Working Papers
www.cesifo.org/wp

Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation

Jeffrey H. Bergstrand
Mario Larch
Yoto V. Yotov

CESIFO WORKING PAPER NO. 4502
CATEGORY 8: TRADE POLICY
DECEMBER 2013

CESifo *Working Paper No. 4502*

Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation

Abstract

Using a novel common econometric specification, we examine the measurement of three important effects in international trade that historically have been addressed largely separately: the (partial) effects on trade of economic integration agreements, national borders, and bilateral distance. First, recent studies focusing on precise and unbiased estimates of effects of economic integration agreements (EIAs) on members' trade may be biased upward owing to inadequate control for exogenous unobservable country-pair-specific technological innovations (decreasing the costs of international relative to intranational trade); we find evidence of this bias using a properly specified gravity equation. Second, our novel methodology yields economically plausible and statistically significant estimates of the declining effect of "national borders" on world trade, now accounting for endogenous EIA formations and unobserved country-pair heterogeneity in initial levels. Third, we confirm recent evidence providing a solution to the "distance-elasticity puzzle," but show that these estimates of the declining effect of distance on international trade are biased upward by not accounting for endogenous EIA formations and unobserved country-pair heterogeneity. We show that these results are robust to a battery of sensitivity analyses allowing for phase-ins of agreements, lagged terms-of-trade effects, reverse causality, various estimation techniques, disaggregation, inclusion of intranational trade, and accounting for firm-heterogeneity and country-selection bias.

JEL-Code: F130, F140, F150.

Keywords: international trade, economic integration agreements, gravity equations.

Jeffrey H. Bergstrand
University of Notre Dame
USA – Notre Dame, Indiana 46556
bergstrand.1@nd.edu

Mario Larch
University of Bayreuth
Bayreuth / Germany
mario.larch@uni-bayreuth.de

Yoto V. Yotov
Drexel University
USA – Philadelphia, PA 19104
yotov@drexel.edu

November 12, 2013

1 Introduction

It's a Small World After All.... (Walt Disney, New York World's Fair, 1964)

Using a novel common econometric specification, we examine the measurement of three important effects on international trade flows that have historically been addressed largely separately: the (partial) effects on trade of economic integration agreements (EIAs), national borders, and bilateral distance.[1] First, one of the most prominent aspects of the global economy over the past 20 years has been the proliferation of economic integration agreements (EIAs) – notably free trade agreements but also some customs unions. Policy makers at national and supra-national government levels increasingly rely on *ex post* estimates of the (partial) effects of EIAs on trade flows based upon gravity equations to evaluate subsequently the welfare effects of EIAs, cf., Berden, Francois, Tamminen, Thelle, and Wymenga (2010).[2] Only recently have economists been able to provide more precise and unbiased *ex post* estimates of the (partial) effects of EIAs on members' international trade flows, in contrast to the highly variable and often economically implausible estimates generated over 45 years from 1962-2007.[3] Using panel data and accounting for the endogeneity of EIAs and prices and for unobserved country-pair heterogeneity, Baier and Bergstrand (2007), or BB, found using a sample spanning 1960-2000 and ordinary least squares (OLS) that a typical EIA increases two members' aggregate goods bilateral trade about 100 percent after 10-15 years – five times the effect estimated using atheoretical gravity equations. Recently, Anderson and Yotov (2011) found similar results using the same BB specification (but using a Poisson quasi maximum likelihood (PQML) estimator) and showed the method also generated eco-

[1] We are concerned in this study only with estimating partial (or direct) effects, not general equilibrium effects as in Anderson and van Wincoop (2003), Baier and Bergstrand (2009), Anderson and Yotov (2011), and Bergstrand, Egger, and Larch (2013).

[2] In an important recent paper, Arkolakis, Costinot, and Rodriguez-Clare (2012) note the importance of gravity equations for estimating trade-cost elasticities, which are then used for calculating welfare gains in several quantitative trade models.

[3] In a meta-analysis of 1,827 earlier studies (including several using flawed specifications), Cipollina and Salvatici (2010) find a range of estimates between 12 percent and 285 percent. Their mean effect is 80 percent and median effect is 46 percent.

nomically plausible, precise, and statistically significant effects for disaggegate trade flows. The key in BB was accounting for unobserved heterogeneity in exporters' and importers' time-varying multilateral influences (such as countries' prices and GDPs) and for unobserved heterogeneity in time-invariant bilateral influences. However, both of these studies failed to account for possible exogenous unobservable *time-varying bilateral* influences owing to technological innovations that likely decreased bilateral variable and fixed costs of international relative to intranational trade and may have resulted in estimates of EIAs' effects being biased upward. In this paper, we address this potentially important shortcoming using a properly specified gravity equation motivated by formal theoretical foundations. In doing so, we also contribute to two related literatures: "(national) border effects" and the "distance-elasticity puzzle."

The "border puzzle" refers to the seminal estimate using traditional atheoretical gravity equations in McCallum (1995) that the Canada-U.S. international border caused Canadian inter-province trade to be *22 times* – or 2100 percent greater than – province-state international trade in 1988, other things equal. This result implied that national borders imposed dramatic costs on international relative to intranational trade. This finding inspired an entire literature, including Anderson and van Wincoop's seminal (2003) paper formulating a new theoretical foundation for the gravity equation, building upon formal foundations in Anderson (1979) and Bergstrand (1985). While Anderson and van Wincoop (2003) addressed the importance of accounting properly for endogenous "prices" (in their terms, "multilateral resistances") in estimation and in general equilibrium comparative statics, to date estimates of the border effect are still very large. For instance, de Sousa, Mayer, and Zignago (2012) report that on average a country traded *493 times* more intranationally than internationally in 1990, even dwarfing the McCallum estimate. Moreover, they estimate that on average this effect fell 63 percent to 181 in 2002, that is, in only 12 years. However, using cross-sectional data, they did not control of unobserved country-pair heterogeneity in border effects, did not account for endogenous EIAs as in BB, and, while recognizing multilateral prices in their estimation, did not account for the endogeneity of prices as addressed in Anderson and van Wincoop (2003). In this paper, we use an enhanced version of the BB panel-data methodology to provide economically plausible, consistent, and precise estimates of the average declining effects of national borders on international trade, using a properly specified gravity equation accounting also for the effects of endogenous EIA formations, endogenous prices, and unobserved country-pair heterogeneity in initial border effects.

The "distance-elasticity puzzle" refers to the issue that – despite widespread anecdotal evidence that the effect of distance on international trade is declining over time, as suggested by Thomas Friedman's "flatter world" – systematic academic empirical evidence suggests that the distance elasticity of bilateral international trade has *not* declined, as established in the Disdier and Head (2008) meta analysis of the distance elasticity.[4] While some authors have offered alternative explanations, they have met mixed success (cf., footnote 1 in Yotov (2012) and our discussion later). However, Yotov (2012) recently provided a persuasive solution to the distance-elasticity puzzle by recognizing the importance of including *intranational*, alongside international, trade flows and bilateral distances in estimation (using OLS and PQML), a feature actually common to the "border effects" literature, cf., Wei (1996). By typically excluding intranational trade flows and intranational distances, gravity equation estimates cannot identify the impact on international trade of international trade costs *relative to* intranational trade costs; previous studies of the distance-elasticity puzzle ignored this. However, Yotov (2012) suffered from two shortcomings. The study did not account for unobserved heterogeneity across country pairs and omitted controls for EIAs, potentially biasing upward his estimates of the declining effect of distance. Recently, Bosquet and Boulhol (2013) using PQML included country-pair fixed effects to account for country-pair heterogeneity and exporter-time and importer-time fixed effects, but could not find declining distance elasticities; however, they did not allow for intranational trade and distances. In this paper, we address all these shortcomings and find economically plausible

[4]The national border puzzle differs from the distance-elasticity puzzle in the following respect. Typically, the border puzzle is associated with arguably economically implausible estimates of the *level* effect of an international border on international trade flows. By contrast, the distance-elasticity puzzle is only concerned with an absence of declines in the distance elasticity of international trade, not the average *level* of the distance elasticity *per se*.

and statistically significant *declining* distance elasticities that are robust to several sensitivity analyses, including alternative estimators, and that indicate the upward bias in estimates in Yotov (2012).

This is the first paper to address all three related – but historically often disjointed – issues using a common econometric framework.[5] Using a state-of-the-art gravity equation, we reconcile methodologically *and empirically* all three issues. We provide three potential contributions. First, some bilateral fixed and variable export costs are unobservable, such as bilateral costs of information and technology (IT). *If* some IT-based bilateral fixed and variable export costs are declining over time – as anecdotal evidence suggests – the Melitz (2003) model suggests that aggregate bilateral trade of existing exporters should expand (i.e., intensive margin response to lower variable trade costs), some previously domestic firms should select into exporting (i.e., extensive margin response to lower fixed and variable trade costs), and the number of domestic firms should decrease (due to market competition). Hence, omission of such variables that account for some of the increase in international relative to intranational trade – other than an EIA – could bias upward EIA coefficient estimates under the approach in BB. Thus, the resounding "yes" claimed in BB in response to their question, "Do free trade agreements actually increase members' international trade?", may have been premature, as this issue of declining bilateral costs of IT was ignored.[6] In this paper, we use three alternative methods to account more fully for *time-varying* unobserved heterogeneity in bilateral international trade costs relative to intranational trade costs. We find that previous estimates of EIAs' partial effects have been *biased upward*. In fact, when using the econometrically preferred PQML estimator the effects of an EIA are nearly halved.

[5]As we will clarify later, we note that studies of the border puzzle typically do not address the distance-elasticity puzzle, and vice versa. Also, some studies of the border puzzle and some of the distance-elasticity puzzle include dummy variables for economic integration agreements, but typically do not examine in detail how EIAs' effects are sensitive to the specifications. This paper addresses all three issues using a unified framework.

[6]Unfortunately, there is little systematic evidence of *observed* declines in bilateral IT costs causing larger trade. A few studies have explored the issue of declining IT costs, cf., Freund and Weinhold (2004), Tang (2006), and Berthelon and Freund (2008). However, all these studies include time-varying *multilateral*, not bilateral, indexes of IT factors. Such multilateral factors will be accounted for in our estimation using exporter-year and importer-year fixed effects. No study has yet accounted for time-varying declines in *bilateral* trade costs, except those associated with EIAs or currency unions.

Second, we draw upon the notion used in Anderson and van Wincoop (2003), or AvW, that national border dummies imbed international trade costs relative to intranational trade costs in gravity equations. In the cross-sectional context of AvW, $BORDER_{ij}$ was a dummy variable that measured 1 if two sub-national regions (Canadian province or U.S. state) were from different countries, and 0 otherwise (and hence from the same country). Thus, $BORDER_{ij}$ was an exogenous index of whether the trade flow was an international versus intranational flow. Here, we use a panel data set of international *and intranational* trade flows for a large number of country pairs for a large number of years. Since intranational trade is a nation's gross output less exports, we confine our analysis to manufactures trade, since exports are measured on a gross basis and we have data on manufactures gross output.[7] Similar to AvW, we then construct an exogenous dummy variable $BORDER_{ij}$ that assumes the value 1 if the source (i) and destination (j) countries are different ($i \neq j$) and the value 0 if i and j are the same ($i = j$). By incorporating this variable interacted with a set of year dummies – creating $BRDR_{ij,t}$, $BRDR_{ij,t+1}$, etc. – and then using the BB panel approach, we can isolate the effect of EIAs on bilateral trade to determine how much an EIA actually increases two members' trade, but now accounting for any trends in declining bilateral (fixed and variable trade) IT costs that have increased international relative to intranational trade. Moreover, the coefficient estimates for the multiple $BRDR_{ij,t}$ dummies also provide direct estimates of the changing (partial) effect of an international border on a pair's trade flow. A novel aspect of our approach – accounting for unobserved country-pair heterogeneity in a panel – is to allow the *level* of the border effect to vary across every country pair using pair fixed effects (or pair fixed effects interacted with a trend); by contrast, previous studies constrain the level border effect to be identical across all country pairs or certain groups of country pairs using cross-sectional data (excluding country-pair fixed effects).[8] We find direct estimates of the (average) falling partial effect of a national border (after accounting

[7] We will focus on aggregate and disaggregate manufactures trade. In one robustness analysis, we will look at aggregate merchandise trade allowing GDP (a value added measure) to be an imperfect proxy for gross output. As we will discuss later, a few studies in the border-effects literature, starting with Wei (1996), have incorporated correctly intranational trade.

