Is the International Border Effect Larger than the Domestic Border Effect? Evidence from U.S. Trade

Cletus C. Coughlin and Dennis Novy*

October 2009

Abstract
Many studies have found that international borders represent large barriers to trade. But how do international borders compare to domestic border barriers? We investigate international and domestic border barriers in a unified framework. We consider a unique data set of exports from individual U.S. states to foreign countries and combine it with trade flows within and between U.S. states. After controlling for distance and country size, we find that relative to state-to-state trade, crossing an individual U.S. state's domestic border entails a *larger* trade barrier than crossing the international U.S. border. This finding highlights the concentration of trade flows at the local level and the importance of factors such as informational barriers and transportation costs even for the relatively short distances associated with state-to-state trade.

JEL classification: F10, F15
Keywords: International Border Effects, Intranational Home Bias, Domestic Borders, Gravity, Trade Costs

*Coughlin: Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166-0442, USA, coughlin@stls.frb.org. Novy: Department of Economics, University of Warwick, Coventry CV4 7AL, UK and CESifo, d.novy@warwick.ac.uk. We are very grateful to Lesli Ott for outstanding research assistance. Novy gratefully acknowledges research support from the Economic and Social Research Council, Grant RES-000-22-3112, and support as a visiting scholar at the Federal Reserve Bank of St. Louis.

1. Introduction

In a seminal paper, McCallum (1995) found that Canadian provinces trade up to 22 times more with each other than with U.S. states. This astounding result, also known as the *international border effect*, has led to a large literature on the trade impediments associated with international borders. More recently, Anderson and van Wincoop (2003) revisited the U.S.-Canadian border effect with new micro-founded estimates. Although they are able to reduce the border effect considerably, there is widespread consensus that the international border remains a large impediment to trade.[1]

A parallel and somewhat smaller literature has documented that border effects also exist within a country, known as the *domestic border effect* or intranational home bias. For example, Wolf (2000) finds that trade within individual U.S. states is significantly larger than trade between U.S. states even after he controls for economic size, distance and a number of additional determinants. Likewise, Nitsch (2000) finds that domestic trade within the average European Union country is about ten times larger than trade with another EU country.[2]

It is important to understand the nature of domestic and international trade barriers since they might impede the integration of markets and have negative welfare consequences. Accurately identifying the magnitudes of border effects at the domestic and international levels is a necessary step for assessing their economic significance. The contribution of this paper is to merge the two strands of literature about border effects into a unified framework. We construct a unique data set that includes three tiers of U.S. trade flows: a) trade within individual U.S. states,

[1] Anderson and van Wincoop (2004) report 74 percent as an estimate of representative international trade costs for industrialized countries (expressed as a tariff equivalent). About two-thirds of these costs can be attributed to border-related trade barriers such as tariffs and non-tariff barriers. The remainder represents transportation costs.
[2] An earlier study by Wei (1996) finds similar results for OECD countries. Nikolaus Wolf (2009) finds sizeable domestic border barriers in the historical context for Germany in the late 19th and early 20th centuries. Chen (2004) documents significant intra-European Union border effects at the industry level.

e.g., Minnesota-Minnesota; b) trade between U.S. states, e.g., Minnesota-Texas; and c) trade between U.S. states and foreign countries, e.g., Minnesota-Canada.

We use gravity theory to estimate the relative size of the domestic and international border effects. As is typical in the literature, the domestic border effect indicates how much a U.S. state trades with itself relative to state-to-state trade, while the international border effect indicates how much a U.S. state trades with foreign countries relative to state-to-state trade. After controlling for distance and country size, we find that relative to state-to-state trade, crossing an individual U.S. state's domestic border entails a *larger* trade barrier than crossing the international U.S. border. That is, the domestic border effect is bigger than the international border effect.

As an example, consider exports from Minnesota to Texas and Canada (see Figure 1). Although Texas and Canada have roughly the same gross domestic products, during the year 2002 Minnesota exported about twice as much to Texas as to Canada ($5.7bn vs. $2.9bn). This gap is the familiar international border effect. However, in the same year Minnesota traded over ten times as much with itself as with Texas ($69.1bn vs. $5.7bn). This gap is the domestic border effect, and it is bigger than the international gap, both in absolute and relative terms.

By using a more complete data set of trade flows, we overcome one of the shortcomings of previous studies. Specifically, our paper addresses the potential sample selection bias that arises if domestic border effect studies ignore international trade flows. Thus, if we estimate the domestic border effect ignoring international trade flows, we find a domestic border effect that is smaller than the one we obtain based on the complete sample. This downward bias arises because the estimation systemically leaves out pairs with relatively low trade flows so that the

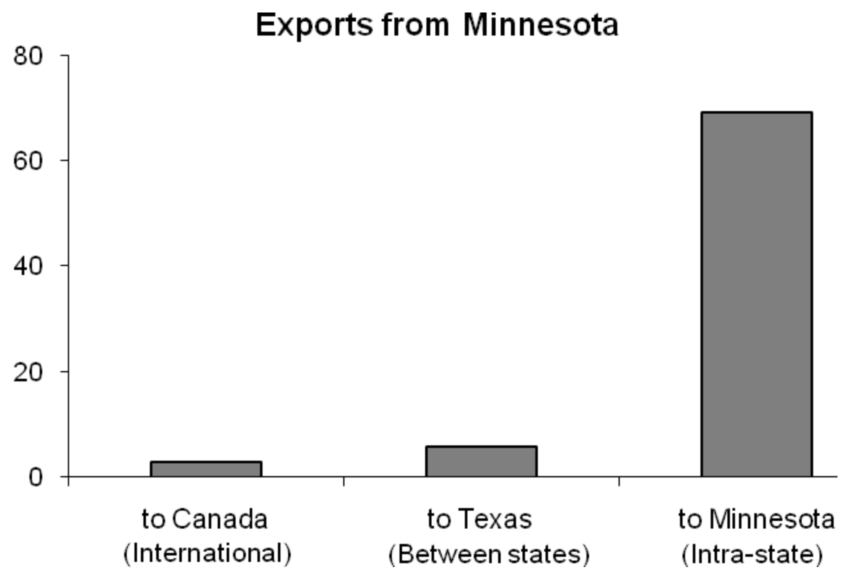

Figure 1: Exports from Minnesota to three destinations. Data for the year 2002, in bn U.S. dollars.

extent of the domestic border effect is underestimated.[3]

What are the economic reasons behind the large domestic border effect? International trade economists traditionally emphasize trade barriers associated with international borders such as tariffs, bureaucratic hurdles and informational barriers.[4] However, beginning with Wolf (2000) and Nitsch (2000), the empirical literature has also demonstrated that borders within a country are associated with a significant trade-impeding effect. One potential explanation is related to work by Hillberry and Hummels (2008). Based on ZIP-code level domestic U.S. trade flows, they document that trade within the United States is heavily concentrated at the local level. In particular, trade within a single ZIP code is on average three times higher than trade with partners outside the ZIP code. This concentration might be due to the prevalence of trade in

[3] Similarly, there is also a bias if the international border effect is estimated ignoring internal trade flows (see section 4.2).

