INTERNATIONAL ECONOMIC REVIEW
Vol. 62, No. 4, November 2021 DOI: 10.1111/iere.12520

# ESTIMATING BORDER EFFECTS: THE IMPACT OF SPATIAL AGGREGATION*

By Cletus C. Coughlin and Dennis Novy

*Federal Reserve Bank of St. Louis, U.S.A.; University of Warwick, U.K.*

Trade data are typically aggregated across space. In this article, we investigate the sensitivity of gravity estimation to spatial aggregation. We build a model in which micro regions are aggregated into macro regions. We then apply the model to the large literature on border effects in domestic and international trade. Our theory shows that aggregation leads to border effect heterogeneity. Larger regions and countries are systematically associated with smaller border effects. We test our theory with aggregate and industry-level trade flows for U.S. states. Our results confirm the model's predictions, with strong heterogeneity patterns.

## 1. INTRODUCTION

By how much do borders impede international trade? It has been a major objective of research in international trade to identify the frictions that hinder the international integration of markets, and many policymakers across the globe are keen on reducing them.

Ever since the seminal paper by McCallum (1995), many researchers have used the gravity equation as a workhorse model to estimate so-called border effects. The aim is to estimate by how much borders reduce international trade. In their simplest form, gravity equations with border dummies are estimated based on aggregate bilateral trade data. As aggregates, these data combine the trade flows of spatial subunits (such as boroughs, municipalities and counties) into trade flows at higher levels of spatial aggregation (such as regions, states, and countries). The question we address in this article is how this process of aggregation affects the estimation of border effects. How do border effects depend on the spatial units we find in any given data set? Put differently, how do border effects depend on the way we slice up the map?

To understand the effects of spatial aggregation, we build a theoretical framework based on a large number of "micro" regions that trade with each other subject to spatial frictions.

*Manuscript received May 2019; revised March 2021.

A previous version of this paper was entitled "Domestic and International Border Effects: State Size Matters." Jacob Haas and Lesli Ott provided excellent research assistance. Alexander Klein generously shared data with us. We thank the editor and three anonymous referees for constructive comments. We are grateful for comments on the previous version by James Anderson, Esther Bøler, Oana Furtuna, Jason Garred, Keith Head, John Helliwell, David Jacks, Chris Meissner, Peter Neary, Krishna Pendakur, Frédéric Robert-Nicoud, Trevor Tombe, Nikolaus Wolf, as well as seminar participants at Simon Fraser University, Kiel International Economics Meeting, University of Sussex, ETSG, Midwest Trade Conference at Penn State, Federal Reserve Bank of St. Louis, CESifo Measuring Economic Integration Conference, Loughborough University Workshop on Distance and Border Effects, ESRC Scottish Centre on Constitutional Change, SIRE Workshop on Country Size and Border Effects in a Globalised World, Edinburgh, LACEA/LAMES Meetings, North American Meetings of the Regional Science Association, Higher School of Economics in Saint Petersburg, NUI Maynooth, Dundee, Strathclyde, Lancaster, London School of Economics, Oldenburg, UC Louvain, Tilburg, Hong Kong University of Science and Technology, National University of Singapore, University of Amsterdam, Essex, OECD, UC Davis, RMET, Oxford, St. Gallen, Hitotsubashi. Novy gratefully acknowledges support as a visiting scholar at the Federal Reserve Bank of St. Louis and research support from the Economic and Social Research Council (ESRC grant ES/P00766X/1) as well as the Centre for Competitive Advantage in the Global Economy (CAGE, ESRC grant ES/L011719/1) at the University of Warwick. The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors. Please address correspondence to: DENNIS NOVY, Department of Economics, University of Warwick, Coventry CV$ 7AL, U.K. E-mail: d.novy@warwick.ac.uk.

We then aggregate these regions into larger "macro" regions. Due to the spatial frictions, the more micro regions we combine, the more we increase the cost of trading within the newly aggregated macro regions. As a result, aggregation increases the relative cost of trading within borders as opposed to across borders.

Our theory shows how this shift in relative cost leads to heterogeneous border effect estimates: smaller regions are associated with relatively strong border effects, and larger regions are associated with relatively weak border effects. We call this the *spatial attenuation effect*. It represents an important testable implication for the estimation of border effects. A second implication is that as standard border effects are averages of the underlying individual border effects, sample composition effects occur. That is, samples that happen to include many large regions (or countries) tend to have moderate border effects, and vice versa. A third implication is that when regions are combined into larger aggregates, their associated estimated border effects should weaken in magnitude.

In the empirical part of the article, we test these theoretical implications with domestic and international trade flows at the level of U.S. states, both for aggregate data and at the industry level. Our results confirm the model's predictions, in particular the systematic heterogeneity of border effects across states. For example, we find that for a large state like California, removing the U.S. international border would lead to an increase of bilateral trade on average by only 13%, whereas for a small state like Wyoming trade would go up over four times as much (61%). We also find evidence of systematic sample composition effects. This means that border effects are typically not directly comparable across studies as samples inevitably vary as they contain different sets of spatial units. We also carry out a hypothetical scenario of aggregating U.S. states into larger spatial units such as the nine Census divisions defined by the U.S. Census Bureau. Consistent with our model, we obtain smaller estimated border effects at the level of Census divisions. Overall, we find that spatial aggregation has a strong, first-order quantitative impact on border effects.

The estimated border effects represent the direct impact of border frictions on trade flows. But in general equilibrium there are also indirect effects that operate through price indices. In the trade literature, these are typically referred to as multilateral resistance effects, as highlighted by Anderson and van Wincoop (2003). We demonstrate that our mechanism of spatial aggregation is separate from such price index effects. In our model, due to the symmetric location of micro regions, every location faces the same price index, and aggregation does not affect this equilibrium structure. We therefore obtain border effect heterogeneity that is not related to price index effects. In the data, when we have to keep track of varying price indices across space, we find that the heterogeneity of border effects stemming from spatial aggregation clearly dominates the heterogeneity coming from general equilibrium effects. Moreover, we use our model to numerically simulate border effects. This further illustrates the spatial attenuation effect in operation. Finally, we explore potential determinants of border effects, namely transport infrastructure.

Our theory and empirical results on spatial aggregation apply to both branches of the border effects literature: the international border effect and the domestic border effect. McCallum (1995) found that Canadian provinces trade up to 22 times more with each other than with U.S. states. This astounding result has led to a large literature on the trade impediments associated with international borders. Anderson and van Wincoop (2003) famously revisit the U.S.–Canadian border effect with new theory-consistent estimates. Although they are able to reduce the border effect considerably, the international border remains a large impediment to trade. Havránek and Iršová (2017) provide an overview of this extensive literature.[1]

[1] Anderson and van Wincoop (2004) report 74% as an estimate of representative international trade costs for industrialized countries (expressed as a tariff equivalent). Hillberry (2002) and Chen (2004) document significant but varying border effects at the industry level. Anderson and van Wincoop (2004), Section 3.8, provide guidance and intuition for border effects in the case of aggregation across industries with industry-specific elasticities of substitution and possibly also industry-specific border barriers. We provide industry-level evidence in Section 4.9.

A parallel and somewhat smaller literature has explored the existence of border effects within a country, known as the domestic border effect or intranational home bias. For example, Wolf (2000) and Millimet and Osang (2007) find that after controlling for economic size, distance and a number of additional determinants, trade within individual U.S. states is significantly larger than trade between U.S. states. Similarly, Nitsch (2000) finds that domestic trade within the average European Union country is about 10 times larger than trade with another EU country. Nitsch and Wolf (2013) find a persistent domestic border effect between East and West Germany that has declined only slowly after reunification. Wrona (2018) documents an east–west border effect within Japan.

Our approach is inspired by Hillberry and Hummels (2008) who find empirically that border effects at the ZIP code level within the United States would be enormous, by far eclipsing the magnitude of traditional border effects typically found in the literature. Havránek and Iršová's (2017) meta-analysis of border effects highlights a related pattern, that is, smaller economies tend to have stronger estimated international border effects than larger economies. To the best of our knowledge, our article is the first in the literature to provide a formal explanation of these patterns. We show that the underlying mechanism operating through spatial aggregation applies to both domestic and international border effects.

Our results on heterogeneous border effects and spatial attenuation illustrate a more general issue known in the geography literature as the Modifiable Areal Unit Problem.[2] Briant et al. (2010) systematically highlight this problem for empirical work in economic geography. In the context of Canadian provincial border effects, Bemrose et al. (2020) find that these border effects decline when geographic units become more similar in size and shape. Our contribution is to provide a theoretical foundation for spatial aggregation that allows us to obtain precise analytical results for the size of estimated border effects. Our article can thus be seen as an attempt to apply the general notion of the Modifiable Areal Unit Problem to the specific context of gravity estimation of border effects.