[8] The level of a country-pair's border effect is imbedded in our fixed effect estimate.

for EIAs) – using a specification motivated by a formal theoretical foundation for the gravity equation and avoiding the endogeneity bias (attributable to endogenous prices and endogenous EIA formations) present in several studies.[9] One of our estimates suggests that the cost of a national border (in terms of trade flows) has declined on average by an economically plausible 27.6 percent from 1990 to 2002, or about 2.7 percent per year.

Third, we will provide a battery of sensitivity analyses to determine the robustness of the results to phase-ins of agreements, lagged terms-of-trade effects, reverse causality, alternative measures to our border dummy variable, various estimation techniques, and accounting for firm-heterogeneity and country-selection biases introduced potentially by using aggregated data. In one sensitivity analysis, our alternative measure to our border dummies to control for declining trade costs is an interaction of bilateral distance with year dummies. These results confirm the findings in Yotov (2012) that the effect of bilateral distance – owing to likely falling bilateral variable and fixed export costs – *is* declining over time. Like Yotov (2012), we use a more appropriate measurement of intranational trade (using manufactures gross output and (gross) exports rather than aggregate trade and GDPs). However, unlike Yotov (2012) our specification accounts for unobserved time-varying bilateral heterogeneity and for the effects of EIA formations, which potentially biased upward Yotov's estimates of the declining effect of distance. We find evidence of this upward bias and estimate that the effect of distance on international trade has fallen by an economically plausible 1.2 percent per year, which still may have had a substantive effect on growth rates of nations' total factor productivity.[10]

The remainder of the paper is as follows. Section 2 provides theoretical background for the estimating equation that will be used. Section 3 addresses econometric issues and provides a data description. Section 4 discusses the empirical results for EIAs' partial effects on total bilateral manufactures trade flows, including the results of a series of sensitivity

[9] For instance, Head and Mayer (2000) and de Sousa, Mayer, and Zignago (2012) accounted for relative prices using measured national prices, but not for the endogeneity of prices as raised by Anderson and van Wincoop (2003).

[10] This contrasts with the stable estimates of the distance elasticity over time using PQML in Bosquet and Boulhol (2013), likely due to their omission of intranational trade and distances in their sample.

analyses, and provides economically plausible estimates of the declining effect of national borders on trade. Section 5 discusses the results for disaggregate manufactures trade flows. Section 6 evaluates the sensitivity of the results to using aggregate *goods* trade flows and a longer time series. Section 7 introduces an alternative variable to account for declining trade costs other than EIAs, and provides economically plausible estimates of the declining effect of distance on international relative to intranational trade. In section 8, we conclude that the specifications suggested six years ago in Baier and Bergstrand (2007) to account for endogenous EIAs can be substantively improved by including border dummies (or distances interacted with year dummies) in panel specifications to account for systematic declines over time in *unobserved time-varying bilateral* costs of international relative to intranational trade – while *simultaneously* accounting for the heteroskedasticity bias in OLS estimates, requiring inclusion of intranational trade and distance in samples, and accounting for unobserved time-invariant heterogeneity in country-pair border level effects. In short, just as BB showed six years ago that panel techniques along with a properly specified gravity equation were critical to finding economically plausible, unbiased, and precise EIA estimates, this paper is designed to show that an enhanced version of BB using panel techniques are critical to finding economically plausible, unbiased, and precise estimates of EIA effects on international trade, of the declining effect of national borders on trade, and of the declining distance-elasticity of international trade.

2 Motivating the Gravity-Equation Specification

The gravity equation has become the empirical workhorse for estimating partial effects of EIAs on members' trade flows. Recently, Arkolakis, Costinot, and Rodriguez-Clare (2012) demonstrated that a gravity equation surfaces for a large class of "quantitative trade models" that feature four main assumptions: (1) Dixit-Stiglitz preferences; (2) one factor of production (typically, labor): (3) linear cost functions; and (4) perfect or monopolistic competition.[11] Trade models satisfying these four assumptions are Armington (cf., Anderson and van Wincoop (2003)), Ricardian (cf., Eaton and Kortum (2002)), Krugman (1980), and Melitz (2003). Arkolakis, Costinot, and Rodriguez-Clare (2012) concluded that the gravity equation provides a common method for estimating the trade elasticity across these different approaches.

Arkolakis, Costinot, and Rodriguez-Clare (2012) show, for instance, that the Melitz model yields a theoretical gravity equation of the form:

$$X_{ij,t}^m = N_{i,t}^m Y_{j,t}^m \left(\frac{(a_{Li,t}^m)^{-\gamma^m} w_{i,t}^{-\gamma^m} \tau_{ij,t}^{-\gamma^m} f_{ij,t}^{-[\gamma^m/(\sigma^m-1)-1]}}{\sum_{k=1}^K N_{k,t}^m (a_{Lk,t}^m)^{-\gamma^m} w_{k,t}^{-\gamma^m} \tau_{kj,t}^{-\gamma^m} f_{kj,t}^{-[\gamma^m/(\sigma^m-1)-1]}} \right), \tag{1}$$

where $X_{ij,t}^m$ is the trade flow from exporter i to importer j in year t in "good" (industry) m, $N_{i,t}^m$ is the number of firms in i (exporting and non-exporting) that produce (differentiated) products in good m, $Y_{j,t}^m$ is the expenditure in j on good m, $a_{Li,t}^m$ is the lower bound of the Pareto distribution of productivities in m in i, γ^m is an index of productivity heterogeneity among firms in good m, $w_{i,t}$ is the wage rate in i, $\tau_{ij,t}$ is variable trade costs of country i's products into j, $f_{ij,t}$ is fixed export costs from i to j, σ^m is the elasticity of substitution in consumption, and $\gamma^m > \sigma^m - 1$.[12] Note that the term in large parentheses is a standard representation of relative prices in the gravity equation, but now also reflecting productivity heterogeneity (through γ^m) and fixed exporting costs ($f_{ij,t}$). In the context of these models, variable trade costs, $\tau_{ij,t}$, affect $X_{ij,t}^m$ via both the intensive and extensive margins, but fixed export costs, $f_{ij,t}$, affect trade via the extensive margin. As Chaney (2008) demonstrates in one Melitz-type model, $\gamma^m = (\sigma^m - 1) + [\gamma^m - (\sigma^m - 1)]$, where $\sigma^m - 1$ represents the intensive-margin elasticity of variable trade costs whereas $\gamma^m - (\sigma^m - 1)$ is the extensive-margin elasticity of variable trade costs.

For the purposes of this paper, the variables of interest are $\tau_{ij,t}$ and $f_{ij,t}$. Typically,

[11] Arkolakis, Costinot, and Rodriguez-Clare (2012) also note three macro-level restrictions: (1) trade is balanced; (2) aggregate profits are a constant share of aggregate revenues; and (3) the import demand system is CES.

[12] For finite means in the theory, $\gamma^m/(\sigma^m - 1)$ must exceed 1. We assume the case where fixed export costs are paid by importers, i.e., the case of $\mu = 0$ in Arkolakis, Costinot, and Rodriguez-Clare (2012), equation (23).

researchers have assumed that the formation of an EIA (such as a free trade agreement) between i and j lowers $\tau_{ij,t}$. However, EIAs are broad agreements reaching beyond elimination of tariff rates and variable trade costs; they likely also lower fixed export costs, $f_{ij,t}$. Yet, in reality, advances in IT likely also lower $\tau_{ij,t}$ and $f_{ij,t}$. Thus, the use of *time-invariant* pair-specific fixed effects, as in BB, is insufficient to isolate an unbiased partial effect of an EIA's formation on trade (via lowering $\tau_{ij,t}$ and $f_{ij,t}$) because $\tau_{ij,t}$ and $f_{ij,t}$ may be influenced also by falling IT costs.

Moreover, as noted in Arkolakis, Costinot, and Rodriguez-Clare (2012) and Yotov (2012), every theoretical quantitative trade model yielding a gravity equation embodies "intranational trade" ($X_{ii,t}$), i.e., a country's domestic spending on its own products. Our novel approach in this paper – using international *and* intranational trade flows – is to introduce a variable $BRDR_{ij,t}$ to account for average (across all pairs of different countries) trend declines in unobservable bilateral (fixed and variable) international trade costs relative to intranational trade costs (unassociated with EIAs), as well as account for unobservable country-pair-specific trend declines in bilateral trade costs (using a random trend econometric model), but in the context of a properly specified gravity equation motivated by a formal theoretical foundation. $BRDR_{ij,t}$ is defined as the product of a year dummy, D_t, and a binary variable, $BORDER_{ij}$, which assumes the value 1 if the source and destination countries, i and j, respectively, are different countries ($i \neq j$) and the value 0 if i and j are the same country ($i = j$). The economic motivation is explained most easily by considering AvW.[13] First, consider AvW's cross-sectional context of trade flows between and among Canadian provinces and U.S. states in 1993. In the context of AvW's Armington framework, trade costs were determined by two variables, bilateral distance ($DISTANCE_{ij}$) and a dummy ($BORDER_{ij}$) for whether the two regions were in different countries (=1, if $i \neq j$) or the same country (=0, if $i = j$). In their paper, they used non-linear least squares to estimate their gravity equation to account for endogenous non-linear multilateral price terms. However, since then most researchers, such as Feenstra (2004), have focused on consistent

[13] An excellent treatment of this model and these issues is in chapter 5 of Feenstra (2004).

estimation of the bilateral border dummy and bilateral distance coefficient estimates using a specification such as:

$$\ln X_{ij} = \beta_0 + \beta_1 \ln DISTANCE_{ij} + \beta_2 BORDER_{ij} + \eta_i + \theta_j + \epsilon_{ij}. \tag{2}$$

In AvW's cross-sectional context, $BORDER_{ij}$ captures any factor influencing international relative to intranational trade. A national border imposes considerable costs. Thus, $BORDER_{ij}$ would capture any cross-sectional variation in *bilateral* trade costs, beyond the role of bilateral distance which is also present in equation (2).[14] Of course, many other factors influence bilateral flows (international or intranational), so it is feasible to replace $BORDER_{ij}$ with a country-pair fixed effect (γ_{ij}). This is a novel aspect of our approach as previous border-effect studies have *not* used country-pair fixed effects to allow for variation across pairs in the *level* border effect.[15] However, in a cross section, both $\ln DISTANCE_{ij}$ and $BORDER_{ij}$ would be perfectly correlated with these fixed effects, and so could not be included. In the absence of country-pair fixed effects in AvW, $BORDER_{ij}$ has a negative effect on trade flows. Thus, $BORDER$'s coefficient estimate was interpreted as the "cost" of an international border (or of international relative to intranational trade).

Our approach is to estimate each of the following equations using a panel of international and intranational trade flows:

$$X_{ij,t} = \exp[\beta_0 + \beta_1 EIA_{ij,t} + \beta_2 BRDR_{ij,t} + \eta_{i,t} + \theta_{j,t} + \gamma_{ij}] + \epsilon_{ij,t}, \tag{3}$$

[14] As noted earlier, any observable or unobservable multilateral IT factors would be accounted for in η_i and in θ_j.