[4] See Anderson and van Wincoop (2004) for an overview.

intermediate goods at the local level, arguably as a result of supply chain optimization as companies seek to minimize transportation costs and suppliers co-locate with final goods producers. In turn, this high concentration of trade at the local level implies large domestic border barrier estimates. As we discuss in section 5, other reasons for the strong local concentration of trade include informational and search costs, for example in the form of business, social and immigration networks, increasing returns at the local level as well as location-specific tastes. The insight from our paper is that estimates of domestic border effects are also large when compared to international border barrier estimates.

The paper is organized as follows. In section 2 we carefully examine the economic theory of trade in general equilibrium with trade barriers and derive our empirical estimation framework. We show that theory does not provide an *a priori* expectation as to the relative size of the domestic and international border effects. We thus affirm that a relatively large domestic border effect—as we find it in the data—is by no means a foregone conclusion. In section 3 we describe our data set. In section 4 we jointly estimate the international and domestic border effects so that they are directly comparable. In section 5, we discuss a number of potential explanations for our empirical results. Section 6 concludes.

2. Gravity theory and the estimation framework

2.1 Gravity theory

The seminal contribution of McCallum (1995) has led to a large number of papers that estimate border effects based on a gravity estimation framework. Gravity theory describes how trade flows are determined in general equilibrium with trade barriers. It is well-known that gravity equations can be derived from a variety of trade models (see Feenstra, Markusen and

Rose, 2001; Evenett and Keller, 2002). Gravity is also consistent with more recent trade theories, including the gravity framework with multilateral resistance by Anderson and van Wincoop (2003), the Ricardian trade model by Eaton and Kortum (2002), Chaney's (2008) extension of the Melitz (2003) heterogeneous firms model as well as the heterogeneous firms model by Melitz and Ottaviano (2008) with a linear demand system.[5] To obtain results that are easily comparable to the previous literature on border effects, we adopt the widely used gravity framework by Anderson and van Wincoop (2003). Our results, however, would also be generated with the other frameworks.

Anderson and van Wincoop's (2003) parsimonious model rests on the Armington assumption that countries produce differentiated goods and trade is driven by consumers' love of variety. They derive the following gravity equation for exports x_{ij} from region i to region j:

$$(1)\quad x_{ij} = \frac{y_i y_j}{y^W} \left(\frac{t_{ij}}{\Pi_i P_j} \right)^{\sigma-1},$$

where y_i and y_j denote output of regions i and j, y^W denotes world output, t_{ij} is the bilateral trade cost factor (one plus the tariff equivalent), Π_i is the outward multilateral resistance term and P_j is the inward multilateral resistance term. The parameter $\sigma > 1$ is the elasticity of substitution. The bilateral trade costs t_{ij} capture a variety of trade frictions such as transportation costs, tariffs and bureaucratic barriers and they also include the border barriers.

2.2 The estimation framework

We follow McCallum (1995) and other authors by hypothesizing that trade costs t_{ij} are a loglinear function of geographic distance, $dist_{ij}$, and a border dummy, $INTERNATIONAL_{ij}$, which takes on the value 1 whenever regions i and j are located in different countries. In addition, we

[5] See Chen and Novy (2008) for an overview.

hypothesize that domestic trade costs within a region's own territory might be systematically different from bilateral trade costs. We therefore include an ownstate dummy variable, $OWNSTATE_{ij}$, that takes on the value 1 for $i=j$. Our trade cost function can thus be expressed as

$$(2)\ \ln(t_{ij}) = \beta' \, INTERNATIONAL_{ij} + \gamma' \, OWNSTATE_{ij} + \delta' \ln(dist_{ij}),$$

where β' and γ' reflect the international and the ownstate (i.e., domestic) border effects, respectively, and δ' is the elasticity of trade costs with respect to distance.

The trade cost function (2) nests the trade cost functions used by Wolf (2000), Hillberry and Hummels (2003) and Anderson and van Wincoop (2003). Wolf (2000) and Hillberry and Hummels (2003) only consider trade flows within the U.S. so that an international border effect cannot be estimated. This corresponds to $\beta'=0$ in equation (2). Anderson and van Wincoop (2003) follow McCallum's (1995) specification that does not allow for a domestic border effect ($\gamma'=0$).

We log-linearize equation (1) so that we obtain:

$$(3)\ \ln(x_{ij}) = \ln(y_i) + \ln(y_j) + (1-\sigma)\ln(t_{ij}) + (\sigma-1)\ln(\Pi_i P_j) + \alpha_t + \varepsilon_{ij},$$

where the logarithm of world output is captured by year dummies α_t and where we add a white-noise error term ε_{ij}. Substituting the trade cost function (2) yields the following estimating equation:

$$\begin{aligned}(4)\ \ln(x_{ij}) = {} & \ln(y_i) + \ln(y_j) + \beta \, INTERNATIONAL_{ij} + \gamma \, OWNSTATE_{ij} \\ & + \delta \ln(dist_{ij}) + (\sigma-1)\ln(\Pi_i P_j) + \alpha_t + \varepsilon_{ij},\end{aligned}$$

where $\beta=(1-\sigma)\beta'$, $\gamma=(1-\sigma)\gamma'$ and $\delta=(1-\sigma)\delta'$. In the empirical analysis we control for multilateral resistance by fixed and random effects in various specifications.

2.3 Border effects in theory

The empirical literature typically finds that international borders impede trade. This corresponds to β<0 in estimating equation (4). Trading within a state is typically associated with higher trade flows, corresponding to γ>0. We first examine whether gravity theory allows us to predict whether the international border effect β is larger or smaller in absolute value than the domestic border effect γ, i.e., whether $|\beta| \gtrless |\gamma|$.

As we explain below in more detail, our data set comprises three tiers of trade flows:

a) Ownstate trade: trade flows within a U.S. state, for example within Minnesota, such that $OWNSTATE_{ij}=1$ and $INTERNATIONAL_{ij}=0$,

b) National trade: trade flows between two U.S. states, for example from Minnesota to Texas, such that $OWNSTATE_{ij}= INTERNATIONAL_{ij}=0$, and

c) International trade: trade flows from a U.S. state to a foreign country, for example from Minnesota to Canada, such that $OWNSTATE_{ij}=0$ and $INTERNATIONAL_{ij}=1$.