Our results highlight theoretically and empirically that border effect coefficients capture a relative cost, that is, the cost of trading across relative to within borders (also see Agnosteva et al. (2019)), and that this relative cost systematically shifts with the size of spatial units. More generally, our article is also related to the recent literature in international trade that explicitly models internal trade costs (Ramondo et al., 2016)), or models space as a continuum (Allen and Arkolakis, 2014). Ramondo et al. (2016) are concerned with endogenous growth models and thus they address a distinct set of questions. However, in their framework as in ours, it is key to move away from the crude assumption of zero internal trade costs. As do they, we depart from the assumption that a country is fully integrated domestically with zero trade costs, in which case it can no longer be treated as a single dot.

The article is organized as follows. In Section 2, we briefly describe the typical estimation of border effects in the literature. In Section 3, we use the theoretical gravity framework to outline the general problem of spatial aggregation. We then present a formal model of spatial aggregation with a closed-form solution for the international border effect. In Section 4, we take the theory to the data and explore the testable implications, using domestic and international trade flows for U.S. states. We also discuss multilateral resistance effects in general equilibrium and explore transport infrastructure as a potential determinant of border effects. In Section 5, we discuss the implications of our analysis for the interpretation of border effects, in particular to what extent they could be seen as statistical artifacts. We conclude in Section 6 by providing practical recommendations for border effect estimation. In the Online Appendix, we provide further details on the theory including a separate model of spatial aggregation for the domestic border effect, and we also describe our data and provide numerical simulations.

[2] See Fotheringham and Wong (1991).

## 2. BORDER EFFECTS IN GRAVITY ESTIMATION

The seminal contribution of McCallum (1995) has led to a large number of papers that estimate border effects based on a gravity framework. We follow the canonical structural gravity model by Anderson and van Wincoop (2003). They derive their model from an endowment economy under the Armington assumption of goods differentiated by country of origin. It is well known that a near-isomorphic gravity structure can be derived from different types of trade models.[3] We first briefly review how domestic and international border effects are typically defined in the literature.

2.1. *The Structural Gravity Framework.* We adopt the widely used structural gravity framework by Anderson and van Wincoop (2003). They derive the following gravity equation for the value of exports $x_{ij}$ from region $i$ to region $j$:

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

where $y_i$ and $y_j$ denote nominal income of regions $i$ and $j$, and $y^W$ denotes world income. The bilateral trade cost factor is given by $t_{ij} \geq 1$ (one plus the tariff equivalent). It is assumed symmetric for any given pair (i.e., $t_{ij} = t_{ji}$). $P_i$ and $P_j$ are the multilateral resistance terms, which can be interpreted as average trade barriers of regions $i$ and $j$.[4] The parameter $\sigma > 1$ is the elasticity of substitution across goods from different countries. There are $N$ regions in the sample.

In the theory and the data, we will deal with three different tiers of trade flows: *international* trade flows that cross an international border, *domestic* bilateral trade flows between different regions of the same country, and *internal* trade flows within regions.[5]

2.2. *The Trade Cost Function.* We follow McCallum (1995) and other authors by hypothesizing that trade costs $t_{ij}$ are a log-linear function of bilateral geographic distance $dist_{ij}$, and an international border barrier represented by the dummy $INT_{ij}$ that takes on the value 1 whenever regions $i$ and $j$ are located in different countries, and zero otherwise. The $INT_{ij}$ variable is therefore an *international border dummy*. We also include a dummy variable $DOM_{ij}$ for bilateral domestic trade flows that takes on the value 1 whenever regions $i$ and $j$ are in the same country but distinct ($i \neq j$), and 0 otherwise. In a sample without international flows, we therefore refer to the $DOM_{ij}$ dummy as the *domestic border dummy* as the case of $DOM_{ij} = 1$ implies that a domestic border has been crossed.[6]

We can express our trade cost function as

$$\ln\left(t_{ij}^{1-\sigma}\right) = \beta INT_{ij} + \gamma DOM_{ij} + \rho \ln\left(dist_{ij}\right), \tag{2}$$

[3] Head and Mayer (2014) state the structural gravity framework more generally. Arkolakis, Costinot and Rodríguez-Clare (2012) analyze the properties of the underlying Armington model in more detail. In addition, they demonstrate under which conditions near-isomorphic gravity equations hold for Ricardian trade models such as Eaton and Kortum (2002) and trade models with heterogeneous firms such as Melitz (2003) and Chaney (2008).

[4] As explained by Anderson and van Wincoop (2003), footnote 12, symmetry between outward and inward multilateral resistance price indices implies a particular normalization. Alternative normalizations are possible such that outward and inward multilateral resistance terms differ but they would differ by a factor of proportionality that is constant across countries.

[5] Some authors use *domestic* to describe trade flows within a region. We stick to *internal* here.

[6] The $DOM_{ij}$ dummy corresponds to the ‘ownstate’ dummy in Hillberry and Hummels (2003) and the ‘home’ dummy in Nitsch (2000), with the 0 and 1 coding swapped.

where $\beta$ and $\gamma$ are dummy coefficients, and $\rho$ is the distance elasticity of trade. We log-linearize gravity equation (1) and insert the trade cost function (2) to obtain

$$\ln(x_{ij}) = \ln(y_i) + \ln(y_j) - \ln(y^W) + \ln(P_i^{\sigma-1}) + \ln\left(P_j^{\sigma-1}\right) + \beta INT_{ij} + \gamma DOM_{ij} + \rho \ln(dist_{ij}), \tag{3}$$

where $\beta$ and $\gamma$ are the coefficients of interest used to compute border effects.[7] Both coefficients are typically found to be negative, and we will reproduce such standard estimates in the empirical Section 4. The distance elasticity is typically estimated around $\rho \approx -1$ based on OLS estimation and closer to $\rho \approx -0.7$ with PPML estimation.

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

## 3. A THEORY OF SPATIAL AGGREGATION

We now explore formally how border dummy coefficients are affected when regions are spatially aggregated. We first describe the problem of spatial aggregation in general terms before turning to a more specific model.

3.1. *The General Problem of Spatial Aggregation.* Our modeling strategy is to imagine a world of many "micro" regions as the basic spatial unit. We then aggregate these micro regions into larger "macro" regions that more closely resemble those we observe in the data. We can think of large regions as a cluster of many micro regions combined. For example, we can imagine California as a cluster of a fairly large number of micro regions, but in comparison Vermont is a cluster of only a few micro regions.

3.1.1. *The basic framework.* We model the world as consisting of an arbitrary number of small micro regions denoted by the subscripts $k$ and $l$. Each region is endowed with a differentiated good as in the Armington framework of Anderson and van Wincoop (2003). As in expression (1), the standard gravity equation for trade flows $x_{kl}$ holds:

$$x_{kl} = \frac{y_k y_l}{y^W}\left(\frac{t_{kl}}{P_k P_l}\right)^{1-\sigma}, \tag{4}$$

where bilateral trade costs are assumed symmetric (i.e., $t_{kl} = t_{lk}$). We assume the same trade cost function as in expression (2):

$$t_{kl}^{1-\sigma} = \exp(\zeta BORDER_{kl})\, dist_{kl}^{\rho}, \tag{5}$$

where $\zeta BORDER_{kl}$ is a stand-in that could equal either $\beta INT_{kl}$ for the international border effect or $\gamma DOM_{kl}$ for the domestic border effect.

How can we measure a border effect? It is implied by the comparison of a trade flow across the border relative to a trade flow within the border. For example, we can compare $x_{kl'}/x_{kl}$

[7] For instance, suppose $\beta = -0.5$. As a back-of-the-envelope calculation ignoring price index effects, all else equal international trade flows would only be 61% as large as other trade flows as $\exp(-0.5) = 0.61$. In partial equilibrium, this would typically be interpreted as a border effect equivalent to a reduction of international trade by 39%.

where trade of region $k$ with region $l'$ crosses the border and trade with region $l$ does not.[8] Using Equations (4) and (5), we can solve for the implied border dummy coefficient as

$$\exp(\zeta) = \frac{x_{kl'}}{x_{kl}} \frac{y_l/P_l^{1-\sigma}}{y_{l'}/P_{l'}^{1-\sigma}} \frac{dist_{kl}^{\rho}}{dist_{kl'}^{\rho}}. \tag{6}$$

That is, the border barrier coefficient $\zeta$ is a function of relative trade flows, other relative trade cost components (represented by distance) and region-specific factors (income and multilateral resistance price indices). The implied border coefficient is by construction the same for all possible pairings of trade flows across and within the border, and equivalent versions of expression (6) would hold for all these pairings, not just for the particular pairing $x_{kl'}/x_{kl}$. In other words, the gravity system (4) and (5) implies a common border coefficient that can be identified with standard regression methods.[9] We will now see how spatial aggregation changes this conclusion.