[15] It is important to note that some previous border-effect studies appropriately accounted for intranational as well international trade flows, such as Wei (1996), Head and Mayer (2000), Fontagne, Mayer, and Zignago (2005), and de Sousa, Mayer, and Zignago (2012). However, all such studies used a cross-sectional approach (for multiple years), including typical bilateral variables such as distance but constraining all country-pairs to have the same border-effect level. Our panel approach using country-pair fixed effects (and alternatively such effects interacted with a time trend) accounts for unobserved heterogeneity across country-pairs in initial border-effect levels. Moreover, by using our panel approach, we also account for the endogeneity of EIAs. Finally, all four studies noted above (except Wei (1996)) included variables representing prices, but did not account for the endogeneity of prices, i.e., the AvW critique. Wei (1996) included a linear approximation of the two countries' multilateral prices, but the approximation was not complete because it used only bilateral distances, cf., Baier and Bergstrand (2009).

and alternatively

$$X_{ij,t} = \exp[\beta_0 + \beta_1 EIA_{ij,t} + \beta_2 BRDR_{ij,t} + \eta_{i,t} + \theta_{j,t} + (\gamma_{ij} \times Trend)] + \epsilon_{ij,t}, \qquad (4)$$

where $EIA_{ij,t}$ is a dummy assuming the value 1 (0) if an EIA exists (does not exist) in year t between countries i and j. Let exp denote the exponentiated value of the term in brackets and, for now, we allow the error term to enter additively. This specification allows for estimation in levels using Poisson estimators and allows for zeros in trade.[16] As discussed in BB, $\eta_{i,t}$ captures all time-varying multilateral factors of exporting country i, such as – in the context of equation (1) above (and suppressing superscripts m) – $N_{i,t}$, $a_{Li,t}$, and $w_{i,t}$ (and including any multilateral IT factors in i). $\theta_{j,t}$ captures all time-varying multilateral factors of importing country j, such as $Y_{j,t}$ and the denominator of the relative price term in equation (1) (and including any multilateral IT factors in j).[17] Accordingly, our paper addresses the issue raised in AvW that direct inclusion of price variables creates potential endogeneity bias.[18] The inclusion of time-invariant country-pair fixed effects γ_{ij} captures all time-invariant factors that might otherwise be picked up by $EIA_{ij,t}$. However, there may exist trends over time in the effects of unobserved bilateral heterogeneity. Consequently, following Wooldridge (2010), in equation (4) we interact the bilateral fixed effects (γ_{ij}) with a time trend ($Trend$), known as a "random trend" model.[19]

Yet, as equation (1) reveals, $\tau_{ij,t}$ and $f_{ij,t}$ are time-varying and reflect both policy-based and "natural" trade costs (such as falling bilateral IT costs) influencing international relative to intranational trade. However, there is a way to account for time-varying changes in these bilateral IT costs, separate from policy-based trade liberalizations such as EIA formations. The variable $BRDR_{ij,t}$ is defined as $BRDR_{ij,t} = D_t$ x $BORDER_{ij}$, where D_t is a year dummy. In the presence of time-invariant pair fixed effects γ_{ij}, variation in $BRDR_{ij,t}$ will

[16] We address these specification issues later.

[17] It is important to note that reductions in MFN tariff rates in importer j would be accounted for by $\theta_{j,t}$ as well.

[18] See footnote 15.

[19] Note that inclusion of an ij fixed effect *for each year* is infeasible; it would perfectly predict trade flows.

capture all bilateral factors influencing international relative to intranational trade over time *on average* relative to the base period (hence, deviations over time relative to the pair fixed effect). Thus, any time-varying pair-specific variable such as $EIA_{ij,t}$ will capture only the effects on trade over time associated with the EIA's formation, and not other factors causing $\tau_{ij,t}$ and $f_{ij,t}$ to decline over time. Consequently, the addition of $BRDR_{ij,t}$ alongside incorporating *intranational* trade flows and distances – consistent with theoretical foundations for the gravity equation explaining intranational as well as international flows – to the otherwise identical specifications in BB will essentially "purge" the partial EIA effects estimated in BB of omitted variables bias caused by general (or average) declines in $\tau_{ij,t}$ and $f_{ij,t}$ (unassociated with trade policies). Moreover, the country-pair fixed effect γ_{ij} will capture the average trade-depressing effect of an international border, allowed to vary across all pairs. If trade costs unassociated with EIAs are falling over time, raising international relative to intranational trade, $BRDR_{ij,t}$ will have a positive coefficient estimate.

3 Econometric Issues and Data Description

3.1 Econometric Issues

The previous section dealt with many econometric issues, especially concerning the definition of our key variable, $BRDR_{ij,t}$. However, one issue omitted above was the estimation approach. Historically, gravity equations have been estimated using ordinary least squares (OLS). The original analysis of the Canadian-U.S. "border puzzle" in McCallum (1995) used ordinary least squares (OLS). Baier and Bergstrand (2007) used OLS. Because of the introduction of a two-equation "structural" gravity model in AvW where one of the equations was non-linear, AvW used non-linear least squares. But with most of the literature focusing first on estimating the partial (or direct) effect of a border, most cross-section estimates have used OLS employing exporter and importer fixed effects (cf., Feenstra (2004)) and recently panel estimates have used OLS employing exporter-year, importer-year, and country-pair fixed effects (cf., BB).

However, a large sub-literature of the gravity equation, starting with Felbermayr and Kohler (2006) and Santos Silva and Tenreyro (2006), has addressed the importance of zeros in international trade flows, cf., Head and Mayer (2013). Moreover, Santos Silva and Tenreyro (2006) showed using empirical specifications and a Monte Carlo analysis that, even in the absence of zeros, log-linear estimates of gravity equations suffered from heteroskedasticity bias (owing to Jensen's inequality). They showed that a Poisson quasi-maximum-likelihood (PQML) estimator could eliminate this heteroskedasticity bias as well as allow for inclusion of zeros. Consequently, as in many recent studies, we prefer the PQML estimator for equations (3) and (4) above. However, we will show that our results are qualitatively the same using (less-preferred) OLS.[20]

Summarizing, the key features of our specifications, for which previous analyses have excluded at least one dimension, are:

(1) Exporter-year and importer-year fixed effects to account for endogenous prices and unobserved time-varying exporter and importer multilateral heterogeneity;

(2) Country-pair fixed effects or country-pair fixed effects interacted with a time trend to account for unobserved time-invariant or time-varying, respectively, bilateral effects, including pair-specific initial border effect levels;

(3) Intranational as well as international trade flows and bilateral distances, so that the border dummies can account for average declining international *relative to* intranational bilateral trade costs; and

(4) PQML estimation to account for heteroskedasticity bias, owing to Jensen's inequality.

[20]See the recent survey of the gravity equation literature in Head and Mayer (2013) for a useful discussion of new directions on these estimation issues. See also Santos Silva and Tenreyro (2011) on the robustness of PQML.

3.2 Data Description

Unlike the original estimates in BB which examined aggregate bilateral goods trade flows, our analysis here focuses on manufactures trade flows.[21] The reason is that the key RHS variable, $BRDR_{ij,t}$, captures the effect *over time* of the (likely declining) average cost of international relative to intranational trade. Hence, as in McCallum (1995) and AvW, the LHS variable needs to include observations on *intranational* trade. Since exports are measured on a "gross" (not value added) basis, national output needs to be measured on a comparable basis to estimate intranational trade. The data used are the sectoral manufacturing data from Anderson and Yotov (2011). These data cover 41 trading partners (40 separate countries and a Rest-of-Word (ROW) aggregate, consisting of 24 additional nations).[22] The eight manufacturing sectors are classified according to the United Nations' 2-digit International Standard Industrial Classification (ISIC) Revision 2.[23] The period of investigation is 1990-2002. For our analysis, we use only the years 1990, 1994, 1998, and 2002, akin to BB's use of data for every five years. We use a shorter four-year interval than BB's five-year interval due to the shorter time-series for our data. However, the use of every four years (or five years in BB) addresses the concern raised in Cheng and Wall (2005) that "Fixed-effects estimations are sometimes criticized when applied to data pooled over consecutive years on the grounds that dependent and independent variables cannot fully adjust in a single year's time" (p. 8).

[21]However, for robustness we will also examine the sensitivity of our findings to aggregate trade flows and GDPs.

[22]The 40 main countries are Argentina, Australia, Austria, Belgium-Luxembourg, Bolivia, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Denmark, Ecuador, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea (South), Mexico, Morocco, Netherlands, Norway, Poland, Portugal, Romania, Spain, Sweden, Switzerland, Tunisia, Turkey, United Kingdom, United States, and Uruguay. The 24 countries in the ROW aggregate are Cameroon, Cyprus, Egypt, Hong Kong, Indonesia, India, Iran, Jordan, Kenya, Kuwait, Sri Lanka, Macao, Malta, Myanmar, Malawi, Malaysia, Niger, Nepal, Philippines, Senegal, Singapore, Trinidad and Tobago, Tanzania, and South Africa.

[23]The nine 2-digit ISIC manufacturing categories are (short labels, used for convenience throughout the paper, are reported in parentheses): 31. Food, Beverages, and Tobacco Products (Food); 32. Textile, Apparel, and Leather Products (Textile); 33. Wood and Wood Products (Wood); 34. Paper and Paper Products (Paper); 35. Chemicals, Petroleum, Coal, Rubber, and Plastic Products (Chemicals); 36. Other Non-metallic Products (Minerals); 37. Basic Metal Industries (Metals); 38. Fabricated Metal Products, Machinery, Equipment (Machinery); 39. Other manufacturing. Inspection of the output data at the 3-digit and 4-digit ISIC level of aggregation reveals that many countries report Equipment production, and especially Scientific Equipment production, under the category Other Manufacturing. Therefore, to avoid inconsistencies, we combine the last two 2-digit categories into one, which we label Machinery.

Also, Wooldridge (2009) confirms the reduction in standard errors of coefficient estimates using changes over longer periods of time than using "year-to-year" changes (p. 459).

Bilateral international trade flows are defined as the value of exports from exporter i to importer j. We use the CEPII *Trade, Production and Bilateral Protection Database*[24] (TradeProd) as the main trade data source because it implements a consistent procedure for mapping the CIF (cost, insurance and freight) values reported by the importing countries in COMTRADE to the FOB (free on board) values reported by the exporters in COMTRADE.[25] This decreases the number of missing observations in the sample.[26] To further decrease the number of missing trade flows, we add export values from the United Nation Statistical Division (UNSD) Commodity Trade Statistics Database (COMTRADE).[27] Internal commodity-level (intranational) trade for each country is constructed as the difference between total manufactures output and aggregate manufactures exports to all trading partners, which come from the same data sources. The number of zero trade flows in the sample is very small and we will document this later. This suggests that the consequences of throwing information away by using the standard log-linear OLS estimator should not be severe. Nonetheless, the PQML estimator is still preferable because, in addition to accounting for the zero trade flows, it also controls for heteroskedasticity bias introduced due to Jensen's inequality.

Industrial output level data comes from two sources. The primary source is the United Nations' UNIDO Industrial Statistics database, which reports industry-level output data at the 3-digit and 4-digit level of ISIC Code (Revisions 2 and 3). We use the CEPII TradeProd

[24]For details regarding this database see Mayer, Paillacar, and Zignago (2008).

[25]The TradeProd database is based on the CEPII *Base pour l'Analyse du Commerce International* (BACI) data. For details regarding BACI see Gaulier and Zignago (2008).

[26]As noted in Anderson and Yotov (2010), in principle, gravity theory calls for valuation of exports at delivered prices. In practice, valuation of exports FOB avoids measurement error arising from poor quality transport cost data.