The second tier is thus the omitted category in equation (4), implying that the ownstate border effect is estimated relative to the benchmark of trade between U.S. states. We choose this benchmark to obtain coefficients that are directly comparable to those in the literature (Wolf, 2000; Nitsch, 2000). The sign and magnitude of the ownstate border effect can therefore be gauged by comparing trade costs t_{ii} within a typical U.S. state i to bilateral trade costs t_{ij} with another U.S. state j. We draw this comparison by considering their ratio t_{ii}/t_{ij}. Abstracting from any stochastic elements, equations (1) and (2) imply that this ratio is given by

$$\frac{t_{ii}}{t_{ij}} = \left(\frac{x_{ij}}{x_{ii}} \frac{y_i}{y_j} \right)^{\frac{1}{\sigma-1}} \frac{P_i}{P_j} = \frac{\exp(\gamma')(dist_{ii})^{\delta'}}{(dist_{ij})^{\delta'}}.$$

Using $\gamma'=\gamma/(1-\sigma)$ and $\delta'=\delta/(1-\sigma)$ this can be rewritten as

$$(5)\ \ \exp(\gamma) = \frac{x_{ii}}{x_{ij}} \frac{y_j}{y_i} \left(\frac{P_j}{P_i} \right)^{\sigma-1} \left(\frac{dist_{ij}}{dist_{ii}} \right)^{\delta}.$$

As a simple example, first assume the symmetric case where $y_i=y_j$, $P_i=P_j$ and $dist_{ii}=dist_{ij}$. A positive ownstate effect $\gamma>0$ would follow only if $x_{ii}/x_{ij}>1$. Now assume the more realistic case where bilateral distance $dist_{ij}$ exceeds domestic distance $dist_{ii}$. Given that the distance elasticity of trade is negative ($\delta<0$), an even bigger ratio x_{ii}/x_{ij} would be required to ensure $\gamma>0$. More generally, we conclude that given the distance element of trade costs as well as the output and multilateral resistance variables, the sign and magnitude of the domestic border effect parameter γ will primarily depend on the extent of domestic trade x_{ii} relative to bilateral trade x_{ij}.

As in the literature we also use the benchmark of trade between U.S. states for estimating the international border effect. To gauge its sign and magnitude, we compare bilateral trade costs t_{ik} between a typical U.S. state i and a typical foreign country k to trade costs t_{ij} between two U.S. states. Their ratio is given by

$$\frac{t_{ik}}{t_{ij}} = \left(\frac{x_{ij}}{x_{ik}} \frac{y_k}{y_j} \right)^{\frac{1}{\sigma-1}} \frac{P_k}{P_j} = \frac{\exp(\beta')(dist_{ik})^{\delta'}}{(dist_{ij})^{\delta'}},$$

or

$$(6)\ \ \exp(\beta) = \frac{x_{ik}}{x_{ij}} \frac{y_j}{y_k} \left(\frac{P_j}{P_k} \right)^{\sigma-1} \left(\frac{dist_{ij}}{dist_{ik}} \right)^{\delta}.$$

As before, assume the simple symmetric case where $y_k=y_j$, $P_k=P_j$ and $dist_{ik}=dist_{ij}$. A negative international border effect $\beta<0$ would follow only if $x_{ik}/x_{ij}<1$. In the more common case where international distance $dist_{ik}$ (say, between Minnesota and Japan) exceeds inter-state distance $dist_{ij}$ (say, between Minnesota and Texas), an even smaller ratio x_{ik}/x_{ij} would be required to ensure $\beta<0$. Given distances as well as the output and multilateral resistance variables, the international border effect parameter β will therefore mainly depend on the extent of international trade x_{ik} relative to inter-state trade x_{ij}.

Thus, equations (5) and (6) can in principle yield either sign for γ and β. The fact that most empirical studies find $\gamma>0$ and $\beta<0$ is consistent with but by no means implied by gravity theory. Neither does gravity theory make a prediction about the absolute magnitudes of β and γ. *A priori* we can therefore not infer whether $|\beta|\gtreqless|\gamma|$.

3. Data

To obtain comparable results, we use the same data sets as Wolf (2000) and Anderson and van Wincoop (2003) for domestic trade flows within the United States. The novelty of our approach is to combine these domestic trade flows with international trade flows from individual U.S. states to the 50 largest U.S. export destinations. Thus, our data set comprises, for instance, trade flows within Minnesota, exports from Minnesota to Texas as well as exports from Minnesota to Canada. We take data quality seriously and below, we describe in detail the data sources, potential concerns and how we address these concerns.

3.1 Domestic exports: Commodity Flow Survey

For our measures of the shipments of goods within and across U.S. states, we use data from the Commodity Flow Survey, which is a joint effort of the Bureau of Transportation Statistics and the Bureau of the Census. We use survey results from 1993, 1997, and 2002.[6] The survey covers the origin and destination of shipments of manufacturing, mining, wholesale trade, and selected retail establishments. The survey excludes shipments in the following sectors:

[6] This survey was most recently conducted in 2007. Preliminary data are available; final data are scheduled to be available in December 2009.

services, crude petroleum and natural gas extraction, farm, forestry, fishery, construction, government, and most retail. Shipments from foreign establishments are also excluded; import shipments are excluded until they reach a domestic shipper. U.S. export (i.e., trans-border) shipments are also excluded. For each shipment, the survey collects information such as value, weight, transportation mode, origin, and destination.[7]

3.2 International exports: Origin of Movement

Our data on exports by U.S. states to foreign destinations are from the Origin of Movement series.[8] These data are compiled by the Foreign Trade Division of the U.S. Bureau of the Census. The data in this series identify the state from which an export begins its journey to a foreign country; however, we would like to know the state is which the export was produced. Below we provide details on the Origin of Movement series and its suitability as a measure of the origin of production.[9]

Beginning in 1987, the Origin of Movement series provides the current-year export sales, or free-alongside-ship (f.a.s.) costs if not sold, for 54 ‘states’ to 242 foreign destinations. These export sales are for merchandise sales only and do not include services exports. The 54 ‘states’ include the 50 U.S. states plus the District of Columbia, Puerto Rico, U.S. Virgin Islands, and unknown. Following Wolf (2000), we use the 48 contiguous U.S. states. Rather than all 242

[7] Erlbaum and Holguin-Veras (2006) note that sample size has been a major issue. The 1993 survey collected data from 200,000 establishments and the size was subsequently reduced to 100,000 in 1997 and 50,000 in 2002. In response to complaints from the freight data users community, the sample size was increased to 100,000 in 2007.
[8] Other studies that have used the Origin of Movement series include Smith (1999), Coughlin and Wall (2003), Coughlin (2004) and Cassey (2007).
[9] The highlighted details as well as much additional information can be found in Cassey (2006).