3.1.2. *Aggregation.* We aggregate micro regions into macro regions. Micro regions with a $k$ subscript are aggregated into a macro region denoted with an $i$ subscript, and micro regions with an $l$ subscript are aggregated into a macro region with a $j$ subscript. The resulting aggregate trade flow between macro region $i$ and macro region $j$ follows as the sum of all underlying micro flows, given by

$$x_{ij} = \sum_{k\epsilon i} \sum_{l\epsilon j} x_{kl} = \sum_{k\epsilon i} \sum_{l\epsilon j} \frac{y_k y_l}{y^W} \left( \frac{t_{kl}}{P_k P_l} \right)^{1-\sigma}. \tag{7}$$

If $i$ and $j$ are made up of exactly the same set of micro regions, then $x_{ii} = x_{jj}$ is an internal flow. If $i$ and $j$ are disjoint sets, then $x_{ij}$ represents a bilateral flow with $x_{ij} = x_{ji}$. We assume there is no aggregation of micro regions across countries. That is, all micro regions aggregated into $i$ belong to the same country, and the same holds for the micro regions subsumed into $j$.

To simplify Equation (7) we define the weight parameter $\theta_{k,i} \equiv (y_k/P_k^{1-\sigma})/(y_i/P_i^{1-\sigma})$ so that we arrive at

$$x_{ij} = \frac{y_i y_j}{y^W} \frac{\sum_{k\epsilon i} \sum_{l\epsilon j} \theta_{k,i} \theta_{l,j} t_{kl}^{1-\sigma}}{(P_i P_j)^{1-\sigma}}.$$

When we further define aggregate trade costs $t_{ij}$ as a weighted CES trade cost index with

$$t_{ij}^{1-\sigma} \equiv \sum_{k\epsilon i} \sum_{l\epsilon j} \theta_{k,i} \theta_{l,j} t_{kl}^{1-\sigma}, \tag{8}$$

we obtain

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

This equation looks the same as the standard gravity equation (1). But the key difference is that aggregate trade costs $t_{ij}$ are endogenous. They contain the weight parameters $\theta_{k,i}$ and $\theta_{l,j}$

[8] In case of the domestic border effect, we would have $l = k$ such that the benchmark trade flow becomes $x_{kk}$.

[9] For example, a log-linearized gravity specification as in Equation (3) can be estimated where region-specific factors are absorbed by fixed effects. See the empirical analysis in Section 4.

that are functions of endogenous income and price index terms. Inserting trade cost function (5) into the trade cost index, we further find

$$t_{ij}^{1-\sigma} = \exp(\zeta BORDER_{ij})\, \widetilde{dist}_{ij}^{\rho} \tag{9}$$

with

$$\widetilde{dist}_{ij}^{\rho} \equiv \sum_{k\epsilon i}\sum_{l\epsilon j} \theta_{k,i}\theta_{l,j} dist_{kl}^{\rho}, \tag{10}$$

where $\widetilde{dist}_{ij}$ is a CES distance index with endogenous weight parameters.[10]

To measure the border effect with spatially aggregated data, we perform the same type of comparison as in Equation (6). We obtain

$$\exp(\zeta) = \frac{x_{ij'}}{x_{ij}} \frac{y_j/P_j^{1-\sigma}}{y_{j'}/P_{j'}^{1-\sigma}} \frac{\widetilde{dist}_{ij}^{\rho}}{\widetilde{dist}_{ij'}^{\rho}}, \tag{11}$$

where $x_{ij'}$ is an aggregate cross-border trade flow and $x_{ij}$ is an aggregate within-border flow. As with micro regions, this type of comparison yields the same common border coefficient for all possible pairings but only if we use the theoretically consistent distance index correctly constructed as indicated by Equation (10). If researchers use alternative distance measures—perhaps because the relevant micro-level distances and the corresponding weights in Equation (10) are not observable—then the implied border coefficients will generally deviate from the true underlying micro friction $\zeta$.

If we set Equations (6) and (11) equal, we get a condition for obtaining the same common implied border coefficients from micro-level and macro-level data. This condition is

$$\frac{x_{kl'}}{x_{kl}} \frac{y_l/P_l^{1-\sigma}}{y_{l'}/P_{l'}^{1-\sigma}} \frac{dist_{kl}^{\rho}}{dist_{kl'}^{\rho}} = \frac{x_{ij'}}{x_{ij}} \frac{y_j/P_j^{1-\sigma}}{y_{j'}/P_{j'}^{1-\sigma}} \frac{\widetilde{dist}_{ij}^{\rho}}{\widetilde{dist}_{ij'}^{\rho}} \tag{12}$$

for all possible pairings as described above. This condition can be seen as a "neutrality result" in the sense that it would have to be met for aggregation not to affect estimated border dummy coefficients. But it is unlikely to hold in practice given that popular empirical distance measures such as distances between capital cities are not based on the theoretically consistent distance aggregator in Equation (10).

We can build intuition by looking at a special case. Imagine all micro regions were of symmetric size and had symmetric price indices. Then all weight parameters would be equal (i.e., $\theta_{k,i} = \theta_{l,j} = 1$ for all $k, l, i, j$). If in addition all micro-level distances aggregated for a particular flow were equal (i.e., $dist_{kl} = dist_{ij}$), then the distance aggregator would be the same as the micro-level distances (i.e., $\widetilde{dist}_{ij} = dist_{kl}$). The neutrality result (12) would hold, and aggregation would have no impact on the implied border coefficient. The intuition is that aggregation makes no difference as long as the aggregated micro regions are economic "clones" with the same size and the same trade costs over all micro-level flows that need to be aggregated to obtain the macro-level flow. However, as we show in the remainder of Section 3, the conditions for this special case are very restrictive and unrealistic, and they are inconsistent with the empirical evidence in Section 4.[11]

[10] This distance index is similar to the one in Head and Mayer (2009) but it accounts for multilateral resistance terms.

[11] It is important to note how restrictive this special case is. For example, for an aggregated internal trade flow $x_{ii}$, the special case would imply that the internal friction of a micro region, $t_{kk}$, is the same as the bilateral friction between micro regions, $t_{kk'}$, with $k, k' \epsilon i$. For more details, see our analytical results in the remainder of Section 3.

For further intuition, let us assume we initially measure the border effect in Equation (11) correctly by using the appropriate distance aggregator on the right-hand side. Then, all else being equal, suppose we mismeasure bilateral distance for the within-border flow by employing a measure $dist_{ij}$ that is shorter than suggested by the distance aggregator, that is, $dist_{ij} < \widetilde{dist}_{ij}$. Due to the negative distance elasticity ($\rho < 0$), the smaller distance measure would lead to an increase in the implied border friction $\zeta$. Given that $\zeta$ is typically negative (see Section 2.2), the inferred border friction would seem less severe (i.e., less negative). Intuitively, if the distance for trade within borders is relatively short, then we would expect more trade within and less trade across borders. To reconcile this view with the observed trade ratio $x_{ij'}/x_{ij}$, we would conclude that the implied border friction cannot be as burdensome. Thus, underestimating within-border distances will lead us to underestimate the border friction. Vice versa, underestimating cross-border distances will lead us to overestimate the border friction.

3.1.3. *The aggregate trade cost function.* To summarize the effect of trade cost mismeasurement on border effects, we can write the distance employed by the researcher as a function of the theoretically consistent distance aggregator and a residual:

$$dist_{ij} = \widetilde{dist}_{ij}\, e_{ij}^{-\frac{1}{\rho}},$$

where $e_{ij}$ captures the difference between the two distance measures. Substituting this expression into trade cost function (9), we obtain

$$t_{ij}^{1-\sigma} = \exp\left(\zeta BORDER_{ij}\right) dist_{ij}^{\rho}\, e_{ij} \tag{13}$$

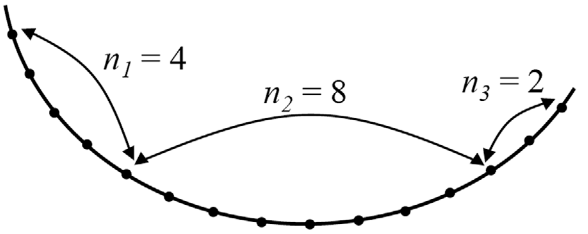
as the aggregate trade cost function. We stress that for purposes of interpretation, distance should be seen as representing all potential trade cost components apart from the border barrier itself. In our theoretical setting, we use distance as the only other trade cost component, and trade cost mismeasurement is therefore by construction focused on distance. But as the empirical gravity literature has demonstrated, in practice there are likely additional trade cost components.

