially reduces the number of observations and leads to many right-hand side variables becoming insignificant. We therefore include all ν_k's to ensure the comparability between samples.

[44] We include year and 3-digit industry fixed effects in both cases. For the gravity equation, domestic outputs are added as regressors and we do not control for multilateral resistance, otherwise variables such as $cfob_t^k$ and wv_t^k would drop out. We also try a specification as in Eq. (12) with controls for multilateral resistance, dropping the domestic outputs, $cfob_t^k$ and wv_t^k variables. The policy implications from the variance decomposition remain unchanged.

[45] This is available at www.warwick.ac.uk/go/nachen.

[46] In the case of zero trade observations, we drop the entire series to keep a balanced sample over time. This results in 12 fewer observations in column (2).

[47] This explains why 4-digit fixed effects cannot be included in Eq. (16) as σ_k varies across 4-digit industries only.

[48] Those industries can be identified in Table 1 by a NACE code ending with a zero.

[49] The results remain unchanged if we further restrict the sample to the 2-digit industries within which there is some variation in σ_k. This reduces the sample by 19 industries or 895 observations.

level as the weights of each SITC product within each NACE sector are unknown (the 163 NACE industries match with 1,555 SITC level elasticities).[50] The NACE level elasticities are thus simply computed as the average of the SITC elasticities. This significantly affects the results as the estimated coefficients are sometimes less significant or change sign, and the R-squared is lower (column 10).[51]

A final concern relates to the exogeneity of lagged productivity in explaining θ_{ij}^k. We instrument lagged productivity using three instruments (also lagged by one year) which are (mostly) measured for a country that is not in our sample, i.e., the US. We believe that such data, in addition to being arguably exogenous, should provide some useful information about the productivity distribution across EU industries. First, we use sectoral real labor productivity for the US. Second, as increasing returns to scale would suggest a positive effect of firm size on productivity, we consider average firm size for US industries (missing observations reduce the sample by 14 series). Third, we compute the US shares of R&D spending in sectoral value-added. In order to instrument the cross-country variation in productivity, we further interact the US shares of R&D spending with aggregate R&D spending as a proportion of GDP for each of the EU partners. Overall, the first-stage regression explains up to 64% of the variation in sectoral EU productivity, and column (11) shows that our results remain robust to Instrumental Variables.[52]

5. Concluding remarks

This paper explores the sources and size of trade barriers across countries and industries. For this purpose we employ the monopolistic competition gravity framework by Anderson and van Wincoop (2004) with heterogeneous trade costs and heterogeneous elasticities of substitution across industries. The model yields a micro-founded measure of bilateral trade integration that nets out multilateral resistance terms, controls for cross-industry heterogeneity and can be computed as a function of observable trade and output data only. We show that this trade integration measure is also consistent with other classes of models including the Ricardian model by Eaton and Kortum (2002) and recent heterogeneous firms models.

In the empirical part of the paper, we use data for 163 manufacturing industries across the European Union for the years 1999 to 2003. We first estimate the elasticities of substitution across industries since these are required to construct the trade integration measure. We find a substantial degree of heterogeneity across industries both for the substitution elasticities and the degree of trade integration. For example, trade integration appears particularly low for industries characterized by high transportation costs such as "Bricks," "Plaster" and "Cement." On the contrary, those industries where trade costs do not appear to be an important consideration include a number of high-tech industries such as "Aircraft and spacecraft," "Engines and turbines" or "Computers."

We also relate our trade integration measure to observable trade cost proxies. In particular, we show that cross-country trade integration is lower for those countries that joined the EU most recently and that have not yet implemented the Schengen Agreement that abolishes physical border controls. Most importantly, we also document a considerable variation in trade integration across industries. 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. For public policy our findings suggest that gains from improved international trade integration are possible especially through the elimination of TBTs.

Of course, ultimately 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 trade impediments at the firm or product level. In our view, analyzing trade barriers at these very fine levels of disaggregation would be a natural and important next step.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10. 1016/j.jinteco.2011.07.005.

References

Anderson, J.E., 1979. A theoretical foundation for the gravity equation. American Economic Review 69 (1), 106–116.
Anderson, J.E., van Wincoop, E., 2002. Borders, trade, and welfare. In: Collins, S.M., Rodrik, D. (Eds.), Brookings Trade Forum. Brookings Institution, pp. 207–244.
Anderson, J.E., van Wincoop, E., 2003. Gravity with gravitas: A solution to the border puzzle. American Economic Review 93 (1), 170–192.
Anderson, J.E., van Wincoop, E., 2004. Trade costs. Journal of Economic Literature 42 (3), 691–751.
Baldwin, R.E., 2000. Regulatory protectionism, developing nations, and a two-tier world trade system. In: Collins, S.M., Rodrik, D. (Eds.), Brookings Trade Forum. Brookings Institution, pp. 237–293.
Baldwin, R.E., 2006. The Euro's trade effects. European Central Bank Working Paper 594.
Baldwin, R.E., Taglioni, D., 2006. Gravity for dummies and dummies for gravity equations. NBER Working Paper 12516.
Baldwin, R.E., Forslid, R., Martin, P., Robert-Nicoud, F., 2003. The core-periphery model: Key features and effects. In: Heijdra, B., Brakman, S. (Eds.), The Monopolistic Competition Revolution in Retrospect. Cambridge University Press, pp. 213–235.
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., Schott, P.K., 2006. Trade costs, firms and productivity. Journal of Monetary Economics 53 (5), 917–937.
Broda, C., Weinstein, D.E., 2006. Globalization and the gains from variety. Quarterly Journal of Economics 121 (2), 541–585.
Caliendo, L., Parro, F., 2011. Estimates of the Trade and Welfare Effects of NAFTA. Princeton University, mimeo.
Chaney, T., 2008. Distorted gravity: The intensive and extensive margins of international trade. American Economic Review 98 (4), 1707–1721.