destinations, we use the 50 leading export destinations for U.S. exports for 2005.[10] We use the annual data from 1993, 1997, and 2002 for total merchandise exports.[11]

Concerns about using the Origin of Movement series to identify the location of production are especially pertinent for agricultural and mining exports.[12] We, however, focus on manufactured goods. Cassey (2006) has examined the issue of the coincidence of the state origin of movement and the state of production for manufactured goods.[13] The reason for restricting the focus to manufacturing is that the best source for location-based data on export production, "Exports from Manufacturing Establishments," covers only manufacturing.[14]

Cassey's key finding relevant to our analysis is that overall, the Origin of Movement data is of sufficient quality to be used as the origin of the production of exports. Nonetheless, the data for specific states may not be of sufficient quality as the origin of production. These states are: Alaska, Arkansas, Delaware, Florida, Hawaii, New Mexico, South Dakota, Texas, Vermont, and Wyoming. He recommends the removal of Alaska and Hawaii in particular. As we use the 48 contiguous U.S. states, our data set is consistent with this recommendation. The next two candidates for removal would be Delaware and Vermont. Cassey further highlights that the consolidation of export shipments might systematically affect the Origin of Movement estimates (relative to the origin of production). Specifically, consolidation tends to bias upward the

[10] In alphabetical order, these countries are Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Colombia, Costa Rica, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Finland, France, Germany, Guatemala, Honduras, Hong Kong, India, Indonesia, Ireland, Israel, Italy, Japan, Kuwait, Malaysia, Mexico, Netherlands, New Zealand, Norway, Panama, Peru, Philippines, Russia, Saudi Arabia, Singapore, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Thailand, Turkey, United Arab Emirates, United Kingdom, and Venezuela.

[11] We have also tried the data for manufacturing only (as opposed to total merchandise). The two series are very highly correlated (99 percent). The regression results are almost identical and we therefore do not report them.

[12] See http://www.trade.gov/td/industry/otea/state/technote.html.

[13] For the initial work on this issue, see Coughlin and Mandelbaum (1991) and Cronovich and Gazel (1999). As Cassey's (2006) analysis refers to manufactured goods, we note that we have also tried the Origin of Movement manufacturing data (as opposed to total merchandise) with virtually identical results.

[14] The data in the "Exports from Manufacturing Establishments" is available at http://www.census.gov/mcd/exports/ but does not contain destination information, so it cannot be used for the current research project.

estimates for Florida and Texas and to bias downward the estimates for Arkansas and New Mexico.[15] As a robustness check, we drop these states from the sample (see section 4.3).

3.3 Adjustments to the trade data

Our simultaneous use of the intra-state and inter-state shipments data from the Commodity Flow Survey and the merchandise international trade data from the Origin of Movement series requires an adjustment to increase the comparability of these data sets. Such an adjustment arises because of three important differences between the data sources. First, the merchandise international trade data measures a shipment from the source to the port of exit just once, whereas the commodity flow data likely measures a good in a shipment more than once. For example, a good may be shipped from a plant to a warehouse and, later, to a retailer. Second, goods destined for foreign countries, when they are shipped to a port of exit, are included in domestic shipments. Third, the coverage of sectors differs between the data sources. The Commodity Flow Survey includes shipments of manufactured goods, but it excludes agriculture and part of mining. Meanwhile, the merchandise trade data includes all goods.[11]

Identical to Anderson and van Wincoop (2003), we scale down the data in the Commodity Flow Survey by the ratio of total domestic merchandise trade to total domestic shipments from the Commodity Flow Survey. Total domestic merchandise trade is approximated by gross output in the goods-producing sectors (i.e., agriculture, mining, and manufacturing) minus international merchandise exports.[16] This calculation yields an adjustment factor of 0.495

[15] The problems for Wyoming and South Dakota are primarily in individual sectors—chemicals for Wyoming and computers in South Dakota.
[16] See Helliwell (1997, 1998) and Wei (1996).

for 1993, 0.508 for 1997, and 0.430 for 2002.[17] Similar to Anderson and van Wincoop (2003), we point out that our adjustment to the commodity flow data does not solve all the measurement problems but it is the best feasible option.

3.4 Other data

The rest of the data used in our estimations can be characterized as well-known. For individual U.S. states we use state gross domestic product data from the U.S. Bureau of Economic Analysis. For foreign countries, we use data on gross domestic product taken from the IMF World Economic Outlook Database (October 2007 edition).

We use the standard great circle distance formula to cover inter-state and international distances between capital cities in kilometers. As intra-state distance, we use the distance between the two largest cities in a state. As alternatives for intra-state distance, we also try the measure used by Wolf (2000) that weights the distance between a state's two largest cities by their population, as well as the measure suggested by Nitsch (2000) that is based on land area (see section 4.2 for details).

4. Empirical results

First, we show that our data exhibit a substantial domestic border effect, as established by Wolf (2000). In separate regressions, we also show that the data exhibit a substantial international border effect, as established by McCallum (1995). Second, we combine the domestic U.S. trade data with the international observations. This allows us to estimate the

[17] The difference between our adjustment factor for 1993 and that of Anderson and van Wincoop--0.495 vs. 0.517--is due to data revision.

domestic and international border effects in a joint framework and directly compare them. Finally, we carry out a number of robustness checks.

Due to the data quality concerns for Alaska, Hawaii and Washington, D.C., we drop these states so that we are left with 48 contiguous states. We form a balanced sample over the years 1993, 1997 and 2002. This yields 1801 trade observations per cross-section within the U.S.[18] Adding 50 foreign countries as export destinations increases the number of trade observations by 2338 so that our sample includes 4139 observations per cross-section, or 12417 in total.[19]

4.1 Domestic and international border effects estimated separately

In columns 1 and 2 of Table 1 we replicate the intranational home bias. As Wolf (2000), in column 1 we only use data for 1993. In column 2 we add the data for 1997 and 2002. Like Hillberry and Hummels (2003) we use exporter and importer fixed effects such that the output regressors drop out. Our estimates are very close to Wolf's baseline coefficient of 1.48 for the ownstate indicator variable. The interpretation of this coefficient is that given distance and economic size, ownstate trade is 4.4 times higher than state-to-state trade (exp(1.48) = 4.4). We conclude that our data are characterized by the typical intranational home bias in trade. Hillberry and Hummels (2003) reduce the ownstate coefficient by about a third when excluding wholesale shipments from the Commodity Flow Survey data. The reason is that wholesale shipments are predominantly local so that their removal disproportionately reduces the extent of ownstate

[18] The maximum possible number of U.S. observations would be 48*48 = 2304. The missing observations are due to the fact that Commodity Flow Survey estimates did not meet publication standards because of high sampling variability or poor response quality.