Overall, if trade cost mismeasurement as captured by $e_{ij}$ is random, then using observed distances instead of the theoretical distance measure should not matter for border effect estimation. However, as we will show below based on a theoretical model with an analytical closed-form solution for aggregate trade costs, mismeasurement is typically not random. Instead, $e_{ij}$ is likely correlated with the border dummy, leading to systematically mismeasured border effects. In Section 4 , we test and confirm this result empirically.

3.2. *Spatial Aggregation and the International Border Effect.* We now develop a theory of spatial aggregation in the context of the international border effect. Our aim is to set up a model that is sufficiently simple to yield analytical results for the effect of spatial aggregation on estimated border effects.

We proceed in two steps. First, as in Section 3.1 we model trade flows at the level of small geographical units, which we call "micro" regions, based on a standard gravity setting. Second, we aggregate these micro regions into larger "macro" regions. Gravity also holds at the macro level, and we map the trade flows and trade costs of the micro regions onto the larger spatial units of macro regions.

3.2.1. *Micro and macro regions on a circle.* The model consists of two symmetric countries, Home and Foreign. Each country consists of symmetric micro regions denoted by superscript $S$ for "small." Micro regions get aggregated into macro regions denoted by the superscript $L$ for "large." As in Section 3.1, whenever necessary we use the subscripts $k$ and $l$ for micro




NOTES: Micro regions are symmetric arcs on the circle. Macro regions are aggregates of adjacent micro regions. Here, the first macro region consists of $n_1 = 4$ micro regions, the second consists of $n_2 = 8$ micro regions, and the third consists of $n_3 = 2$ micro regions.

FIGURE 1

ILLUSTRATION OF A SECTION OF THE CIRCLE REPRESENTING AN ECONOMY

regions and the subscripts $i$ and $j$ for macro regions. Each micro region is endowed with a differentiated good as in the Armington framework of Anderson and van Wincoop (2003) and has common income and multilateral resistance terms $y_k^S$ and $P_k^S$. Gravity equation (14) holds at the micro level:

$$x_{kl}^S = \frac{y^S y^S}{y^W} \left( \frac{t_{kl}^S}{P^S P^S} \right)^{1-\sigma}, \tag{14}$$

where we drop the subscripts for all region-specific variables. As outlined by Anderson and van Wincoop (2003), footnote 12, by setting outward and inward multilateral resistance price indices equal to each other, we adopt a particular normalization.[12]

We introduce a spatial topography such that frictions increase between more distant micro regions. For the purpose of obtaining analytical solutions, we adopt a symmetric setting. As the simplest case of such a topography, we model each country as a circle. Micro regions are symmetric arcs of the circle, each surrounded by two neighbors. Bilateral trade costs $t_h^S$ are equal to $\delta^h$ where $\delta \geq 1$ represents a spatial distance friction with $h \geq 1$ denoting the number of "steps" between micro regions. Adjacent regions are one step apart with $h = 1$, and so on. Thus, bilateral trade costs between micro regions increase in distance as long as $\delta > 1$. Internal trade costs within a micro region are lower than or equal to bilateral costs between micro regions, that is, $t_{kk}^S \leq t_h^S$ for any $h$.[13]

We aggregate $n \geq 1$ micro regions into a macro region denoted by superscript $L$ with income $y^L = ny^S$. Due to symmetry at the micro level, gravity also holds at the macro level. The aggregated micro regions are adjacent on the circle such that the macro region has no "holes." Figure 1 illustrates an example of macro regions on the circle.

**Aggregate bilateral trade costs**

Here we focus on bilateral trade between macro regions both within and across borders. Those are the relevant flows for the international border effect. But for completeness, in Appendix A.1 we also derive the internal trade flows of an aggregated macro region and the associated internal trade costs.

[12] In Appendix A.2, we show that our theory holds up under the more general treatment that retains separate outward and inward multilateral resistance variables. But to keep the exposition as simple as possible, we use the more parsimonious version here that adopts the normalization of equal outward and inward multilateral resistances.

[13] The theory for the domestic border effect in Appendix B can be seen as a one-country special case. The simple binary difference between bilateral and internal trade costs at the micro level can be achieved by setting $\delta = 1$ such that all bilateral trade costs become unity ($t_h^S = t^S = 1$) and by normalizing internal trade costs to a smaller value $t_{kk}^S < 1$.

Bilateral trade costs at the macro level are sensitive to aggregation. Suppose we observe two macro regions of different sizes, one comprising $n_1$ micro regions and the other $n_2$. Gravity commands the bilateral trade relationship:

$$x_{1,2}^L = \frac{y_1^L y_2^L}{y^W}\left(\frac{t_{1,2}^L}{P^L P^L}\right)^{1-\sigma}, \tag{15}$$

where $x_{1,2}^L$ denotes the trade flow from the first to the second macro region with bilateral costs $t_{1,2}^L$, and $y_1^L$ and $y_2^L$ are their respective incomes. More specifically, the bilateral macro flow from the first to the second macro region is the aggregate of $n_1 n_2$ bilateral micro flows:

$$x_{1,2,h}^L = \sum_{v=1}^{n_1}\sum_{w=1}^{n_2} x_{h+v+w-2}^S, \tag{16}$$

where the subscript $h$ in $x_{1,2,h}^L$ indicates the number of steps that the two macro regions are apart ($h$ is the smallest distance between a micro region from $i$ and a micro region from $j$). For instance, $x_{1,2,1}^L$ for $h = 1$ means that the two macro regions are adjacent (i.e., one step apart), and $x_{1,2,2}^L$ for $h = 2$ means the two macro regions are two steps apart. We have to add the micro flows $x_{h+v+w-2}^S$ with step length $h + v + w - 2$, summed over $v$ and $w$, to yield the bilateral macro flow.

Combining expressions (14) through (16), we can therefore write

$$x_{1,2,h}^L = \frac{n_1 y^S n_2 y^S}{y^W}\left(\frac{t_{1,2,h}^L}{P^L P^L}\right)^{1-\sigma} = \frac{y^S y^S}{y^W}\frac{\sum_{v=1}^{n_1}\sum_{w=1}^{n_2}\left(t_{h+v+w-2}^S\right)^{1-\sigma}}{\left(P^S P^S\right)^{1-\sigma}},$$

where $t_{1,2,h}^L$ denotes aggregate bilateral trade costs for the two macro regions $h$ steps apart. It turns out that aggregation does not change the multilateral resistance price indices, that is, $P^S = P^L$. In Appendix A.2, we show this result formally. The intuition is that aggregation does not affect the underlying trade cost structure and equilibrium of trade flows. Thus, the expression for bilateral trade costs at the macro level follows from the previous equation as

$$\left(t_{1,2,h}^L\right)^{1-\sigma} = \frac{1}{n_1 n_2}\sum_{v=1}^{n_1}\sum_{w=1}^{n_2}\left(t_{h+v+w-2}^S\right)^{1-\sigma}. \tag{17}$$

A key result is that these bilateral macro trade costs rise in the number of aggregated micro regions, that is, $\partial t_{1,2,h}^L/\partial n_1 > 0$ and $\partial t_{1,2,h}^L/\partial n_2 > 0$. That is, all else equal, larger macro regions tend to have larger trade costs with other regions in that country. The only exception would be the special case of no spatial gradient when bilateral trade costs between micro regions are the same regardless of distance, that is, when $\delta = 1$ such that $t_h^S = t^S$ for all $h$. In that case, bilateral trade costs would be the same at the micro and macro levels.