[27]We access COMTRADE through the World Integrated Trade Solution (WITS) software, http://wits.worldbank.org/witsweb/. The software reports trade data in three different concordances including Harmonized System (HS) Revisions 1989/92 and 1996, and the Standard International Trade Classification (SITC), which are automatically converted to ISIC Rev. 2. To obtain maximum number of observations, we combine the data from the different concordances.

database as a secondary source of product-level output data.[28] We interpolate some of the missing output values for the sample countries, which account for 15.6% of the observations.

Data on EIA dummies comes from the Database on Economic Integration Agreements on Jeffrey Bergstrand's website (www.nd.edu/ jbergstr). Baier and Bergstrand's EIA database categorizes bilateral EIA relationships from 1950-2005 for pairings of 195 countries using a multichotomous index. In this study, $EIA_{ij,t}=1$ denotes a free trade agreement between a pair of countries ij in year t or deeper integration (such as a customs union, common market, or economic union), or 0 otherwise, as in BB. Table 1 lists the agreements. In our ROW aggregate, there are no countries with EIAs with the main 40 countries.[29]

4 Empirical Results for Total Manufacturing Trade Flows

Table 2 presents our main results using aggregate international and intranational manufactures bilateral trade flows. Table 2 is partitioned into two panels, 2A and 2B. Panel 2A provides estimates omitting our key variable $BRDR_{ij,t}$. Panel 2B includes $BRDR_{ij,1994}$, $BRDR_{ij,1998}$, and $BRDR_{ij,2002}$; $BRDR_{ij,1990}$ is omitted due to the inclusion of a constant.[30]

Even though the recent empirical gravity equation literature has been focusing upon PQML estimation, we report first OLS estimates (using positive trade flows only) for comparability to BB and other studies. Column (1) of Panel 2A (OLS1(+)) reports the coefficient estimates using only the current $EIA_{ij,t}$ dummy in equation (3) but omitting the $BRDR$ dummies; this specification is similar to that in BB, Table 5, specification (2). In reality, most EIAs are phased in over time and often the consequent changes in terms-of-trade affect trade flows with a lag. Accordingly, to allow for these effects, we also include lagged values of EIA. Column (2) of Panel 2A reports the coefficient estimates using current and two lags ($EIA_{ij,t-4}$ and $EIA_{ij,t-8}$) of the EIA dummy in equation (3) but omitting the $BRDR$

[28] TradeProd uses the OECD STAN Industrial Database in addition to UNIDO's Industrial Statistics Database.

[29] In one sensitivity analysis using aggregate goods trade flows, Table 4, EIAs include one-way and two-way preferential agreements also.

[30] However, due to the inclusion of the other fixed effects, the constant cannot be interpreted as an estimate of $BRDR_{ij,1990}$.

dummies; this specification is similar to that in BB, Table 5, specification (4). For brevity, we focus on Column (2) results in Panel 2A. While the total partial effect in BB is 0.76, the total partial effect here is 0.35 (including the second lag, which is statistically insignificant). The difference is attributable to three factors: different sample time periods, different country-pair samples, and inclusion in our results of intranational trade. The BB estimate is based upon data spanning 1960-2000 for pairings of 96 countries whereas the estimates in Table 2 are based upon manufactures data spanning only 1990-2002 for pairings of 40 countries and a ROW aggregate, and we include also intranational trade.[31]

Columns (1) and (2) of Panel 2B report the partial effects of using only current and current with two lags of the EIA dummy, respectively, but now including the $BRDR$ dummies ($BRDR_{ij,1994}$, $BRDR_{ij,1998}$ and $BRDR_{ij,2002}$). The first notable finding is that the comparable estimates from Panel 2A are *biased upward*, once one accounts for (unobserved) changes in the cost of international relative to intranational trade, as captured by the $BRDR$ dummies, although the differences in the EIA estimates between the two panels are not statistically significant. The second notable finding is that the $BRDR$ dummies are positive, economically and statistically significant, and increasing over time. These results suggest that – although the exclusion of the $BRDR$ dummies did not have a dramatic effect on the EIA coefficient estimates – the effect of international borders on international trade, on average across country pairs, has been declining dramatically. The coefficient estimate of 0.675 for $BRDR_{ij,2002}$ in column (2) implies that – after accounting for EIAs – international trade relative to intranational trade has almost doubled over 12 years, suggesting that the effect of the international border on trade has fallen by *half* over 12 years (i.e., 0.49 = 1 - $[1/e^{0.675}]$).[32] In comparison to other studies with multiple years, the results here are com-

[31] These differences are consistent with systematically falling (partial) effects of EIAs over time. In a recent examination of EIA effects using OLS, Baier, Bergstrand, and Clance (2013) show that earlier EIAs' effects were larger – approximately 0.68 in 1965 – consistent with earlier self-selection of country-pairs that expected to gain the most from EIAs. The partial effects decline systematically over time. EIA effects of later agreements have been smaller, approximately 0.41 in 2005, consistent with the findings here using OLS.

[32] We also estimated the first-difference version of the OLS specification, eliminating the bilateral fixed effects. The coefficient estimates were very similar quantitatively (and identical qualitatively) to the OLS results presented in column (2). These results were omitted from presentation for brevity, but are available upon request.

parable. de Sousa, Mayer, and Zignago (2012), for instance, found for a large sample that the border effect fell from 493 times ($e^{6.2}$) to 181 times ($e^{5.2}$), or by 63 percent, over the period 1990-2002. Their estimates, however, are based upon a specification not accounting for unobserved bilateral heterogeneity and for the endogeneity of prices. However, without pair fixed effects, border effect levels in their study were constrained to be identical across all country pairs.

As noted earlier, OLS results are likely to be biased owing to potential heteroskedasticity bias due to Jensen's inequality, cf., Santos Silva and Tenreyro (2006). The only distinction between columns (2) and (3) is that column (3) uses PQML. In Panel 2A there is a notable consequence of this difference. The effects of EIAs are much larger using PQML relative to OLS. Since the sample is the same as in column (2), with only positive trade flows, the larger EIA coefficient estimates are due solely to using PQML; this is consistent with Santos Silva and Tenreyro (2006), or SST.[33]

Panel 2B shows our novel inclusion of the $BRDR$ dummies influences the effects of EIAs on trade flows. Using PQML, the sum of the EIA coefficient estimates in column (3) falls from 0.795, implying an increase in trade after eight years of 121 percent ($=[e^{0.795}-1] \times 100$ percent), to 0.523, implying an increase in trade of 69 percent. Two points are worth noting. First, the estimated partial effect of an EIA after eight years of 69 percent is below the BB estimate, but still economically and statistically significant. Estimation in BB using OLS tends to bias downward EIA estimates relative to using PQML. However, ignoring unobserved time-varying costs of international trade relative to intranational trade (i.e., ignoring the $BRDR$ dummies) tends to bias upward EIA estimates. Recall, however, that the BB sample differed from our sample in terms of period covered, manufactures vs. aggregate trade flows, the countries included in the sample, and results here using intranational trade.

[33]Using an atheoretical gravity equation in a cross-section, SST found that their EIA coefficient estimate was lower using PQML than OLS. However, using a properly specified cross section with exporter and importer fixed effects to account for GDPs and prices, SST found that their EIA coefficient estimate was *larger* using PQML relative to OLS, indicating the importance of allowing for unobserved heterogeneity in the specification. However, SST did not account for the endogeneity of EIAs nor for intranational trade and distances.

Second, the coefficient estimates for the $BRDR$ dummies are considerably smaller using PQML than OLS. The coefficient estimate for $BRDR_{ij,2002}$ of 0.323 indicates that international trade has increased relative to intranational trade on average by about 38 percent (=[$e^{0.323}$-1] $\times$ 100 percent) after 12 years, or an economically plausible 2.7 percent per year. Since no other study has used PQML to evaluate the declining border effect, we have no other estimates to compare this against.[34]

The remainder of our findings in this section relate to sensitivity analyses for total manufactures trade flows. Our first sensitivity analysis is to include zeros in the analysis. As discussed earlier, estimation using PQML addresses two shortfalls of previous work using OLS on a log-linear version of equation (3). One is heteroskedasticity bias and the other is country-selection bias attributable to ignoring zeros. In both panels of Table 2, Column (4) provides estimates including zeros, and they are comparable to those in column (3) which excludes zeros. These estimates indicate that ignoring zeros, and consequently country-selection bias, has little effect on our results. However, note from comparing columns (3) and (4) that there are only 85 zeros (= 6,724-6,639). Hence, our results are not influenced by country-selection bias in this particular case because of the small number of zeros.

Our second sensitivity analysis is to include a (four-year) "lead" EIA effect. It is quite possible that exogeneity runs from trade flows *to* EIAs. Wooldridge (2010, p. 325) suggests that it is easy to test for the "strict exogeneity" of EIAs in our context.[35] To do this, we add a future level of EIA to the model. In the panel context here, if EIA is exogenous to trade flows, $EIA_{ij,t+4}$ should be uncorrelated with the current trade flow. If EIA is endogenous to trade, we would expect a positive coefficient estimate. Column (5) reports the results. In Panel 2A, we find that the future level of EIA has no economically or statistically significant

[34] de Sousa, Mayer, and Zignago (2012) only report declining border effects using OLS. Bosquet and Boulhol (2013) do not estimate changing border effects, but only changing distance effects, using PQML; however, as discussed later, they do not find falling distance elasticities, likely due to omitting intranational trade. Note that the initial border effect levels for each country pair in our study are allowed to differ across pairs, and are subsumed in the pair fixed effects.

[35] An empirical finding that trade leads an EIA need not imply that trade "causes" an EIA. Trade may increase in anticipation of an EIA, as infrastructure and delivery systems involving sunk costs are redirected. Alternatively, trade may decrease – or be delayed – in anticipation of the benefits of an EIA.

effect on current trade. This is actually a stronger result than found in BB using a longer time series and aggregate goods trade flows; BB found a negative coefficient estimate of -0.04. Column (5) in Panel 2B, however, reports a result more similar to that in BB, a small negative, but statistically significant, effect of an impending EIA on current trade. However, a small negative effect is not a problem; our concern would be if the lead effect was significant *and positive*, as this could be interpreted as trade causing EIAs. A small negative effect is easily interpreted, as in BB, as anticipation of an impending EIA delaying trade today. Moreover, we find that the exclusion of the lead EIA effect biased downward slightly the concurrent effect of an EIA on trade. We note that the effects of the border dummies are unchanged materially relative to the previous results.

Our third sensitivity analysis is a comparison of equation (3) versus equation (4). It is possible that unobserved country-pair heterogeneity is not time-invariant; in fact, our coefficient estimates for the $BRDR$ dummies suggest that the costs of international relative to intranational trade are decreasing – independent of EIA formations. Equation (4) includes country-pair fixed effects interacted with a linear time trend. Note that inclusion of ij fixed effects interacted with year dummies is infeasible; it would perfectly predict trade flows. Column (6) reports the results using equation (4), which are directly comparable to those in column (4). It is evident that there is no material difference in the results using equation specifications (3) or (4).

One potential bias we have not accounted for is firm-heterogeneity bias. As discussed in Helpman, Melitz, and Rubinstein (2008), or HMR, and Egger, Larch, Staub, and Winkelmann (2011), the existence of firm heterogeneity may bias coefficient estimates in gravity equations using aggregate data. One of the advantages of HMR's two-stage approach is that it accounts for zeros, but also for firm heterogeneity, when using aggregate trade flows. HMR concluded that firm-heterogeneity bias mattered even more than country-selection bias in their cross-section estimates.[36] However, accounting also for endogeneity (self-selection) bias

[36] HMR also estimated their model pooling several cross-sections over time. However, their estimation never included bilateral fixed effects in their second stage, which is critical to our discussion below.

of EIAs, Egger, Larch, Staub, and Winkelmann (2011) found that firm-heterogeneity bias hardly mattered at all. We argue here that – for our panel specification shown in equation (3) – the results are not likely influenced materially by firm-heterogeneity bias, due to the inclusion of the bilateral fixed effects. This is an issue explored only recently in Baier, Bergstrand, and Feng (2013).