[19] The maximum possible number of international observations would be 48*50 = 2400. Sixty-two observations are missing mainly because exports to Malaysia were generally not reported in 1993. Only 18 of the observations not included in our sample are most likely zeros (as opposed to missing). We therefore deem it highly unlikely that omitting these few zeros distorts our results.

trade.[20] However, Nitsch (2000) reports higher home bias coefficients by comparing trade within European Union countries to trade between EU countries. He finds home bias coefficients in the range of 1.8 to 2.9.

In columns 3-6 we do not consider ownstate trade and instead focus on the international border effect, with country fixed effects controlling for multilateral resistance. Motivated by the prominence of Canadian data in the border effect literature, in columns 3 and 4 we present regression results for a subsample that only includes Canada as a foreign country. Columns 5 and 6 include all 50 foreign countries. In column 4 we estimate an international border coefficient of -1.34, implying that after we control for distance and economic size, exports from U.S. states to Canada are about 74 percent lower than exports to other U.S. states (exp(-1.34) = 0.26). This estimate is comparable to the estimate of -1.65 obtained by Anderson and van Wincoop (2003, Table 2). In the full sample with 50 foreign countries, however, the international border effect is notably smaller. We will now revisit these findings by considering the complete sample of domestic and international trade flows.

4.2 Is the international border effect larger than the domestic border effect?

In Table 2 we estimate the domestic and international border effects jointly so that their magnitudes are directly comparable. For this purpose, we combine ownstate and international trade flows whilst keeping inter-state trade as the reference group as in Table 1. When we pool the data over the years 1993, 1997 and 2002 in columns 2 and 4, we add random effects instead

[20] We do not have access to the private-use coding of wholesale shipments and thus cannot replicate their finding with our data. However, our main result on the relative size of the domestic and international border effects would seem robust to a reduction by a third in the ownstate coefficient magnitudes (see Tables 2-4). Hillberry and Hummels (2003) further reduce the ownstate coefficient by using an alternative distance measure that is based on actual shipping distances. We refer to section 4.3 where we employ such a measure, but our main result is unchanged.

of country fixed effects. The reason is that country fixed effects are perfectly collinear with the ownstate indicator variable. Exporter and importer fixed effects would also be impractical because those are perfectly collinear with the international border dummy. This arises because the foreign countries in our data set are only importers but never exporters.

Columns 1 and 2 report regression results for the Canadian subsample. The joint estimation yields substantial domestic border coefficients (1.75 and 1.78). But although crossing the international border impedes trade, the coefficients are low and even insignificant in the pooled regression (-0.33 and -0.03). A test of whether the absolute values of the domestic and international border coefficients are equal, $|OWNSTATE_{ij}|=|INTERNATIONAL_{ij}|$, is clearly rejected (p-value = 0.00). Columns 3 and 4 report results for the full sample. Both the ownstate and the international coefficients are larger in absolute magnitude than in columns 1 and 2. The hypothesis that the two border coefficients are equal is again clearly rejected (p-value = 0.00). In summary, a key finding in Table 2 is that the domestic border effect is larger than the international border effect. That is, relative to inter-state trade, crossing an individual U.S. state's domestic border entails a larger trade barrier than crossing the international U.S. border.

A second key finding is that the joint estimation of domestic and international border effects yields substantially different estimates than when these effects are estimated separately. For example, the coefficient on $OWNSTATE_{ij}$ is 1.48 (Table 1, column 2) when estimated separately, and 2.05 (Table 2, column 4) when estimated jointly. Meanwhile, the coefficient on $INTERNATIONAL_{ij}$ shows an even larger change. This coefficient changes from -0.05 (Table 1, column 6) when estimated separately to -1.10 (Table 2, column 4) when estimated jointly.

Why do these differences arise? We can think of them in terms of a sample selection problem. The ownstate border regressions in Table 1 systematically leave out observations that trade the least (i.e., international observations). Thus, while trade flows within states are still very high, comparatively they appear more moderate if international observations are ignored. This imparts a downward bias, leading to an underestimated domestic border effect. By the same token, the domestic border effect is also smaller in the Canadian subsample (Table 2, columns 1 and 2) than in the full sample (Table 2, columns 3 and 4).

In a similar vein, the international border regressions in columns 5 and 6 of Table 1 systematically leave out those observations that are characterized by exceptionally high trade flows (i.e., the ownstate observations). International trade flows therefore do not seem as low as they would in the complete sample that includes ownstate observations as in columns 3 and 4 of Table 2, leading to a downward bias in the magnitude of the international border effect.

4.3 Robustness

Various authors have pointed out that the estimation of border effects is sensitive to how distance is measured, see Helliwell and Verdier (2001) and Head and Mayer (2009). For example, if the economic distance within a U.S. state is much shorter than indicated by the conventional measures—perhaps because economic activity is highly concentrated in two nearby cities—then it might no longer be surprising if that state trades considerably more within its boundaries than with partners further away. To address this concern we employ three alternative distance measures that have been suggested in the literature.

Columns 1 and 4 of Table 3 use the alternative measure for ownstate distance proposed by Wolf (2000). This measure weights the distance between a state's two largest cities by their population. It thus better reflects heavy concentration of economic activity in relatively small areas. For example, most economic activity in Utah is concentrated around Salt Lake City such that the conventional great circle distance measure could easily overstate actual shipping distances. As expected, on average this alternative measure results in shorter ownstate distances (109 km vs. 179 km) so that it reduces the domestic border effects compared to Table 2. For example, the coefficient on $OWNSTATE_{ij}$ declines from 1.78 (Table 2, column 2) to 1.37 (Table 3, column 1) for the Canadian sample, and from 2.05 (Table 2, column 4) to 1.64 (Table3, column 4) for the 50-country sample. Despite the smaller magnitudes of the domestic border effects, they are still significantly different from the absolute values of the international border estimates in columns 1 and 4 of Table 3 (the p-values are 0.00 and 0.02). [21] Note also that compared to Table 2, the results for the international border effects in Table 3 are virtually the same. A similar comment can be made concerning the coefficients estimated for distance.

In columns 2 and 5 we employ a measure of ownstate distance that is based on land area, as proposed by Nitsch (2000). His measure is based on a hypothetical circular economy with three equal-sized cities, one in the centre and the other two on opposite sides of the circle. The average internal distance of such an economy, and also other economies with more complex structures, can be approximated by the radius of the circle. In the data this is computed as $1/\sqrt{\pi}=$ 0.56 times the square root of the area in km^2 and on average results in roughly similar ownstate distances (170 km vs. 179 km). Nevertheless, the ownstate dummy estimates increase slightly compared to the results in Table 2. They increase to 1.95 from 1.78 for the Canada sample and to 2.23 from 2.05 for the 50-country sample.