To see more clearly how bilateral trade costs depend on region size $n_1$ and $n_2$, we substitute the spatial friction $t^S_{h+v+w-2} = \delta^{h+v+w-2}$.[14] We can then decompose bilateral trade costs at the macro level into three elements as

$$t^L_{1,2,h} = \underbrace{\delta^h}_{\text{bilateral distance}} \underbrace{\left(\frac{1}{n_1}\sum_{v=1}^{n_1}\left(\delta^{v-1}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}_{\alpha_1} \underbrace{\left(\frac{1}{n_2}\sum_{w=1}^{n_2}\left(\delta^{w-1}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}_{\alpha_2}. \tag{18}$$

The first element $\delta^h$ denotes the bilateral distance between the two macro regions. The remaining elements $\alpha_1$ and $\alpha_2$ are region-specific, and more importantly they rise in the sizes $n_1$ and $n_2$ of the macro regions.[15] These terms can be interpreted as the costs of reaching the domestic borders of macro regions. For instance, suppose the first macro region consists of only one micro region ($n_1 = 1$). It follows $\alpha_1 = 1$, meaning that no distance has to be incurred to reach the domestic border. But for a macro region consisting of several micro regions ($n_1 > 1$), we get $\alpha_1 > 1$ as long as $\delta > 1$ because of the rising average internal distances of individual micro regions to the domestic border.

In summary, bilateral trade costs at the macro level increase in the size of the underlying regions because more spatial frictions within the macro regions have to be overcome. Only in the limiting case where the macro regions are micro regions ($n_1 = n_2 = 1$) does the bilateral distance $\delta^h$ fully represent the bilateral trade costs.

**International trade costs**

Both countries have the same internal structure of micro regions, and we therefore have two circles. We assume that bilateral international trade costs between micro regions $t^S_{int}$ consist of a common international distance $\delta_{int}$. The common distance can be motivated by a central port for international trade in each country. Then for each micro region the distance to the port is the same.[16] In addition, we assume a cost for crossing the international border so that we can write

$$t^S_{int} = \delta_{int} \exp\left(\frac{\beta}{1-\sigma}\right), \tag{19}$$

where $\beta \leq 0$ captures the international border barrier. This structure translates into the same level of international trade costs at the aggregate level between two macro regions of size $n_1$ and $n_2$, that is, $t^S_{int} = t^L_{1,2,int}$. The intuition is that identical trade costs are aggregated such that the appropriate theoretical average is the same. This stands in contrast to aggregate bilateral trade costs within countries as in Equation (18) that do vary by macro region size.

We should briefly comment on a possible generalization. As an alternative modeling strategy, instead of just two circles representing two countries we could assume multiple circles representing multiple countries. To preserve symmetry we could have a "pearl necklace" of countries where each pearl represents a circular economy. That is, we could arrange countries in a circular fashion similar to the way micro regions are arranged within countries. International distances would then vary by country pair in contrast to our simple common distance $\delta_{int}$. However, this expanded model would not yield any qualitatively new insights. We therefore work with the simpler two-country setting.

[14] We assume that the two macro regions are in the same semi-circle so that the shortest direction of trade is always either clockwise or counterclockwise. If the two regions straddled different semi-circles, the resulting expression for $t^L_{1,2,h}$ would be more complicated.

[15] Formally, $\partial\alpha_1/\partial n_1 > 0$ and $\partial\alpha_2/\partial n_2 > 0$.

[16] As a generalization, we could allow for bilateral distance gradients between micro regions at the international level. The relevant case would be a friction parameter that differs from the corresponding parameter $\delta$ for domestic flows.

**The trade cost function**

Comparing expressions (18) and (19) for bilateral trade costs at the domestic and international levels, we can see that region-specific terms only appear for domestic trade costs. In logarithmic form and scaled by the elasticity of substitution, we can therefore write the overall trade cost function that arises from our model as

$$\ln\left(t_{ij}^{1-\sigma}\right) = \ln\left(\delta_{ij}^{1-\sigma}\right) + \beta INT_{ij} + \phi(1 - INT_{ij})\left\{\ln(\alpha_i^{1-\sigma}) + \ln(\alpha_j^{1-\sigma})\right\} \tag{20}$$

with $\phi = 1$ where region $i$ denotes an exporter and region $j$ is an importer. If $ij$ is a domestic pair, then $\delta_{ij}$ equals $\delta^h$, and $\delta_{int}$ otherwise. "Regions" here refer to macro regions (those will be U.S. states and foreign countries in our empirical analysis), and for simplicity we drop the $L$ superscript. The trade cost function (20) is a special case of the general trade cost function (13) developed in Section 3.1.[17]

3.2.2. *Heterogeneous international border effects.* The key feature of trade cost function (20) is the interaction term between the international border dummy and the region-specific terms $\ln(\alpha_i^{1-\sigma})$ and $\ln(\alpha_j^{1-\sigma})$. This interaction is absent in standard trade cost functions such as (2). It implies that in gravity estimation, the impact of the border on bilateral trade becomes heterogeneous. More specifically, the direct effect of $INT_{ij}$ on bilateral trade follows as

$$\frac{\Delta \ln(x_{ij})}{\Delta INT_{ij}} = \beta + \phi\left\{\ln(\alpha_i^{\sigma-1}) + \ln(\alpha_j^{\sigma-1})\right\}, \tag{21}$$

where $\Delta INT_{ij}$ indicates a comparison of $INT_{ij} = 0$ with $INT_{ij} = 1$ and where for the moment we ignore general equilibrium effects including multilateral resistance effects operating through the price indices.

If an international border barrier exists, we have $\beta < 0$. In the limiting case when regions $i$ and $j$ are micro regions with no aggregated spatial frictions, we have $\alpha_i = \alpha_j = 1$ and the second term disappears. This would also happen if the domestic economies were frictionless in the sense of $\delta = 1$. In Appendix A.3, we show formally that only if the $\alpha_i$ and $\alpha_j$ terms are unity can we obtain an unbiased estimate of $\beta$ in a gravity regression with a standard international border effect. But in the more realistic case when $i$ and $j$ are macro regions and spatial frictions are present, the second term becomes positive and counteracts the negative effect stemming from $\beta$. Thus, larger macro regions have weaker (i.e., less negative) border effects. We call this the spatial attenuation effect. The effect of an international border dummy is therefore driven by the "internal resistance" of the regions in question, inducing systematic heterogeneity. In the empirical part of the article, we illustrate the heterogeneity by reporting the full range of border effects. We find that the heterogeneity is quantitatively substantial.

We note that this form of heterogeneity operates independently of heterogeneity induced by multilateral resistance effects in general equilibrium. We discuss general equilibrium effects in more detail in Section 4.7.

**Estimating heterogeneous international border effects**

Estimation of trade cost function (20) is straightforward. The $\alpha_i$ and $\alpha_j$ terms are region-specific. We can therefore capture them with a set of region fixed effects $\alpha_r$ that equal unity

[17] The $\zeta BORDER_{ij}$ term in Equation (13) corresponds to $\beta INT_{ij}$ in Equation (20), and $dist_{ij}^{\rho}$ corresponds to $\delta_{ij}^{1-\sigma}$. The trade cost mismeasurement term is equal to $e_{ij} = 1$ for $INT_{ij} = 1$ and equal to $e_{ij} = \alpha_i^{1-\sigma}\alpha_j^{1-\sigma}$ for $INT_{ij} = 0$ with $\phi = 1$, meaning trade costs are correctly measured for $INT_{ij} = 1$ but mismeasured for $INT_{ij} = 0$. This yields trade cost function (20).

whenever $i$ or $j$ appear regardless of the direction of trade.[18] As the empirical specification, we obtain

$$(22)\quad \ln\left(t_{ij}^{1-\sigma}\right) = \underbrace{INT_{ij}\Big(\beta + \phi\Big\{\ln(\alpha_i^{\sigma-1}) + \ln(\alpha_j^{\sigma-1})\Big\}\Big)}_{\beta_r INT_{ij}\alpha_r \;=\; \beta_r INT_{ij}^r} + \ln\left(\delta_{ij}^{1-\sigma}\right) - \underbrace{\phi\Big\{\ln(\alpha_i^{\sigma-1}) + \ln(\alpha_j^{\sigma-1})\Big\}}_{\alpha_r},$$

where $\beta_r$ indicates region-specific international border coefficients and where $INT_{ij}^r \equiv INT_{ij}\alpha_r$. A simple test of border effect heterogeneity comes down to the hypothesis that the $\beta_r$ coefficients differ from each other. We note that the $\beta$ parameter cannot be identified due to collinearity with the fixed effects.