To understand why, we first review briefly the HMR approach, which was used in a cross section (1986 trade flows). The two-stage methodology entails estimating first a probit equation to determine the probability of a positive observation between a country pair. The probit estimates are then used to construct inverse Mills' ratios (denoted $IMR_{ij,t}$) to capture selection bias and variables $Z_{ij,t}$, $Z_{ij,t}^2$, and $Z_{ij,t}^3$ to control for heterogeneous productivities of firms. $IMR_{ij,t}$, $Z_{ij,t}$, $Z_{ij,t}^2$, and $Z_{ij,t}^3$ are then used as additional regressors in the second-stage gravity-equation specification to control for country-selection and firm-heterogeneity biases. Both stages of estimation used exporter and importer fixed effects to account for multilateral variables, but did not use bilateral country-pair fixed effects.

In our panel specification of equation (3), we have time-invariant country-pair *bilateral* fixed effects. If most of the variation in the predicted probit probabilities of trading (i.e., selection of country-pairs into positive trade) is cross-sectional in nature, then time-invariant country-pair bilateral fixed effects in the second stage will account for most of the variation in $IMR_{ij,t}$, $Z_{ij,t}$, $Z_{ij,t}^2$, and $Z_{ij,t}^3$. The only possible bias in gravity equation coefficient estimates using our panel attributable to selection and firm-heterogeneity would be time variation in $IMR_{ij,t}$, $Z_{ij,t}$, $Z_{ij,t}^2$, and $Z_{ij,t}^3$. It becomes an empirical issue then to determine if such bias is material.

Baier, Bergstrand, and Feng (2013), or BBF, recently addressed the HMR two-stage estimation procedure in a panel with bilateral fixed effects (and alternatively first-differencing) in the second stage. Akin to HMR, BBF first estimated eight individual cross-section probits for the years 1965, 1970,..., 2000 to generate predicted probabilities of positive aggregate goods trade flows for a large number of country pairs. They then used these predicted probits to construct for each year $IMR_{ij,t}$, $Z_{ij,t}$, $Z_{ij,t}^2$, and $Z_{ij,t}^3$. In the second stage, they estimated a specification similar to equation (3), but excluding $BRDR_{ij,t}$ and using OLS (or the first-difference analogue). Their results from the second stage regressions were reported in BBF's Appendix Table A4, which can be readily compared to the results from omitting $IMR_{ij,t}$, $Z_{ij,t}$, $Z^2_{ij,t}$, and $Z^3_{ij,t}$ which were presented in Table 1 of BBF. A comparison of the results from the two tables reveals clearly that there is very little quantitative and *no* qualitative differences between the respective coefficient estimates. The reason is the presence in the second stage of the first-differencing (or bilateral fixed effects). Put simply, most of the variation in the predicted probabilities of positive trade flows is cross-sectional, not time-varying; first-differencing (or bilateral fixed effects) accounts largely for the influences of country-selection and firm-heterogeneity. Based upon those results, we argue our results are likely robust to firm-heterogeneity bias. Moreover, one feature of our data is that there are very few zero trade flows that are not perfectly explained by our fixed effects (country-pair, exporter-time, and importer-time). This implies that there are few country-pairs that start or stop trading in our data set.[37]

5 Empirical Results for Disaggregate Manufactures Trade

One dimension which BB ignored entirely is the sensitivity of the findings to disaggregation of trade flows. The empirical literature on partial effects of EIAs using disaggregate data is actually *very* small. Anderson and Yotov (2011) is one of the few studies using the BB approach to analyze disaggregate trade flow effects, and our data set allows us to explore disaggregation since it is based upon the same data. Table 3 provides the results of estimating the partial EIA effects using equation (3) for the eight 2-digit ISIC categories of manufactures. Table 3 is divided into four panels (3A.1, 3A.2, 3B.1, 3B.2). Panels 3A.1 and 3A.2 provide the main PQML specification results using each of eight manufactures industries and 3B.1 and 3B.2 provide the OLS results (for robustness). We address each panel in turn.

[37]For this issue, it is critical to note that identification of the HMR terms $IMR_{ij,t}$, $Z_{ij,t}$, $Z^2_{ij,t}$, and $Z^3_{ij,t}$ relies on the variation of the export-status (positive or zero). The HMR approach was explored using our data set; however, convergence could not be achieved precisely because of the very small number of zeros.

Panels 3A.1 and 3A.2 provide the results of EIA effects in the analysis of disaggregate trade flows using PQML. The results are largely consistent with those in the previous section for total manufactures trade flows. First, compare the results in Panel 3A.1 of Table 3 with those in column (4) of Panel 2A in Table 2. We find positive, economically significant, and statistically significant partial effects of EIAs on trade flows in *all eight* sectors, with the total ATE of the eight sectors ranging from 0.286 (Wood Products) to 0.960 (Machinery).

Panel 3A.2 provides the results including lagged EIA effects as well as the border dummies. As in previous estimates, the EIA effects are diminished, but remain economically and statistically significant. For instance, the partial effect for Machinery Products falls from 0.960 to 0.583. Yet, this still implies that an EIA increases trade by 79 percent (absent any general equilibrium effects). Estimates of the sums of current and lagged effects range from 0.343 to 0.664.[38] Moreover, we find that the coefficient estimates for the border dummies are positive (except for Wood) and, in many cases, statistically different from zero. On net, we find the coefficient estimates for the $BRDR$ dummies are economically feasible using PQML, as we found in Table 2. We do not present the results for PQML with positive flows only, as these results were very close to those including zeros, as we established in Table 2. And as mentioned above, inclusion of a lead EIA effect only had a statistically significant positive lead effect in Textiles.

Panels 3B.1 and 3B.2 present OLS results akin to those in column (2) of Table 2's Panels 2A and 2B. In Panel 3B.1, with the exception of Chemicals, all sectors reveal an economically and statistically significant positive partial effect of EIAs on trade flows. Panel 3B.2 enhances the specifications to include our key variables, the border dummies. We note two important results. First, as for total manufactures trade flows, the inclusion of the border dummies has the effect of lowering the partial effect of an EIA. However, we still find that an EIA has an economically and statistically significant positive effect on trade

[38] We also added a lead EIA effect for every sector, similar to Table 2. While the results are omitted for brevity, we note the following. Only Textile Products showed a statistically significant positive lead effect suggesting reverse causality. Also, three sectors (Paper, Metals, and Machinery) had statistically significant negative lead EIA coefficient estimates, similar in magnitude though to that for total manufactures, indicating trade flows in these sectors were subject to "anticipation" of impending EIAs, falling in advance of the EIA.

in seven of the eight sectors. However, one surprising result is that the effect of EIAs is not very large for Machinery, giving greater credence to the PQML results. Second, we find that the border dummies have economically and statistically significant positive effects in all estimates, which is robust support for the specification in equation (3). However, across columns (1)-(8), the positive and statistically significant coefficient estimates for the border dummies are much larger in many cases than the corresponding PQML results, and in some cases seemingly economically implausible, also lending more credence to the PQML findings.

In sum, the main PQML results shown in column (4) of Panel 2B of Table 2 for total manufactures, and in Panel 3A.2 of Table 3 for disaggregate trade flows, are very robust to an extensive sensitivity analysis. These results suggest that – after accounting for likely declining bilateral variable and fixed trade costs using a novel set of "border dummies" – EIAs still have economically and statistically significant partial effects on trade flows, but that ignoring the border dummies biased upward EIA estimates. Moreover, the novel "border dummies" reveal the average cost of international relative to intranational trade is declining, consistent with falling $\tau_{ij,t}$ and $f_{ij,t}$ in the Melitz model in equation (1).

6 Empirical Results for Aggregate Trade Flows

One of the limitations of the manufactures data set employed is the short time series; we are able to explain the effect of declining bilateral trade costs over a period of only 12 years.[39] However, data is available for a longer time series for bilateral aggregate goods trade flows. The drawback of using this longer time series is that our available measure of gross output, from which to construct intranational trade flows, is GDP – a "value-added" (not gross output) measure. Nevertheless, we thought it worthwhile in this section to evaluate the sensitivity of the earlier results for total manufactures trade flows using a shorter time series to use of aggregate goods trade flows for a longer period.

Table 4 provides the results. In the first column of results using PQML, we show that

[39] The effective constraint is available production data.

as before the current EIA dummy has an economically and statistically significant positive effect. The introduction in the next specification of the "border dummies" causes as expected the coefficient estimates of $EIA_{ij,t}$ and its lags to decline. Note, however, the the total EIA partial effect in column (2) of 0.554 is *very close* to the comparable estimate in column (4) of Table 2, Panel 2B of 0.522. Moreover, it is worth noting the pattern of coefficient estimates for the five border dummies. Except for the coefficient estimate for $BRDR_{ij,1993}$, all of the coefficient estimates are positive. The negative coefficient estimate for $BRDR_{ij,1993}$ is explained by the fact that these represent changes in the effect of international relative to intranational trade for 1993 relative to 1989. 1989 was the last year before a global economic slowdown (and the U.S.-Iraq War) which troughed around 1992-1993; thus, the negative effect may reflect this. Similarly, the coefficient estimate for $BRDR_{ij,2009}$ is below that for $BRDR_{ij,2005}$. This also is not surprising, since the financial and liquidity crisis of fall 2008 and spring 2009 raised the fixed cost of trade of numerous firms (reflected in dramatically higher LIBOR rates), reducing international trade relative to intranational trade. Thus, the coefficient estimates for the "border dummies" make sense and are consistent with earlier results.

The third and fourth columns report the results using instead OLS. As found earlier, EIAs' coefficient estimates are smaller using OLS relative to PQML. However, in this sample we find that the estimates of the declining effect of national borders using OLS are only slightly higher than those using PQML. Thus, using this sample, differences in declining border effect estimates between the two estimation procedures is not as pronounced.

7 The "Distance-Elasticity" Puzzle

One of the potential key contributions of this paper is the introduction of a variable $BRDR_{ij,t}$ to account for likely declining trends in bilateral fixed and variable trade costs that are likely increasing international relative to intranational trade. As pointed out earlier, however, the nature of the year dummies interacted with the dummy for whether trade was international relative to intranational implies that $BRDR_{ij,t}$ can only hold constant the "average effect" of declining relative international trade costs.

However, there is a way to introduce "pair-specificity" to capture these declining bilateral trade costs. We introduce an alternative measure called $DIST_{ij,t}$. We define $DIST_{ij,t} = BRDR_{ij,t}$ x $\ln DISTANCE_{ij}$, where $DISTANCE_{ij}$ was defined earlier as the bilateral distance between the economic centers of the source and destination countries. Recall, our specification in equation (3) still includes country-pair fixed effects. In this case, $DIST_{ij,t}$ still captures the time-varying effects of changing costs in international relative to intranational trade (relative to a pair's mean, captured by the pair fixed effect) as did $BRDR_{ij,t}$. However, whereas $BRDR_{ij,t}$ captured the *average effect* of falling international relative to intranational trade costs (across pairs), $DIST_{ij,t}$ allows this effect to be sensitive to the country-pairs' bilateral distance.[40]

Table 5 presents a set of results (for its first eight columns) in Panel 5A using $DIST_{ij,t}$ that can be compared to those in the eight columns of Panel 3A.2 in Table 3. The basic finding is that the main results in Panel 3A.2 hold up well. Comparison of comparable industries' results between the two panels shows that allowing for declines in the relative costs of international relative to intranational trade to be pair-specific (i.e, sensitive to the pair-specific bilateral distance) does not alter the main findings. However, it is important to note that, because of the interactions with the bilateral distance variable $\ln DISTANCE_{ij}$, the coefficient estimates for $DIST_{ij,t}$ are a different order of magnitude than those using $BRDR_{ij,t}$.