In columns 3 and 6 we employ a third alternative distance measure that is closer to actual shipping distances observed within the U.S. Based on private-use Commodity Flow Survey data at the ZIP code level, Hillberry and Hummels (2003, equation 4 and table 1) provide a statistical

[21] Helliwell and Verdier (2001) develop more sophisticated measures of ownstate distance that account for the population distribution within states more accurately. However, their approach yields ownstate distances that are higher than usual, implying larger domestic border effects.

relationship between the distance measure used by Wolf (2000), an ownstate dummy and an adjacency dummy. They estimate the following equation:

$$(7)\ \ \ln(\textit{actual dist}_{ij}) = \lambda_1 \ln(\textit{Wolf dist}_{ij}) + \lambda_2 OWNSTATE_{ij} + \lambda_3 adjacency_{ij} + e_{ij}$$

with λ_1=0.821, λ_2=-0.498 and λ_3=-0.404. We use these coefficients to approximate actual shipping distances within the U.S. and to Canada and Mexico, and we then use them as an explanatory variable. The resulting distances are on average considerably shorter compared to the great circle distances, both within U.S. states (18 km vs. 179 km) as well as between U.S. states and to Canada and Mexico (450 km vs. 1556 km). The distances to overseas countries are not affected. As a consequence, the ownstate coefficient in column 3 is cut substantially but its absolute size (0.69) is still significantly larger than that of the international border coefficient (-0.07). In column 6 of Table 3, both the ownstate and the international border coefficients are reduced in magnitude to 1.50 from 2.05 for the ownstate coefficient and to -0.24 from -1.10 for the international coefficient such that their absolute difference remains highly significant.

In Table 4 we conduct further robustness checks. Hillberry and Hummels (2008) document that trade is highly concentrated at the local level and that it consists to some extent of local wholesale shipments. In columns 1 and 4 we drop all state-to-state observations that are less than 200 miles apart to check whether they distort the sample. This excludes 100 cross-sections. But the regression results are virtually the same as in columns 2 and 4 of Table 2.

As we explain in section 3.2, Cassey (2006) raises doubts whether the Origin of Movement data are sufficiently similar to the actual origin of production in the case of Delaware, Vermont, Florida, Texas, Arkansas, and New Mexico. In columns 2 and 5 we therefore drop these six states from our sample. But the regression results are overall quite similar to those in Table 2.

In columns 3 and 6 we follow Wolf (2000) in adding an adjacency dummy that takes on the value 1 whenever two states are neighboring (say, Minnesota and Wisconsin). Like Wolf (2000) we find that adding an adjacency dummy reduces the ownstate coefficient. But the domestic border effect nevertheless remains stronger in absolute value than the international border effect. However, we can no longer reject the hypothesis that their absolute values are equal (p-value = 0.30).

Finally, in column 7 we restrict the sample to Mexico as the sole export destination. The ownstate coefficient is similar to the one for the Canadian subsample in Tables 2-4. But the international border effect is more pronounced than for Canada, suggesting larger border-related trade costs with Mexico. Yet, the ownstate coefficient is significantly larger in absolute size than the international border coefficient.

Overall, we conclude that although the point estimates of the domestic and international border effects can change depending on which distance measure and which subsample we use, it is a robust feature of the data that the absolute magnitude of the domestic border effect exceeds that of the international border effect. Their difference is highly significant in almost all specifications.

5. Discussion

We discuss a number of potential explanations for our empirical result that the domestic border effect is comparatively large. One major explanation is related to the work by Hillberry and Hummels (2008). Based on ZIP-code level domestic U.S. trade flows, they document that trade within the United States is heavily concentrated at the local level. Trade within a single ZIP code is on average three times higher than trade with partners outside the ZIP code. As a major reason they point out the co-location of producers in supply chains to minimize transportation costs and exploit informational spillovers. The local concentration of trade might also be related to increasing returns to scale, production externalities and associated agglomeration effects (see Rossi-Hansberg, 2005), as well as to hub-and-spoke distribution systems and wholesale shipments (see Hillberry and Hummels, 2003). Such spatial clustering of economic activity can lead to large domestic border barrier estimates, as we find it in our results.[22]

The concentration of trade at the local level is also borne out in other types of data. Using individual transactions data from online auction websites, Hortaçsu, Martínez-Jerez and Douglas (2009) find that purchases tend to be disproportionately concentrated within a short distance perimeter, with many counterparties based in the same city. Some of these purchases can be explained by their location-specific nature, for example in the case of opera tickets. But the evidence also suggests that direct contract enforcement in case of breach and lack of trust may be major reasons behind the same-city bias, which the authors subsume under 'contracting costs.' They also find evidence for culture and local tastes as factors that shape the local concentration of trade. For example, the same-city effect is most pronounced for local interest items such as sports memorabilia (also see Blum and Goldfarb, 2006).

Business networks and immigration patterns might also be related to strong trade flows between relatively close locations. Combes, Lafourcade and Mayer (2005) report that business and immigrant networks significantly facilitate trade within France. They cite the reduction of information costs and the diffusion of preferences as two main economic mechanisms through which networks may operate. This includes the reduction of search costs associated with

[22] The concentration of trade at the local level might also be related to firms' slicing up their production chains (multi-stage production and vertical specialization). Yi (2008) offers an explanation of the border effect using the vertical specialization argument in a Ricardian framework.

matching buyers and sellers. As an additional facilitating factor for trade, Rauch and Trindade (2002) also mention the possibility of community sanctions that could be imposed amongst members of an ethnic network. In the context of the border effect in U.S. data, Millimet and Osang (2007) find that incorporating migration flows within the U.S. diminishes the estimated intranational home bias. Business and immigrant networks therefore likely play an important role in explaining the trade-reducing effect of distance.

Evans (2006) introduces an important dimension into the study of international border effects. In contrast to the majority of the literature that is based on aggregate data, she distinguishes between the intensive and extensive margins of trade. This is motivated by the well-established fact that only a fraction of goods produced domestically is also exported. To isolate the portion of the international border effect that is attributable to the extensive margin, she re-estimates the international border effect whilst restricting production data to those firms that actually sell abroad. Overall, she finds that reductions in each margin explain about one-half of the conventional aggregate international border effect. However, it is an open question whether this result would also obtain for the domestic border effect. Evidence by Hillberry and Hummels (2008) shows that the extensive margin within the U.S. is much more sensitive to distance than the intensive margin. The extensive margin might therefore play a relatively larger role in explaining the domestic border effect.