3.2.3. *Testable implications.* The previous analysis leads to three testable implications that we will explore in the empirical section:

(1) **Heterogeneous border effects**: According to our theoretical model, we should expect heterogeneous border effects as opposed to a common border effect. Specifically, all else being equal, through Equations (18) and (21) we should expect weaker border effects for larger regions. We can employ trade cost function (22) with region-specific border effects to test this prediction empirically.
(2) **Systematic sample composition effects**: Related to the first implication, border effect coefficients are sensitive to sample composition in a systematic way. Specifically, adding relatively large regions to the sample pushes the border coefficient toward zero. Conversely, adding relatively small regions renders the border coefficient more negative. Such composition effects imply that standard border effect coefficients estimated in the conventional way without allowing for heterogeneity are not directly comparable across different samples. For example, suppose we obtain a coefficient of $\beta_1 = -1.5$ in one sample and a coefficient of $\beta_2 = -1$ in another, and the two coefficients are significantly different. This difference does not necessarily imply that the international border is more detrimental to trade flows in the first sample than in the second. Instead, the difference could be driven by sample composition effects.
(3) **Aggregation effects**: Aggregating adjacent regions generates larger regions associated with weaker border effects. As a consequence of the second implication, we should therefore expect weaker conventional border effect coefficients in a sample with aggregated regions compared to the same sample prior to aggregation.

Before turning to the empirical analysis, we briefly review corresponding theoretical results for the domestic border effect.

3.3. *Spatial Aggregation and the Domestic Border Effect.* In Appendix B, we develop a model of spatial aggregation that is targeted toward the domestic border effect. As in the model for the international border effect in Section 3.2, we work with micro regions that get aggregated into macro regions. But as a simplification, we do not model regions as being located on a circle. Similar to Equation (21), we show that the direct effect of the domestic border dummy $DOM_{ij}$ on trade is heterogeneous:

$$(23)\quad \frac{\Delta \ln(x_{ij})}{\Delta DOM_{ij}} = \gamma + \ln\left(t_{ii}^{\sigma-1} t_{jj}^{\sigma-1}\right)^{\frac{1}{2}},$$

where $\Delta DOM_{ij}$ indicates a comparison of $DOM_{ij} = 0$ with $DOM_{ij} = 1$.[19]

[18] We do not use internal trade flows in the estimation where $i = j$.
[19] See Equation (44) in Appendix B.1.

The key insight is that all else equal, larger internal trade costs lead to a smaller border effect. That is, the term $\ln(t_{ii}^{\sigma-1}t_{jj}^{\sigma-1})^{1/2}$ increases in $t_{ii}$ and $t_{jj}$ and thus counteracts the negative effect stemming from $\gamma < 0$. Ceteris paribus domestic border effects are therefore mechanically driven by internal trade costs and inherently heterogeneous. This is spatial attenuation in the context of the domestic border effect. The intuition is that due to aggregation, larger macro regions have larger internal trade frictions. This increases "internal resistance," leading to relatively less internal trade and relatively more bilateral trade. As a result, the domestic border effect appears smaller.

Similar to Equation (22), we show that the model implies a trade cost function with region-specific domestic border effect coefficients $\gamma_r$.[20] A simple test of border effect heterogeneity comes down to the hypothesis that the $\gamma_r$ coefficients differ from each other. Similar testable implications follow as in Section 3.2.3.[21] We will explore these in the subsequent section.

## 4. EMPIRICAL RESULTS

4.1. *Data.* Our two main data sources are the Commodity Flow Survey and the Origin of Movement series provided by the U.S. Census Bureau. To obtain results that are comparable to the literature, we use the same data sets as Wolf (2000) and Anderson and van Wincoop (2003) for domestic trade flows within the United States, based on the Commodity Flow Survey. 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, based on the Origin of Movement series. Thus, for instance, our data set comprises trade flows within Minnesota, exports from Minnesota to Texas as well as exports from Minnesota to France. We also employ trade data between foreign countries in our sample. We take data quality seriously, and in Appendix C we describe our data sources and adjustments in detail, including our distance measures.

We form a balanced sample over the years 1993, 1997, 2002, and 2007. We drop Alaska, Hawaii, and Washington, D.C. due to data quality concerns raised in the Commodity Flow Survey so that we are left with the 48 contiguous states. This yields 1,726 trade observations per cross-section within the United States, including 48 intrastate observations and 1,678 state-to-state observations per cross-section.[22] The observations that involve the 50 foreign countries are made up of 2,338 export flows from U.S. states to foreign countries as well as 2,233 exports flows among foreign countries per cross-section.[23]

4.2. *Overview.* We first show in Section 4.3 that our data exhibit a significant international border effect, as established by McCallum (1995). We also show that the data exhibit a substantial domestic border effect, as established by Wolf (2000).

We then explore the three testable implications from Section 3.2.3. In Section 4.4, we explore the first implication (heterogeneous border effects) by allowing border effects to vary across states, uncovering a great degree of heterogeneity. Section 4.5 examines the second implication (sample composition effects) where we systematically drop large and small U.S. states from our sample. Section 4.6 covers the third implication (aggregation effects) where we aggregate U.S. states into larger spatial units.

[20] See Equation (45) in Appendix B.1.

[21] See Appendix B.1 for details.

[22] The maximum possible number of U.S. observations would be $48\times48 = 2,304$ per cross-section. The missing observations are due to the fact that a number of Commodity Flow Survey estimates did not meet publication standards because of high sampling variability or poor response quality. To generate a balanced sample, we drop pairs if at least one year is missing.

[23] The maximum possible number of international exports from U.S. states would be $48\times50 = 2,400$ per year. We have 62 missing observations mainly because exports to Malaysia were generally not reported in 1993. Only 18 of these observations not included in our sample are most likely zeros (as opposed to missing). The maximum possible number of exports between foreign countries would be $49\times50 = 2,450$ per cross-section. To generate a balanced sample, we drop pairs if at least one year is missing.

TABLE 1
INTERNATIONAL AND DOMESTIC BORDER EFFECTS

| Sample | U.S. and Foreign Countries | | U.S. only | |
|---|---|---|---|---|
| Year | 1993 | 1993, 1997, 2002, 2007 | 1993 | 1993, 1997, 2002, 2007 |
| | (1) | (2) | (3) | (4) |
| $\ln(dist_{ij})$ | $-1.19^{***}$ | $-1.21^{***}$ | $-1.07^{***}$ | $-1.08^{***}$ |
| | (0.02) | (0.02) | (0.03) | (0.03) |
| $INT_{ij}$ (international border dummy) | $-1.25^{***}$ | $-1.21^{***}$ | | |
| | (0.08) | (0.06) | | |
| $DOM_{ij}$ (domestic border dummy) | | | $-1.47^{***}$ | $-1.48^{***}$ |
| | | | (0.20) | (0.19) |
| Internal trade (within U.S. states) | no | no | yes | yes |
| Domestic trade (between U.S. states) | yes | yes | yes | yes |
| International trade (with foreign countries) | yes | yes | no | no |
| Observations | 6249 | 24,996 | 1726 | 6904 |
| Clusters | – | 6249 | – | 1726 |
| Fixed effects | yes | yes | yes | yes |
| $R^2$ | 0.81 | 0.82 | 0.90 | 0.90 |

*Notes*: The dependent variable is ln(xij). OLS estimation. Robust standard errors are reported in parentheses, clustered around bilateral pairs ij in columns 2 and 4. State and foreign country fixed effects in columns 1 and 2; exporter and importer fixed effects in columns 3 and 4; those fixed effects are time-varying in columns 2 and 4. *** denotes statistical significance at the 1% level.

In Section 4.7, we show that border effect heterogeneity is substantially more important than heterogeneity related to general equilibrium price index effects. Section 4.8 explores potential determinants of border effects, namely transport infrastructure. Finally, in Section 4.9 we present additional evidence of border effect heterogeneity based on industry-level trade data. In Appendix C.7, we provide numerical simulations of our model.

4.3. *Estimating Common Border Effects.* In Table 1, we replicate well-known results on international and domestic border effects. As our estimating equation, we use the log-linear version of gravity equation (1) in combination with trade cost function (2).

In columns 1 and 2 of Table 1, we replicate standard results for the international border effect. As is customary, we do not include trade flows within U.S. states, and the domestic border dummy is dropped as a regressor. To be able to identify the international border dummy coefficient, we follow Anderson and van Wincoop (2003) and others by using state and foreign country fixed effects instead of exporter and importer fixed effects.[24] As the output regressors are collinear with these fixed effects, they are dropped from the estimation. In column 1, we estimate an international border coefficient of $\widehat{\beta} = -1.25$ for the year 1993, implying that after we control for distance and economic size, exports from U.S. states to foreign countries are about 71% lower than trade between U.S. states ($\exp(-1.25) = 0.29$). Assuming a value for the elasticity of substitution of $\sigma = 5$, we can translate this into a tariff equivalent of the domestic border of 37%.[25] In column 2, when we pool the data over the years 1993, 1997, 2002 and 2007, we obtain a similar coefficient of $-1.21$. These estimates are somewhat smaller in absolute magnitude but nevertheless roughly fall in the same ballpark as the estimates of around $-1.6$ reported by Anderson and van Wincoop (2003), Table 2, in their sample involving trade flows of U.S. states and Canadian provinces.