Yet, the results just discussed raise the possibility of addressing another important issue. One of the well-known puzzles in the empirical international trade literature is the "distance-elasticity puzzle." This puzzle is that – in spite of well documented advances in IT that

[40] $DISTANCE_{ij}$ is calculated identically to that in Yotov (2012). Following Mayer and Zignago (2006), bilateral distance – both between countries and (internal distance) within countries – is calculated as $DISTANCE_{ij} = \sum_{k \in i} Pop_k/Pop_i \sum_{l \in j} Pop_l/Pop_j D_{kl}$, where Pop_k is the population of agglomeration k in exporter i, Pop_l is the population of agglomeration l in importer j, and D_{kl} is the bilateral distance in kilometers between agglomeration k and agglomeration l (using Great Circle Distance formula). All data on latitudes, longitudes, and population are from the World Gazetteer web page. A nice feature of this variable is that the same procedure is used to construct (consistently) international as well as intranational distances.

have likely reduced bilateral fixed and variable trade costs – a time series of cross-sectional estimates of a properly-specified "gravity equation" yield *rising* distance elasticities. That is, international trade in such cross sections declines more in response to distance in recent years relative to earlier years, cf., Disdier and Head (2008).

While several researchers have made attempts to solve the puzzle, Yotov (2012) addressed the issue by including observations for *intranational* trade along with including a variable measuring intranational distances. Such intranational trade flows and distances have actually been a common feature of several border-effect studies, but had not yet permeated the distance-elasticity literature. Yotov (2012) "solved" the distance-elasticity puzzle by noting the importance of measuring international distances *relative to* intranational distances, as theoretical foundations for gravity equations actually suggest. Importantly, Yotov (2012) solved the distance-elasticity puzzle using both OLS and PQML.

However, a shortcoming of Yotov (2012) is that – by using a time-series of cross-sections – the author does not control for unobserved bilateral heterogeneity and consequently the results suffer from omitted variables bias. Moreover, the results did not account for endogenous EIAs' effects. In order to estimate the distance elasticity, country-pair fixed effects cannot be used, as they would subsume the cross-country variation in bilateral distance. Consequently, researchers typically include bilateral distance and many other pair-specific variables to explain trade cross-sectionally. This leads to potentially biased results.[41]

To obtain unbiased estimates, our approach uses a panel with pair fixed effects. The pair fixed effects capture the cross-sectional negative impact of bilateral distance on trade flows. We then introduce a set of year dummies interacted with bilateral distance. This variable, termed $DIST2_{ij,t}$, can potentially capture the changing effects of bilateral distance

[41] Yotov (2012) addressed the distance puzzle using several cross-sections (including intranational trade flows as well as international trade flows on the LHS), but included a separate variable to control for intranational distances in an otherwise typical gravity equation. However, Yotov (2012) faced a shortcoming. Yotov (2012) uses only a limited number of time-invariant pair-specific observable determinants of trade flows to capture the full array of time-invariant factors influencing bilateral trade cross-sectionally, and no pair-specific (ij) fixed effects, creating potential omitted variables bias. By contrast, the approach in our paper includes ij fixed effects to control for all time-invariant bilateral observables and unobservables influencing trade flows cross-sectionally.

on trade flows *relative to the initial year*. Another way to look at this variable is that it is a time-varying measure of the changing costs of international trade relative to intranational trade, but using a continuous measure rather than the earlier employed border dummies. Formally, $DIST2_{ij,t} = \ln DISTANCE_{ij}$ x D_t. Table 5, Panel 5B, presents the alternative results using $DIST2_{ij,t}$. Consistent with the results in Panel 5A, the variables $DIST2_{ij,t}$ have economically and statistically significant positive effects. Moreover, the size of the coefficient estimates tend to increase from 1994 to 2002. For total manufactures in column (9), the sizes of the positive coefficients increase monotonically.

What do these coefficient estimates suggest? The country-pair fixed effects (whose coefficient estimates are not shown due to the very large number) pickup that bilateral distance has a negative effect on trade. However, the *positive* and typically increasing over time coefficient estimates for $DIST2_{ij,1994}$, $DIST2_{ij,1998}$, and $DIST2_{ij,2002}$ indicate that the negative effect of bilateral distance is *declining* over time. In the context of the discussion above and gravity equation (1), these results are consistent with the costs of international trade falling over time relative to intranational trade, likely attributable to decreasing – but unobservable – bilateral fixed and variable trade costs, that are increasing international relative to intranational trade. In fact, the coefficient estimate in panel 5B for $DIST2_{ij,2002}$ of 0.147 for total manufactures trade suggests that the average effect of distance on international relative to intranational trade has diminished by 14 percent over 12 years (1 - $[1/e^{0.147}] \times$ 100 percent), an economically plausible 1.2 percent per year. A comparison of our results with those in Yotov (2012, Table 2) also reveals more economically plausible estimates using our approach. For instance, in Yotov (2012) the distance elasticity for textiles fell 57 percent over 10 years; such an estimate seems economically implausible. By contrast, our results in Panel 5B imply that the distance elasticity for textiles fell only 24 percent over a similar 12-year period, and this was the largest estimated distance-elasticity decrease. In fact, the smaller declines in the distance elasticities in our study relative to Yotov (2012) suggest that the estimates in the latter study were biased upward by ignoring the effects of EIAs.

This result is novel because it is generated allowing the country-pair fixed effects to subsume the level effect of distance on trade flows, and allows this effect to differ across country pairs in the initial year.[42] Only two previous studies have included bilateral country-pair fixed effects to address the distance-elasticity puzzle. Carrere, de Melo, and Wilson (2009) account for unobserved bilateral heterogeneity in their OLS estimates as well as linear approximations of the multilateral price terms. They find rising distance elasticities; however, they do not account for EIAs or intranational trade and distances. Bosquet and Boulhol (2013) could not find declining distance elasticities using PQML including bilateral fixed effects, but that is likely attributable to their exclusion of intranational trade and distances. However, unlike Carrere, de Melo, and Wilson (2009) and Bosquet and Boulhol (2013), we include as in Yotov (2012) intranational trade and distances; this feature is important to find declining distance effects on international trade, because we allow for measurement of the effects of distance on international relative to intranational trade. In a robustness check of the importance of accounting for intranational trade as well, we re-estimated the specifications in Panel 5B excluding intranational trade; these are presented in Panel 5C. As seen there, we no longer have general evidence of a declining effect of distance on international trade. Only in the case of Chemicals do we find evidence of a declining distance elasticity. In most industries, we find negative coefficient estimates, with some statistically significant, implying rising distance elasticities. Thus, PQML alone will not solve the distance-elasticity puzzle. Measuring the effects of distance, or borders, on international relative to intranational trade requires inclusion of intranational trade. Importantly, note also that the coefficient estimates of EIAs are biased in Panel 5C relative to Panel 5B by excluding intranational trade, and our approach accounts for the endogeneity of EIAs. Finally, we have also re-estimated the specifications shown in Panel 5B *but excluded EIAs*; we found that the $DIST2$ coefficient

[42] Coe, Subramanian, and Tamirisa (2007) found evidence of declining distance elasticities by employing non-linear estimation (including PQML), although the declines ended between 1990-2000, but could not find declining distance elasticities using OLS. Berthelon and Freund (2008) found rising distance elasticities since 1985 using OLS. Larch, Norbäck, Sirries, and Urban (2012) found a declining distance effect using non-linear estimators, but not using OLS. However, Coe, Subramanian, and Tamirisa (2007), Berthelon and Freund (2008), and Larch, Norbäck, Sirries, and Urban (2012) did not account for unobserved bilateral heterogeneity or for the endogeneity of EIAs, and did not include intranational trade flows and distances.

estimates were biased upward by as much as 40 percent.[43]

8 Conclusions

We have attempted to provide using a common gravity-equation specification consistent, precise, and economically plausible estimates of the (partial) effects of three important concepts in international trade that typically have been addressed in three somewhat separate literatures. First, we have improved upon the specification in BB for estimating the effects of EIAs on international trade flows by controlling now for *time-varying* unobservable bilateral trade costs (such as IT costs) that may have increased international relative to intranational trade; our results suggest that previous estimates of EIAs' effects were biased upward. Using our econometrically preferred estimator (PQML), the partial effect of an EIA is nearly halved.

Second, our novel approach allows us to estimate precisely the declining effect of "national borders" on international relative to intranational trade allowing for unobserved bilateral country-pair heterogeneity and endogenous EIAs. While previous authors have found evidence of declining border effects, such estimates may have been biased upward by ignoring endogenous EIAs. One of the shortcomings of these previous studies is omitted variables bias in initial border effect *levels* and not accounting for endogenous EIAs. Our results suggest that previous estimates of the declining effect of national borders were biased upward, and we find the effects of national borders on international trade have declined an economically plausible 2.7 percent per year from 1990 to 2002.

Third, in an extensive sensitivity analysis, we introduce another method for accounting for unobserved *time-varying* declines in the costs of bilateral international relative to intranational trade. Accounting for endogenous EIAs and unobserved country-pair heterogeneity, we provide economically plausible estimates of the declining effect of distance on international trade, providing empirical support for the elusive declining "distance elasticity"

[43] Results available on request.

of international trade. While our approach recognizes as in Yotov (2012) the importance of including intranational trade and using PQML in estimation, our novel contribution here is to account for unobserved country-pair heterogeneity and endogenous EIAs. We find that previous estimates of the declining effect of distance on international relative to intranational trade have been biased upward by not accounting for endogenous EIAs and unobserved bilateral heterogeneity. Our results suggest that the effect of distance on international trade has declined an economically plausible 1.2 percent annually.

Just as BB contributed to the literature by emphasizing the importance of accounting for exporter-year, importer-year, and country-pair fixed effects in estimating the (partial) effects of EIAs, our hope is that – going forward – subsequent analyses account for all of the following using panel techniques:

(1) Exporter-year and importer-year fixed effects to account for endogenous prices and unobserved time-varying exporter and importer multilateral heterogeneity;

(2) Country-pair fixed effects or country-pair fixed effects interacted with a time trend to account for unobserved time-invariant or time-varying, respectively, bilateral effects (that subsume pair-specific border effect levels);

(3) Intranational as well as international trade flows and bilateral distances, so that the border dummies can account for declining international *relative to* intranational bilateral trade costs other than EIAs; and

(4) PQML estimation to account for heteroskedasticity bias, owing to Jensen's inequality.

References

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

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

ANDERSON, J., AND Y. YOTOV (2010): "The Changing Incidence of Geography," *American Economic Review*, 100(5), 2157–2186.

——— (2011): "Terms of Trade and Global Efficiency Effects of Free Trade Agreements," *NBER Working Paper No. 17003*.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRIGUEZ-CLARE (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

BAIER, S., AND J. BERGSTRAND (2007): "Do Free Trade Agreements Actually Increase Members' International Trade?," *Journal of International Economics*, 71(1), 72–95.

——— (2009): "*Bonus Vetus* OLS: A Simple Method for Approximating International Trade-Cost Effects using the Gravity Equation," *Journal of International Economics*, 77(1), 77–85.

BAIER, S., J. BERGSTRAND, AND M. CLANCE (2013): "Preliminary Examination of Heterogeneous Effects of Economic Integration Agreements," *Working Paper*.

BAIER, S., J. BERGSTRAND, AND M. FENG (2013): "Economic Integration Agreements and the Margins of International Trade," *Working Paper*.