6. Conclusion

We collect a data set of U.S. exports that combines three types of trade flows: trade within an individual state (Minnesota-Minnesota), trade between U.S. states (Minnesota-Texas) and trade flows from an individual U.S. state to a foreign country (Minnesota-Canada). This data set allows us to jointly estimate the effect on trade of crossing an international border and the effect of crossing the domestic border.

Our paper sheds new light on the interpretation of border effects. We qualify the general view that international border effects, as estimated in standard gravity regressions, are abnormally large. While we obtain point estimates typically found in the literature, we show that the international border effect is in fact smaller than the domestic border effect. This result is robust to alternative distance measures and different subsamples. It highlights the effect of trade frictions at relatively short distances. In addition, by using a more complete data set that involves ownstate, national and international trade flows of U.S. states, we also address the potential sample selection bias that arises if domestic border effect studies ignore international trade flows. Specifically, we find that not taking into account international trade flows imparts a downward bias on estimated domestic border effects.

References

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

Anderson, J., van Wincoop, E., 2004. Trade Costs. *Journal of Economic Literature* 42, pp. 691-751.

Blum, M., Goldfarb, A., 2006. Does the Internet Defy the Law of Gravity? *Journal of International Economics* 70, pp. 384-405.

Cassey, A., 2006. State Export Data: Origin of Movement vs. Origin of Production. Unpublished manuscript, University of Minnesota.

Cassey, A., 2007. State Export Behavior and Barriers. Unpublished manuscript, University of Minnesota.

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

Chen, N., Novy, D., 2008. International Trade Integration: A Disaggregated Approach. CEPR Discussion Paper #7103.

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

Coughlin, C., 2004. The Increasing Importance of Proximity for Exports from U.S. States. *Federal Reserve Bank of St. Louis Review* 86, pp. 1-18.

Coughlin, C., Mandelbaum, T., 1991. Measuring State Exports: Is There a Better Way? *Federal Reserve Bank of St. Louis Review* 73, pp. 65-79.

Coughlin, C., Wall, H., 2003. NAFTA and the Changing Pattern of State Exports. *Papers in Regional Science* 82, pp. 427-450.

Cronovich, R., Gazel, R., 1999. How Reliable Are the MISER Foreign Trade Data? Unpublished manuscript.

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

Erlbaum, N., Holguin-Veras, J., 2006. Some Suggestions for Improving CFS Data Products. Commodity Flow Survey Conference, July 8-9, 2005, Transportation Research Circular, Number E-C088, pp. 77-97.

Evans, C., 2006. Border Effects and the Availability of Domestic Products Abroad. *Canadian Journal of Economics* 39, pp. 211-246.

Evenett, S., Keller, W., 2002. On Theories Explaining the Success of the Gravity Equation. *Journal of Political Economy* 110, pp. 281-316.

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

Head, K., Mayer, T., 2009. Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade. In: The Gravity Model in International Trade: Advances and Applications. Editors: Bergeijk and Brakman, Cambridge University Press.

Helliwell, J., 1997. National Borders, Trade and Migration. *Pacific Economic Review* 2, pp. 165-185.

Helliwell, J., 1998. How Much Do National Borders Matter? Washington, D.C., Brookings Institution.

Helliwell, J., Verdier, G., 2001. Measuring Internal Trade Distances: A New Method Applied to Estimate Provincial Border Effects in Canada. *Canadian Journal of Economics* 34, pp. 1024-1041.

Hillberry, R., Hummels, D., 2003. Intranational Home Bias: Some Explanations. *Review of Economics and Statistics* 85, pp. 1089-1092.

Hillberry, R., Hummels, D., 2008. Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data. *European Economic Review* 52, pp. 527-550.

Hortaçsu, A., Martínez-Jerez, F., Douglas, J., 2009. The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre. *American Economic Journal: Microeconomics* 1, pp. 53- 74.

McCallum, J., 1995. National Borders Matter: Canada-U.S. Regional Trade Patterns. *American Economic Review* 85, pp. 615-623.

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

Melitz, M., Ottaviano, G., 2008. Market Size, Trade, and Productivity. *Review of Economic Studies* 75, pp. 295-316.

Millimet, D., Osang, T., 2007. Do State Borders Matter for U.S. Intranational Trade? The Role of History and Internal Migration. *Canadian Journal of Economics* 40, pp. 93-126.

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

Rauch, J., Trindade, V., 2002. Ethic Chinese Networks in International Trade. *Review of Economics and Statistics* 84, pp. 116-130.

Rossi-Hansberg, E., 2005. A Spatial Theory of Trade. *American Economic Review* 95, pp. 1464-1491.

Smith, P., 1999. Are Weak Patent Rights a Barrier to U.S. Exports? *Journal of International Economics* 48, pp. 151-177.

Wei, S., 1996. Intra-National Versus International Trade: How Stubborn are Nations in Global Integration? NBER Working Paper #5531.

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

Yi, K., 2008. Can Multi-Stage Production Explain the Home Bias in Trade? *American Economic Review*, forthcoming.

Wolf, N., 2009. Was Germany Ever United? Evidence from Intra- and International Trade 1885-1933. *Journal of Economic History* 69, pp. 846-881.

Table 1: Domestic and international border effects, estimated separately

Sample	U.S. only		U.S. and Canada		U.S. and 50 countries	
Year	1993	1993, 1997, 2002	1993	1993, 1997, 2002	1993	1993, 1997, 2002
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(y_i)$			1.11**	0.99**	1.20**	0.98**
			(0.03)	(0.03)	(0.04)	(0.04)
$\ln(y_j)$			1.08**	0.98**	0.89**	0.73**
			(0.03)	(0.03)	(0.03)	(0.04)
$\ln(dist_{ij})$	-1.08**	-1.07**	-1.14**	-1.13**	-1.17**	-1.17**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$OWNSTATE_{ij}$	1.46**	1.48**				
	(0.20)	(0.19)				
$INTERNATIONAL_{ij}$			-0.79**	-1.34**	-0.40*	-0.05
			(0.10)	(0.17)	(0.23)	(0.22)
National trade (reference group)	yes	yes	yes	yes	yes	yes
Ownstate trade	yes	yes	no	no	no	no
International trade	no	no	yes	yes	yes	yes
Observations	1801	5403	1801	5403	4091	12273
Clusters	--	1801	--	1801	--	4091
Fixed effects	yes	yes	yes	yes	yes	yes
R-squared	0.90	0.90	0.88	0.88	0.87	0.87

Notes: The dependent variable is $\ln(x_{ij})$. OLS estimation. Robust standard errors are reported in parentheses, clustered around country pairs ij. Exporter and importer fixed effects in columns 1 and 2; country fixed effects in columns 3-6; time-varying in columns 2, 4 and 6. Constants and year dummies are not reported. * significant at 10% level. ** significant at 1% level.