[24] That is, we use fixed effects that are state-specific (in the case of U.S. states) and country-specific (in the case of foreign countries) but that do not distinguish between being an exporter or an importer.

[25] For $\ln(t_{ij}^{1-\sigma}) = -1.25$, it follows $t_{ij} = 1.37$. This is a partial equilibrium calculation in the sense that we ignore price index effects for simplicity. For general equilibrium effects, see Section 4.7.

In columns 3 and 4 of Table 1, we estimate the domestic border dummy coefficient. We only use trade flows within the U.S. International trade flows are not included, and the international border dummy is dropped as a regressor. As typical in the literature, for instance, Hillberry and Hummels (2003), we use exporter and importer fixed effects to control for multilateral resistance and all other country-specific variables such as income. As in Wolf (2000), in column 3 we only use data for 1993. In column 4, we add the data for 1997, 2002, and 2007. Our estimate of $\widehat{\gamma} = -1.48$ in column 4 is the same as Wolf's baseline coefficient.[26]

The interpretation of our coefficient is that given distance and economic size, trade between U.S. states is 77% lower compared to trade within U.S. states ($\exp(-1.48) = 0.23$). The corresponding tariff equivalent is 45%. But as we show in Appendix B.1, this interpretation would only be valid under the assumption that internal trade costs within U.S. states were zero on average. For positive internal trade costs (which is the realistic scenario), the underlying tariff equivalent would be even higher.[27] Put differently, the $\widehat{\gamma}$ estimate only captures the domestic border barrier net of internal trade costs.

Overall, we have replicated international and domestic border coefficient estimates as typically found in the literature. In fact, our domestic point estimate exceeds the international point estimate in absolute magnitude, a finding that is consistent with Fally et al. (2010) in their study of Brazilian trade data as well as Coughlin and Novy (2013).[28]

4.4. *Estimating Individual Border Effects.* The general trade cost function (13) highlights how aggregation can affect the relevant distance measure. In principle, aggregation can lead to mismeasurement of trade costs within borders and across borders (where cross-border trade costs refer to international trade costs in the context of the international border effect, and to bilateral domestic trade costs in the context of the domestic border effect). If aggregation increases within-border trade costs relatively more, then the cost of trading across borders appears relatively smaller and we would infer weaker border effects for larger states. This is the scenario predicted by our theoretical models for the international border effect in Section 3.2 and the domestic border effect in Appendix B. The alternative hypotheses that would in principle be consistent with the general specification (13 ) are either no aggregation effects, or stronger border effects for larger states. We now examine the data to find out which pattern holds empirically.

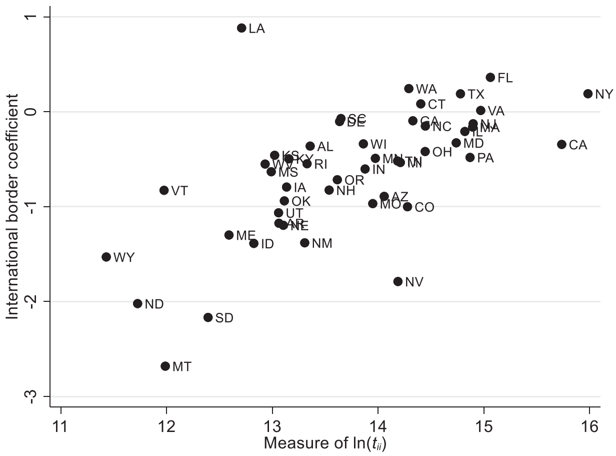
We run the same regression specifications with panel data as in columns 2 and 4 of Table 1, but now allowing the international and domestic border coefficients to vary across states. That is, we estimate individual, state-specific border effects to explore the first testable implication in Section 3.2.3. This approach is consistent with our theoretical results in expressions (21) and (23) where we predict that for larger states, the international and domestic border dummy coefficients should be closer to zero due to spatial attenuation.

4.4.1. *Individual International Border Effects.* We first estimate individual international border dummy coefficients for the 48 U.S. states in our sample. We obtain them by using trade

[26] Wolf's coefficient has a positive sign because his domestic border dummy is coded in the opposite way. Hillberry and Hummels (2003) reduce the magnitude of the national border 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 intra-state trade. However, Nitsch (2000) reports higher coefficients in the range of −1.8 to −2.9 by comparing trade within European Union countries to trade between EU countries.

[27] See expression (42) in Appendix B.1.

[28] Hillberry and Hummels (2008) use ZIP code-level shipments within the United States and show that over very short distances, distance operates in a highly nonlinear fashion due to extensive margin effects. The smallest unit in our sample (a U.S. state) has an average internal distance of 179 kilometers as opposed to around 6 kilometers for a ZIP code. We do not use any internal trade observations for our results on the international border effect.




NOTES: The mean of the coefficients is −0.64. The average standard error is 0.13 (not plotted). More details are provided in the main text.

FIGURE 2

PLOT OF INTERNATIONAL BORDER DUMMY COEFFICIENTS FOR THE 48 CONTIGUOUS U.S. STATES AGAINST THE LOGARITHM OF A MODEL-CONSISTENT PROXY FOR INTERNAL TRADE COSTS $t_{ii}$

cost function (22) in otherwise standard gravity estimation, substituting bilateral distance for $\delta_{ij}$.[29] More precisely, we estimate the following specification

$$\ln(x_{ij}) = \sum_{r=1}^{48} \beta_r INT_{ij}^r + \rho \ln(dist_{ij}) + \alpha_s + \varepsilon_{ij},$$

where $INT_{ij}^r$ is an international border dummy specific to state $r$, $\beta_r$ is the corresponding coefficient, $\alpha_s$ denotes state and foreign country fixed effects, and $\varepsilon_{ij}$ is an error term.[30] As Equation (21) shows, all else equal theory predicts a smaller trade effect of the international border dummy in absolute value (i.e., less negative) for regions with larger (logarithmic) internal trade costs, which are in turn associated with larger economic size.

How can we obtain a measure of internal trade costs that is consistent with the theory? Equation (26) in Appendix A.1 describes how $t_{ii}^L$ depends on the number of aggregated micro regions $n$ and the micro frictions $t_{kk}^S$ and $t^S$. But as these micro frictions are unobservable, instead we resort to gravity equation (24) to obtain a theory-consistent measure of internal trade costs.[31] Given that multilateral resistance terms are the same across macro regions, it follows that $(t_{ii}^L)^{\sigma-1}$ is proportional to the ratio $y^L y^L / x_{ii}^L$. We therefore proxy $\ln(t_{ii})$ with $\ln(y^L y^L / x_{ii}^L)$.

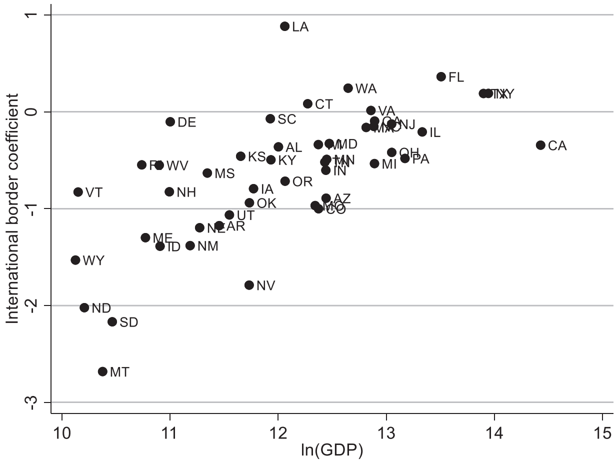
Figure 2 illustrates the individual coefficients plotted against our proxy of internal trade costs. As a more direct measure of economic size, Figure 3 plots the coefficients against logarithmic state GDP. Overall, the figures demonstrate a clear positive relationship. The individual estimates display a large degree of heterogeneity, falling into a range of −2.7 to 0.9. The mean estimate is −0.64.[32] The coefficients are tightly estimated with an average standard

[29] Behrens et al. (2012) also estimate heterogeneous international border dummy coefficients based on a framework that allows for spatial correlation of trade flows.

[30] As in the regressions underlying columns 1 and 2 of Table 1, we follow Anderson and van Wincoop (2003) and others by using state and foreign country fixed effects to identify the international border dummy coefficients. That is, the $\alpha_s$ fixed effects are state-specific (in the case of U.S. states) and country-specific (in the case of foreign countries).