BERDEN, K., J. FRANCOIS, S. TAMMINEN, M. THELLE, AND P. WYMENGA (2010): *Non-tariff measures in EU-US Trade and Investment*. ECORYS, The Netherlands.

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

BERGSTRAND, J. H., P. EGGER, AND M. LARCH (2013): "*Gravity Redux*: Estimation of Gravity-Equation Coefficients, Elasticities of Substitution and General Equilibrium Comparative Statics under Asymmetric Bilateral Trade Costs," *Journal of International Economics*, 89(1), 110–121.

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

Bosquet, C., and H. Boulhol (2013): "What Is Really Puzzling about the "Distance Puzzle"," *working paper*.

Carrere, C., J. de Melo, and J. Wilson (2009): "The Distance Effect and the Regionalization of the Trade of Low-Income Countries," *CEPR Working Paper*.

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

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

Cipollina, M., and L. Salvatici (2010): "Reciprocal Trade Agreements in Gravity Models: A Meta-Analysis," *Review of Economic Literature*, 18(1), 63–80.

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

de Sousa, J., T. Mayer, and S. Zignago (2012): "Market Access in Global and Regional Trade," *Regional Science and Urban Economics*, 42(6), 1037–1052.

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

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

Egger, P., M. Larch, K. Staub, and R. Winkelmann (2011): "The Trade Effects of Endogenous Preferential Trade Agreements," *American Economic Journal: Economic Policy*, 3(3), 113–143.

Feenstra, R. (2004): *Advanced International Trade: Theory and Evidence*. Princeton University Press, Princeton, New Jersey.

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

Fontagne, L., T. Mayer, and S. Zignago (2005): "Trade in the Triad: How Easy is the Access to Large Markets?," *Canadian Journal of Economics*, 38(4), 1401–1430.

Freund, C., and D. Weinhold (2004): "The Effect of the Internet on International Trade," *Journal of International Economics*, 62(1), 171–189.

Gaulier, G., and S. Zignago (2008): "BACI: A World Database of International Trade Analysis at the Product Level," *CEPII Working Paper*.

Head, K., and T. Mayer (2000): "Non-Europe: The Magnitude and Causes of Market Fragmentation in the EU," *Weltwirtschaftliches Archiv*, 136(2), 284–314.

——— (2013): "Gravity Equations: Workhorse, Toolkit, and Cookbook," *CEPR Working Paper No. 9322,* draft of a chapter to appear in the Handbook of International Economics Vol. 4, eds. Gopinath, Helpman, and Rogoff.

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

Krugman, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70(5), 950–959.

Larch, M., P. Norbäck, S. Sirries, and D. Urban (2012): "Heterogeneous Firms, Globalization, and the Distance Puzzle," *working paper*.

Mayer, T., R. Paillacar, and S. Zignago (2008): "TradeProd. The CEPII Trade, Production and Bilateral Protection Database: Explanatory Notes," *CEPII*.

Mayer, T., and S. Zignago (2006): "Notes on CEPII's Distance Measures," *CEPII*.

McCallum, J. (1995): "National Borders Matter," *American Economic Review*, 85(3), 615–623.

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

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

——— (2011): "Further Simulation Evidence on the Poisson-PML Estimator," *Economics Letters*, 112(2), 220–222.

Tang, L. (2006): "Communications Costs and Trade of Differentiated Goods," *Review of International Economics*, 14(1), 54–68.

Wei, S. (1996): "Intra-National versus International Trade: How Stubborn are Nations in Global Interactions?," *National Bureau of Economic Research Working Paper*, (5531).

Wooldridge, J. (2009): *Introductory Econometrics: A Modern Approach*. Cengage Learning Services, Florence, KY, 4th edn.

——— (2010): *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts, 2nd edn.

Yotov, Y. V. (2012): "A Simple Solution to the Distance Puzzle in International Trade," *Economics Letters*, 117(3), 794–798.

Table 1: Economic Integration Agreements

European Union, or EU (1958): Belgium–Luxembourg, France, Italy, Germany, Netherlands, Denmark (1973), Ireland (1973), United Kingdom (1973), Greece (1981), Portugal (1986), Spain (1986), Iceland (1994) Austria (1995), Finland (1995), Sweden (1995)
European Free Trade Association, or EFTA (1960): Austria (until 1995), Denmark (until 1973), Iceland (1970), Finland (1986–1995), Norway, Portugal (until 1986), Sweden (until 1995), Switzerland, United Kingdom (until 1973)
Latin American Free Trade Agreement/Latin American Integration Agreement, or LAFTA/LAIA (1993–): Argentina, Bolivia, Brazil, Chile, Ecuador, Mexico, Uruguay
EU–EFTA Agreement/European Economic Area (1973/1994)
US–Israel (1985)
US–Canada (1989)
EFTA–Israel (1993)
Central Europe Free Trade Agreement, or CEFTA (1993): Hungary, Poland, Romania (1997), Bulgaria (1998)
EFTA–Turkey (1992)
EFTA–Bulgaria (1993)
EFTA–Hungary (1993)
EFTA–Poland (1993)
EFTA–Romania (1993)
Andean Community (1993): Bolivia, Columbia, Ecuador
EU–Hungary (1994)
EU–Poland (1994)
North American Free Trade Agreement, or NAFTA (1994): Canada, Mexico, United States
Bolivia–Mexico (1995)
Costa Rica–Mexico (1995)
EU–Bulgaria (1995)
EU–Romania (1995)
Columbia–Mexico (1995). As part of the Group of Three. The third country, Venezuela, is not in the sample.
Mercosur (1991): Argentina, Brazil, Uruguay (formed in 1991 FTA in 1995)
Mercosur–Chile (1996)
Mercosur–Bolivia (1996)
EU–Turkey (1996)
Canada–Chile (1997)
Canada–Israel (1997)
Hungary–Turkey (1998)
Hungary–Israel (1998)
Israel–Turkey (1998)
Romania–Turkey (1998)
Poland–Israel (1998)
EU–Tunisia (1998)
Mexico–Chile (1999)
EU–Israel Agreement (2000)
EU–Mexico (2000)
EU–Morocco (2000)
EFTA–Morocco (2000)
Poland–Turkey (2000)
Mexico–Israel (2000)
Chile–Costa Rica (2002)

Notes: This table lists, in chronological order, all economic integration agreements (EIAs) used in estimation. Only agreements involving the countries in our sample are included. EIAs that entered into force before 1990 are used, when appropriate, to construct the lagged variables of the EIA dummy variable. For all estimations using total or disaggregate manufactures trade EIAs include free trade agreements and deeper integration agreements based upon the Baier-Bergstrand data set. For the single robustness analysis using aggregate trade flows (Table 4 below), one-way and two-way preferential agreements were included also.

Table 2: Panel Gravity with Exporter-Year, Importer-Year, and Country-Pair FEs

	(1) OLS1(+)	(2) OLS2(+)	(3) PQML(+)	(4) PQML	(5) PQML Lead	(6) PQML(Eq.4)
2A. Standard EIA Effects (No Globalization)						
$EIA_{ij,t}$	0.199	0.166	0.244	0.243	0.245	0.253
	(0.066)**	(0.059)**	(0.035)**	(0.035)**	(0.042)**	(0.036)**
$EIA_{ij,t-4}$		0.089	0.272	0.272	0.272	0.277
		(0.053)+	(0.068)**	(0.068)**	(0.067)**	(0.068)**
$EIA_{ij,t-8}$		0.097	0.279	0.279	0.279	0.284
		(0.067)	(0.054)**	(0.054)**	(0.056)**	(0.055)**
$EIA_{ij,t+4}$					-0.004	
					(0.048)	
EIA_TOTAL		0.352	0.795	0.794	0.796	0.815
		(0.125)**	(0.069)**	(0.070)**	(0.075)**	(0.072)**
2B. EIA Effects and Globalization						
$EIA_{ij,t}$	0.173	0.144	0.097	0.097	0.136	0.098
	(0.060)**	(0.055)**	(0.041)*	(0.041)*	(0.043)**	(0.042)*
$EIA_{ij,t-4}$		0.080	0.195	0.195	0.192	0.194
		(0.052)	(0.052)**	(0.052)**	(0.051)**	(0.052)**
$EIA_{ij,t-8}$		0.089	0.231	0.231	0.224	0.229
		(0.065)	(0.050)**	(0.050)**	(0.051)**	(0.051)**
$EIA_{ij,t+4}$					-0.102	
					(0.051)*	
$BRDR_{ij,1994}$	0.379	0.382	0.122	0.122	0.127	0.130
	(0.028)**	(0.028)**	(0.024)**	(0.024)**	(0.025)**	(0.024)**
$BRDR_{ij,1998}$	0.652	0.649	0.316	0.317	0.320	0.331
	(0.036)**	(0.037)**	(0.033)**	(0.032)**	(0.033)**	(0.033)**
$BRDR_{ij,2002}$	0.695	0.675	0.323	0.323	0.327	0.346
	(0.041)**	(0.045)**	(0.043)**	(0.042)**	(0.043)**	(0.043)**
EIA_TOTAL		0.313	0.523	0.522	0.552	0.522
		(0.115)**	(0.064)**	(0.064)**	(0.066)**	(0.064)**
N	6639	6639	6639	6724	6639	6639

Notes: This table reports panel gravity estimates with data on total manufacturing, 1990-2002. Panel 2A offers different variations of the main specification from Baier and Bergstrand (2007). In Panel 2B, we account for globalization trends using time-varying border variables $BRDR_{ij,t}$. Column $OLS1(+)$ presents OLS estimates with a single EIA covariate using only positive trade flows. In column (2), $OLS2(+)$, we allow for phasing–in of the EIA effects. In column (3), $PQML(+)$, we reproduce the results from column (2) using the PQML estimator and only positive observations. The estimates in column (4), $PQML$, use all observations in the sample. In column (5), $PQML$ *Lead*, we test to reverse causality by introducing a lead EIA effect. Specifications (1)-(5) are estimated with pair (ij), exporter-year (it), and importer-year (jt) fixed effects. Finally, the estimates in column (6), $PQML$ ($Eq.4$), are obtained with pair-fixed effects interacted with a time trend. Fixed effects estimates, including the constant, are not reported, for brevity. Robust standard errors, clustered by country pair, are in parentheses. + $p < 0.10$, * $p < .05$, ** $p < .01$.