Table 2: Domestic and international border effects, estimated jointly

Sample	U.S. and Canada		U.S. and 50 countries	
Year	1993	1993, 1997, 2002	1993	1993, 1997, 2002
	(1)	(2)	(3)	(4)
$\ln(y_i)$	0.96**	0.93**	1.28**	1.21**
	(0.02)	(0.02)	(0.02)	(0.02)
$\ln(y_j)$	0.93**	0.90**	0.82**	0.82**
	(0.02)	(0.02)	(0.01)	(0.01)
$\ln(dist_{ij})$	-0.94**	-0.93**	-0.83**	-0.82**
	(0.03)	(0.03)	(0.03)	(0.03)
$OWNSTATE_{ij}$	1.75**	1.78**	2.04**	2.05**
	(0.19)	(0.19)	(0.20)	(0.20)
$INTERNATIONAL_{ij}$	-0.33**	-0.03	-1.24**	-1.10**
	(0.10)	(0.10)	(0.06)	(0.05)
$\lvert OWNSTATE_{ij}\rvert=\lvert INTERNATIONAL_{ij}\rvert$	[0.00]	[0.00]	[0.00]	[0.00]
National trade (reference group)	yes	yes	yes	yes
Ownstate trade	yes	yes	yes	yes
International trade	yes	yes	yes	yes
Observations	1849	5547	4139	12417
Clusters	--	1849	--	4139
Random effects	no	yes	no	yes
R-squared	0.82	0.82	0.79	0.79

Notes: The dependent variable is $\ln(x_{ij})$. OLS estimation. Robust standard errors are reported in parentheses, clustered around country pairs ij. Random effects in columns 2 and 4. Constants and year dummies are not reported. ** significant at 1% level. The numbers in brackets report p-values for the test $\lvert OWNSTATE_{ij}\rvert=\lvert INTERNATIONAL_{ij}\rvert$.

Table 3: Robustness - Alternative distance measures

Sample	U.S. and Canada			U.S. and 50 countries		
Years: 1993, 1997, 2002						
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(y_i)$	0.92**	0.93**	0.93**	1.21**	1.21**	1.21**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\ln(y_j)$	0.90**	0.91**	0.90**	0.81**	0.82**	0.79**
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
$\ln(dist_{ij})$: Wolf (2000)	-0.89**			-0.80**		
	(0.03)			(0.03)		
$\ln(dist_{ij})$: Nitsch (2000)		-0.97**			-0.84**	
		(0.03)			(0.03)	
$\ln(dist_{ij})$: Actual shipping distance			-1.01**			-0.76**
			(0.03)			(0.02)
$OWNSTATE_{ij}$	1.37**	1.95**	0.69**	1.64**	2.23**	1.50**
	(0.21)	(0.15)	(0.21)	(0.22)	(0.17)	(0.20)
$INTERNATIONAL_{ij}$	-0.02	-0.03	-0.07	-1.13**	-1.06**	-0.24**
	(0.10)	(0.10)	(0.10)	(0.06)	(0.05)	(0.07)
$\lvert OWNSTATE_{ij}\rvert = \lvert INTERNATIONAL_{ij}\rvert$	[0.00]	[0.00]	[0.01]	[0.02]	[0.00]	[0.00]
National trade (reference group)	yes	yes	yes	yes	yes	yes
Ownstate trade	yes	yes	yes	yes	yes	yes
International trade	yes	yes	yes	yes	yes	yes
Observations	5547	5547	5547	12417	12417	12417
Clusters	1849	1849	1849	4139	4139	4139
Random effects	yes	yes	yes	yes	yes	yes
R-squared	0.81	0.83	0.83	0.78	0.79	0.79

Notes: The dependent variable is $\ln(x_{ij})$. OLS estimation. Robust standard errors are reported in parentheses, clustered around country pairs ij. Random effects in all columns. Constants and year dummies are not reported. ** significant at 1% level. The numbers in brackets report p-values for the test |OWNSTATEij|=|INTERNATIONALij|.

Table 4: Further robustness checks

Sample	U.S. and Canada			U.S. and 50 countries			U.S. and Mexico
Years: 1993, 1997, 2002	Distance > 200 m.	Fewer states	Adjacency	Distance > 200 m.	Fewer states	Adjacency	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\ln(y_i)$	0.93**	0.90**	0.93**	1.22**	1.21**	1.21**	0.93**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\ln(y_j)$	0.91**	0.86**	0.92**	0.82**	0.81**	0.81**	0.91**
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
$\ln(dist_{ij})$	-0.97**	-0.92**	-0.71**	-0.81**	-0.78**	-0.65**	-0.94**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$adjacency_{ij}$			0.93**			1.11**	
			(0.06)			(0.07)	
$OWNSTATE_{ij}$	1.67**	1.68**	1.48**	2.08**	2.03**	1.48**	1.75**
	(0.20)	(0.21)	(0.17)	(0.20)	(0.22)	(0.19)	(0.19)
$INTERNATIONAL_{ij}$	-0.06	0.01	-0.16	-1.10**	-1.21**	-1.28**	-0.64**
	(0.10)	(0.11)	(0.10)	(0.06)	(0.06)	(0.06)	(0.11)
$\|OWNSTATE_{ij}\|=\|INTERNATIONAL_{ij}\|$	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.30]	[0.00]
National trade (reference group)	yes	yes	yes	yes	yes	yes	yes
Ownstate trade	yes	yes	yes	yes	yes	yes	yes
International trade	yes	yes	yes	yes	yes	yes	yes
Observations	5247	4362	5547	12117	10368	12417	5547
Clusters	1749	1454	1849	4039	3456	4139	1849
Random effects	yes	yes	yes	yes	yes	yes	yes
R-squared	0.82	0.82	0.84	0.78	0.79	0.79	0.82

Notes: The dependent variable is $\ln(x_{ij})$. OLS estimation. Robust standard errors are reported in parentheses, clustered around country pairs ij. Random effects in all columns. Constants and year dummies are not reported. ** significant at 1% level. The numbers in brackets report p-values for the test |OWNSTATEij|=|INTERNATIONALij|. Columns 1 and 4 drop all pairs of U.S. states that are less than 200 miles apart. Columns 2 and 5 drop states with inferior data quality (AR, DE, FL, NM, TX, VT), see Section 3.2. Columns 3 and 6 add an adjacency dummy that is one if two regions have a land border. Column 7 drops all other foreign countries except Mexico.