---

[50] The matching is achieved using correspondence tables from the SITC Rev.3 to the ISIC Rev.3 and from the ISIC Rev.3 to the NACE Rev.1 level.
[51] Hummels (2001a) uses data on bilateral trade flows, import tariffs and transport costs to estimate substitution elasticities at the 2-digit SITC Rev.3 level. Our results also remain robust to the use of those elasticities in computing θ_{ij}^k.
[52] According to Staiger and Stock's (1997) "rule of thumb" in the case of one endogenous variable to be instrumented, we reject the null hypothesis of weak instruments as the F-statistic on the excluded instruments in the first stage is larger than 10 (and equal to 132). The results of the first stage are available from the authors upon request.

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.

Combes, P.-P., Lafourcade, M., 2005. Transport costs: Measures, determinants, and regional policy implications for France. Journal of Economic Geography 5 (3), 319–349.

Crozet, M., Koenig, P., 2010. Structural gravity equations with intensive and extensive margins. Canadian Journal of Economics 43 (1), 41–62.

Davies, S., Lyons, B., 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: Frankel, J.A. (Ed.), The Regionalization of the World Economy. University of Chicago Press, pp. 7–22.

di Giovanni, J., Levchenko, A., Rancière, R., 2011. Power laws in firm size and openness to trade: Measurement and implications. Journal of International Economics 85 (1), 42–52.

Dunlevy, J.A., 2006. The influence of corruption and language on the protrade effects of immigrants: Evidence from the American states. Review of Economics and Statistics 88 (1), 182–186.

Eaton, J., Kortum, S., 2002. Technology, geography and trade. Econometrica 70 (5), 1741–1779.

European Commission, 2000. EU transport in figures — Statistical pocket book. Directorate General for Energy and Transport.

Evans, C.L., 2003. The economic significance of national border effects. American Economic Review 93 (4), 1291–1312.

Evans, C.L., Harrigan, J., 2005. Distance, time, and specialization: Lean retailing in general equilibrium. American Economic Review 95 (1), 292–313.

Feenstra, R.C., 1994. New product varieties and the measurement of international prices. American Economic Review 84 (1), 157–177.

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., Kneller, R., 2008. Exporting, productivity and agglomeration. European Economic Review 52 (5), 919–939.

Grossman, G.M., 1998. Comment. In: Frankel, J.A. (Ed.), The Regionalization of the World Economy. University of Chicago Press, pp. 29–31.

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., 2010. Airplanes and comparative advantage. Journal of International Economics 82 (2), 181–194.

Head, K., Mayer, T., 2000. Non-Europe: The magnitude and causes of market fragmentation in the EU. Weltwirtschaftliches Archiv 136 (2), 285–314.

Head, K., Mayer, T., 2004. The empirics of agglomeration and trade. Handbook of Regional and Urban Economics. Chapter 59, North-Holland.

Head, K., Ries, J., 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., Rubinstein, Y., 2008. Estimating trade flows: Trading partners and trading volumes. Quarterly Journal of Economics 123 (2), 441–487.

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.

Imbs, J., Méjean, I., 2009. Elasticity optimism. CEPR Discussion Paper 7177.

Kee, H.L., Nicita, A., Olarreaga, M., 2009. Estimating trade restrictiveness indices. Economic Journal 119, 172–199.

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

Limão, N., Venables, A.J., 2001. Infrastructure, geographical disadvantage and transport costs. World Bank Economic Review 15, 451–479.

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., Ottaviano, G.I.P., 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., 2011. Gravity Redux: Measuring International Trade Costs with Panel Data. University of Warwick, mimeo.

Pesaran, H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74 (4), 967–1012.

Rauch, J.E., 1999. Networks versus markets in international trade. Journal of International Economics 48 (1), 7–35.

Rauch, J.E., 2001. Business and social networks in international trade. Journal of Economic Literature 39 (4), 1177–1203.

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., Tenreyro, S., 2006. The log of gravity. Review of Economics and Statistics 88 (4), 641–658.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557–586.

Wei, S.-J., 1996. Intra-national versus international trade: How stubborn are nations in global integration? NBER Working Paper 5531.