Table 3: Sectoral Panel Gravity with Exporter-Year, Importer-Year, and Country-Pair FEs

	(1) Food	(2) Textile	(3) Wood	(4) Paper	(5) Chemicals	(6) Minerals	(7) Metals	(8) Machinery
3A.1. PQML EIA estimates (without Border Dummies, Phasing–in)								
$EIA_{ij,t}$	0.397	0.587	-0.131	-0.084	0.119	0.019	0.322	0.298
	(0.100)**	(0.107)**	(0.067)+	(0.042)*	(0.030)**	(0.050)	(0.049)**	(0.068)**
$EIA_{ij,t-4}$	0.219	0.399	0.130	0.066	0.165	0.274	0.215	0.361
	(0.035)**	(0.095)**	(0.036)**	(0.054)	(0.053)**	(0.055)**	(0.060)**	(0.078)**
$EIA_{ij,t-8}$	0.076	0.305	0.287	0.325	0.212	0.194	0.389	0.301
	(0.075)	(0.047)**	(0.065)**	(0.072)**	(0.071)**	(0.037)**	(0.061)**	(0.087)**
3A.2. PQML EIA estimates (with Border Dummies, Phasing–in)								
$EIA_{ij,t}$	0.301	0.332	-0.112	-0.171	0.043	-0.030	0.157	0.128
	(0.100)**	(0.063)**	(0.066)+	(0.061)**	(0.040)	(0.065)	(0.043)**	(0.065)*
$EIA_{ij,t-4}$	0.170	0.310	0.165	0.059	0.134	0.236	0.154	0.258
	(0.037)**	(0.075)**	(0.043)**	(0.047)	(0.042)**	(0.055)**	(0.047)**	(0.052)**
$EIA_{ij,t-8}$	0.051	0.240	0.325	0.311	0.219	0.188	0.353	0.197
	(0.077)	(0.049)**	(0.065)**	(0.069)**	(0.066)**	(0.043)**	(0.062)**	(0.081)*
$BRDR_{ij,1994}$	0.111	0.398	0.021	0.065	0.037	-0.020	0.172	0.145
	(0.024)**	(0.038)**	(0.046)	(0.024)**	(0.029)	(0.050)	(0.042)**	(0.027)**
$BRDR_{ij,1998}$	0.196	0.680	-0.022	0.187	0.200	0.100	0.385	0.409
	(0.037)**	(0.062)**	(0.040)	(0.065)**	(0.038)**	(0.059)+	(0.052)**	(0.043)**
$BRDR_{ij,2002}$	0.207	0.647	-0.074	0.141	0.122	0.076	0.357	0.499
	(0.053)**	(0.075)**	(0.051)	(0.062)*	(0.055)*	(0.076)	(0.054)**	(0.053)**
3B.1. OLS EIA estimates (without Border Dummies, Phasing–in)								
$EIA_{ij,t}$	0.281	0.280	0.118	0.252	0.048	0.154	0.245	0.187
	(0.075)**	(0.081)**	(0.094)	(0.096)**	(0.070)	(0.076)*	(0.098)*	(0.091)*
$EIA_{ij,t-4}$	0.165	0.312	0.012	0.221	0.089	0.178	0.292	0.096
	(0.072)*	(0.064)**	(0.101)	(0.087)*	(0.062)	(0.080)*	(0.098)**	(0.081)
$EIA_{ij,t-8}$	0.075	0.150	0.245	0.111	0.133	0.035	0.211	-0.007
	(0.099)	(0.100)	(0.127)+	(0.128)	(0.093)	(0.101)	(0.141)	(0.099)
3B.2. OLS EIA estimates (with Border Dummies, Phasing–in)								
$EIA_{ij,t}$	0.262	0.248	0.073	0.225	0.025	0.138	0.231	0.141
	(0.075)**	(0.069)**	(0.080)	(0.089)*	(0.068)	(0.077)+	(0.099)*	(0.066)*
$EIA_{ij,t-4}$	0.159	0.300	-0.008	0.210	0.077	0.171	0.288	0.074
	(0.073)*	(0.065)**	(0.093)	(0.087)*	(0.060)	(0.080)*	(0.099)**	(0.073)
$EIA_{ij,t-8}$	0.070	0.138	0.226	0.098	0.123	0.032	0.210	-0.027
	(0.100)	(0.100)	(0.124)+	(0.127)	(0.091)	(0.102)	(0.143)	(0.090)
$BRDR_{ij,1994}$	0.411	0.577	0.552	0.375	0.266	0.209	0.264	0.562
	(0.047)**	(0.039)**	(0.061)**	(0.053)**	(0.041)**	(0.049)**	(0.066)**	(0.039)**
$BRDR_{ij,1998}$	0.526	0.853	0.941	0.533	0.577	0.454	0.393	1.201
	(0.053)**	(0.048)**	(0.072)**	(0.062)**	(0.049)**	(0.050)**	(0.077)**	(0.047)**
$BRDR_{ij,2002}$	0.477	0.878	0.924	0.582	0.622	0.332	0.280	1.339
	(0.061)**	(0.057)**	(0.083)**	(0.077)**	(0.057)**	(0.062)**	(0.091)**	(0.058)**

Notes: This table reports panel gravity estimates for the eight 2-digit ISIC categories of manufactures, 1990-2002. Results are divided into four panels (3A.1, 3A.2, 3B.1, and 3B.2). Panels 3A.1 and 3A.2 report PQML results. Panels 3B.1 and 3B.2 provide OLS estimates. All specifications are estimated with pair (ij), exporter-year (it), and importer-year (jt) fixed effects and allow for phasing–in of the EIA effects. Fixed effects estimates, constants, and log-likelihood estimates are not reported, for brevity. Robust standard errors, clustered by country pair, are reported in parentheses. + $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 4: EIA Gravity Estimates using Aggregate Goods Trade Data

	PQML		OLS	
	No Glob.	Glob.	No Glob.	Glob.
$EIA_{ij,t}$	0.197	0.138	0.090	0.086
	(0.066)**	(0.065)*	(0.065)	(0.065)
$EIA_{ij,t-4}$	0.157	0.171	0.017	0.015
	(0.082)+	(0.071)*	(0.063)	(0.063)
$EIA_{ij,t-8}$	0.313	0.245	0.254	0.250
	(0.056)**	(0.036)**	(0.070)**	(0.070)**
$BRDR_{ij,1993}$		-0.094		-0.011
		(0.032)**		(0.031)
$BRDR_{ij,1997}$		0.139		0.194
		(0.045)**		(0.033)**
$BRDR_{ij,2001}$		0.214		0.285
		(0.050)**		(0.037)**
$BRDR_{ij,2005}$		0.277		0.361
		(0.052)**		(0.039)**
$BRDR_{ij,2009}$		0.120		0.266
		(0.057)*		(0.042)**
N	24993	24993	23896	23896

Notes: This table reports panel gravity estimates of the effects of economic integration agreements (EIAs, see text for definition) with aggregate data for the period 1989-2009. The left panel uses the $PQML$ estimator and the right panel uses the OLS estimator. All specifications allow for phasing–in of the EIA effects. Columns labeled "No Glob." use the specification of Baier and Bergstrand (2007). In columns labeled "Glob." we account for globalization. All specifications are estimated with pair (ij), exporter-year (it), and importer-year (jt) fixed effects. Fixed effects estimates are not reported, for brevity. Robust standard errors, clustered by country pair, are in parentheses. + $p < 0.10$, * $p < .05$, ** $p < .01$.

Table 5: EIA Effects and Distance

	Food	Textile	Wood	Paper	Chemicals	Minerals	Metals	Machinery	Total
5A. EIA estimates (Globalization, Robustness)									
$EIA_{ij,t}$	0.310	0.369	-0.124	-0.169	0.042	-0.040	0.172	0.151	0.110
	(0.102)**	(0.066)**	(0.066)+	(0.057)**	(0.037)	(0.063)	(0.043)**	(0.061)*	(0.038)**
$EIA_{ij,t-4}$	0.178	0.325	0.155	0.052	0.128	0.230	0.164	0.278	0.202
	(0.037)**	(0.076)**	(0.041)**	(0.046)	(0.043)**	(0.054)**	(0.048)**	(0.050)**	(0.052)**
$EIA_{ij,t-8}$	0.058	0.250	0.316	0.304	0.204	0.179	0.364	0.212	0.230
	(0.077)	(0.047)**	(0.065)**	(0.071)**	(0.067)**	(0.043)**	(0.062)**	(0.077)**	(0.050)**
$DIST_{ij,1994}$	0.015	0.048	0.004	0.010	0.006	-0.000	0.022	0.019	0.017
	(0.003)**	(0.005)**	(0.006)	(0.003)**	(0.003)+	(0.007)	(0.006)**	(0.003)**	(0.003)**
$DIST_{ij,1998}$	0.025	0.081	0.000	0.027	0.027	0.017	0.050	0.051	0.040
	(0.005)**	(0.008)**	(0.005)	(0.008)**	(0.005)**	(0.008)*	(0.007)**	(0.006)**	(0.004)**
$DIST_{ij,2002}$	0.026	0.077	-0.006	0.022	0.020	0.015	0.044	0.062	0.042
	(0.007)**	(0.009)**	(0.007)	(0.008)**	(0.007)**	(0.010)	(0.008)**	(0.007)**	(0.005)**
5B. EIA estimates (Globalization, Distance2)									
$EIA_{ij,t}$	0.339	0.438	-0.139	-0.130	0.038	-0.028	0.203	0.180	0.135
	(0.102)**	(0.070)**	(0.070)*	(0.046)**	(0.037)	(0.056)	(0.045)**	(0.054)**	(0.036)**
$EIA_{ij,t-4}$	0.204	0.378	0.126	0.063	0.145	0.250	0.215	0.348	0.244
	(0.036)**	(0.091)**	(0.039)**	(0.054)	(0.050)**	(0.054)**	(0.053)**	(0.057)**	(0.065)**
$EIA_{ij,t-8}$	0.073	0.290	0.284	0.317	0.197	0.187	0.397	0.288	0.263
	(0.076)	(0.045)**	(0.065)**	(0.070)**	(0.066)**	(0.035)**	(0.059)**	(0.070)**	(0.048)**
$DIST2_{ij,1994}$	0.043	0.155	-0.012	0.035	0.022	-0.019	0.062	0.061	0.052
	(0.009)**	(0.031)**	(0.016)	(0.012)**	(0.008)**	(0.012)	(0.029)*	(0.015)**	(0.010)**
$DIST2_{ij,1998}$	0.078	0.254	0.007	0.072	0.103	0.052	0.190	0.170	0.138
	(0.012)**	(0.049)**	(0.023)	(0.016)**	(0.012)**	(0.018)**	(0.040)**	(0.027)**	(0.015)**
$DIST2_{ij,2002}$	0.075	0.240	0.007	0.064	0.098	0.048	0.129	0.207	0.147
	(0.017)**	(0.042)**	(0.032)	(0.024)**	(0.022)**	(0.026)+	(0.039)**	(0.028)**	(0.017)**
5C. The Distance Puzzle									
$EIA_{ij,t}$	0.259	0.015	-0.145	0.023	0.047	-0.035	0.123	0.099	0.076
	(0.063)**	(0.062)	(0.063)*	(0.033)	(0.035)	(0.043)	(0.054)*	(0.048)*	(0.031)*
$EIA_{ij,t-4}$	0.088	0.154	-0.078	-0.015	0.068	0.065	0.204	0.110	0.105
	(0.041)*	(0.049)**	(0.049)	(0.034)	(0.037)+	(0.050)	(0.061)**	(0.039)**	(0.030)**
$EIA_{ij,t-8}$	0.127	0.061	-0.046	0.138	0.166	-0.176	0.075	-0.120	-0.017
	(0.089)	(0.075)	(0.074)	(0.076)+	(0.078)*	(0.082)*	(0.110)	(0.072)+	(0.057)
$DIST2_{ij,1994}$	-0.068	-0.065	-0.075	-0.007	0.032	-0.049	-0.066	-0.018	-0.022
	(0.027)*	(0.036)+	(0.037)*	(0.019)	(0.014)*	(0.025)+	(0.032)*	(0.016)	(0.013)+
$DIST2_{ij,1998}$	-0.085	-0.101	0.023	-0.047	0.081	0.019	0.021	0.005	0.007
	(0.034)*	(0.044)*	(0.045)	(0.022)*	(0.023)**	(0.027)	(0.060)	(0.023)	(0.021)
$DIST2_{ij,2002}$	-0.068	-0.078	0.054	-0.017	0.167	-0.009	-0.116	-0.012	0.025
	(0.043)	(0.059)	(0.050)	(0.032)	(0.032)**	(0.036)	(0.045)**	(0.027)	(0.024)

Notes: This table reproduces the results from panel 3A.2 of Table 3 with two additions. First, we report total manufacturing estimates in the last column of this table. Second, we use two different specifications to capture the effects of globalization. In Panel 5A, we replace the international border variables from our main analysis with interactions between the border variables and distance. In panel 5B, we use distance (including both internal and international distance) interacted with year dummies. Finally, in panel 5C, we use only the observations for international trade in our sample. All specifications are estimated with pair (ij), exporter-year (it), and importer-year (jt) fixed effects, which are not reported, for brevity. Robust standard errors are in parentheses. + $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.