[31] We refer to Appendix C.7 where we discuss internal trade costs in more detail.

[32] The corresponding common international border coefficient that captures international trade flows of U.S. states only is −0.60 and thus very close to the mean estimate underlying Figures 2 and 3. See Section 4.7 for details.




NOTES: The mean of the coefficients is −0.64. The average standard error is 0.13 (not plotted). More details are provided in the main text.

FIGURE 3

PLOT OF INTERNATIONAL BORDER DUMMY COEFFICIENTS FOR THE 48 CONTIGUOUS U.S. STATES AGAINST THE LOGARITHM OF STATE GDP

error of 0.13 (not plotted in the figure). The larger the state, the closer the individual international border coefficient tends to be to zero. For example, Wyoming as the smallest state is associated with an international border coefficient of −1.53, whereas the value for California as the largest state is −0.34. Under the assumption of $\sigma = 5$, the corresponding tariff equivalents would be 47% and 9%.

We stress that in our model, the international border friction at the micro level, $\beta$, is common across all regions (see Equation 19). Through that lens the substantial difference between the above tariff equivalents can therefore be attributed to spatial aggregation as the primary driving force behind border effect estimates (see Section 5 for a discussion).

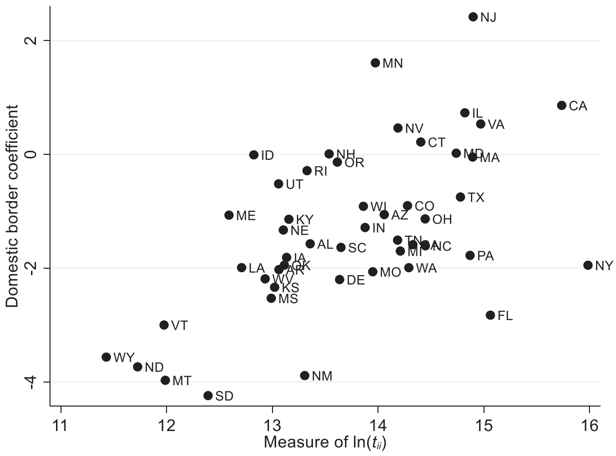
4.4.2. *Individual Domestic Border Effects.* We also estimate individual domestic border dummy coefficients. We obtain them by using the theoretically consistent trade cost function (45) from Appendix B.1 in an otherwise standard gravity regression. That is, we estimate the following specification:

$$\ln(x_{ij}) = \sum_{r=1}^{48} \frac{\gamma_r}{2} DOM_{ij}^r + \rho \ln(dist_{ij}) + \alpha_i + \alpha_j + \varepsilon_{ij},$$

where $DOM_{ij}^r$ is a domestic border dummy specific to state $r$, $\gamma_r$ is the corresponding coefficient, $\alpha_i$ and $\alpha_j$ denote exporter and importer fixed effects, and $\varepsilon_{ij}$ is an error term.[33] As Equation (44) shows, theory predicts that for a given U.S. state, all else being equal we should expect a smaller trade effect of the domestic border dummy in absolute magnitude (i.e., less negative) if the state has larger (logarithmic) internal trade costs.

As an illustration, in Figure 4 we plot the domestic border coefficients $\gamma_r$ against our proxy of internal trade costs. Two main observations can be made. First, as with the international border coefficients there is a large degree of heterogeneity. Although the mean of coefficients is −1.32 and thus close to the point estimates reported in columns 3 and 4 of Table 1, the

[33] See Appendix B.1 for details. As every domestic trade flow is captured by two region-specific border dummies (once on the exporter side and once on the importer side), the estimated coefficients $\gamma_r/2$ must be multiplied by 2 to obtain estimates of $\gamma_r$ that are comparable to the common border dummy coefficient $\gamma$.




NOTES: The mean of the coefficients is -1.32. The average standard error is 0.13 (not plotted). More details are provided in the main text.

FIGURE 4

PLOT OF DOMESTIC BORDER DUMMY COEFFICIENTS FOR THE 48 CONTIGUOUS U.S. STATES AGAINST THE LOGARITHM OF A MODEL-CONSISTENT PROXY FOR INTERNAL TRADE COSTS $t_{ii}$

individual border coefficients span a range of more than six log points. They are tightly estimated, with standard errors of 0.13 on average.[34]

Second, as predicted by our theory, the individual coefficients are positively related to internal trade costs. Given a correlation of 0.92 between internal trade costs and state GDP, this means the coefficients are also positively related to the economic size of states.[35] That is, the smaller the state, the more detrimental the effect of crossing a domestic border appears to be. For example, the five states with the smallest state GDPs (Wyoming, Vermont, North Dakota, Montana, and South Dakota) have border coefficients in the vicinity of −4. The back-of-the-envelope interpretation would be that for those states, crossing a border with another state reduces trade by 98%.[36] At the other extreme, a few economically large states such as New Jersey and California are associated with positive border coefficients.[37] These results underline the importance of spatial attenuation effects.[38]

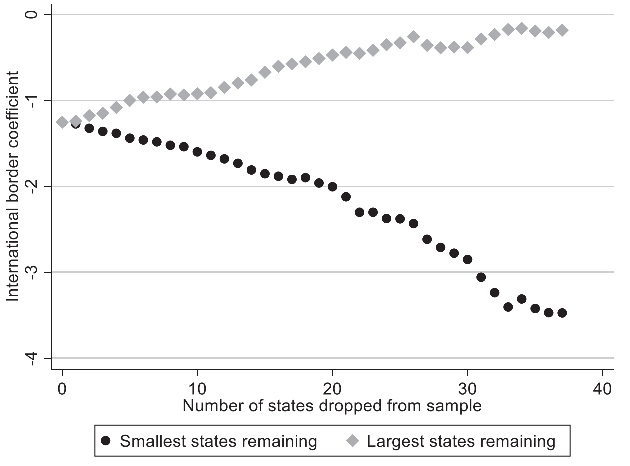
[34] Given that our sample has a number of missing observations as explained in Section 4.1, we carry out a robustness check where we drop the states with the most missing observations (those with fewer than 60 bilateral observations out of a maximum of $94 = 2 \times 47$ per cross-section). Those states are Delaware, Idaho, Montana, North Dakota, New Hampshire, New Mexico, Nevada, Rhode Island, South Dakota, Vermont, and Wyoming. We re-estimate individual domestic border dummy coefficients based on this reduced sample. As average state size is larger, the mean of coefficients shrinks in absolute magnitude (it is −1.21 instead of −1.32). The individual coefficients in the reduced sample are very similar to those in Figure 4. They span a range of five log points and are tightly estimated with an average standard error of 0.10. Their positive correlation with logarithmic internal trade costs is retained (the correlation is 45%). We conclude that the heterogeneity pattern in Figure 4 is not driven by missing observations.

[35] We also refer to Appendix C.7 where we provide numerical simulations confirming the positive relationship between individual coefficients and internal costs.

[36] As $\exp(-4) = 0.02$, all else equal in partial equilibrium the border reduces trade by 98% relative to within-state trade. For general equilibrium effects, see Section 4.7.

[37] The coefficients for California, Illinois, Minnesota, Nevada, New Jersey, and Virginia are positive and significant at the 5% level. In the theory in Appendix B.1, the upper bound for state-specific domestic border coefficients is actually zero. In Equation (37), $t_{ii}^L$ approaches $t^S$ for $n \to \infty$, which is the same as $t_{1,2}^L$ through Equation (39). Therefore, in Equation (45), it follows $\gamma_r = 0$ as $\psi = 1$. In the data, however, it is conceivable that trade costs within some macro regions are sufficiently large relative to bilateral trade costs such that positive domestic border coefficients are estimated (see Equation 44).

[38] As described in Appendix C.4, our measure of intrastate distance is the distance between the two largest cities in a state. We also employed two alternative intrastate distance measures. The first is the distance measure proposed by Wolf (2000) that weights the distance between a state's two largest cities by their population. The second is the mea-




NOTES: When zero states are dropped, the coefficient is -1.25 as in column 1 of Table 1. Black dots plot the coefficients obtained by successively dropping the largest remaining state from the sample such that the smallest states are remaining. The grey diamonds plot the coefficients obtained by successively dropping the smallest remaining state such that the largest states are remaining. More details are provided in the main text.

FIGURE 5

PLOT OF COMMON INTERNATIONAL BORDER DUMMY COEFFICIENTS ESTIMATED FOR DIFFERENT SAMPLES OF U.S. STATES

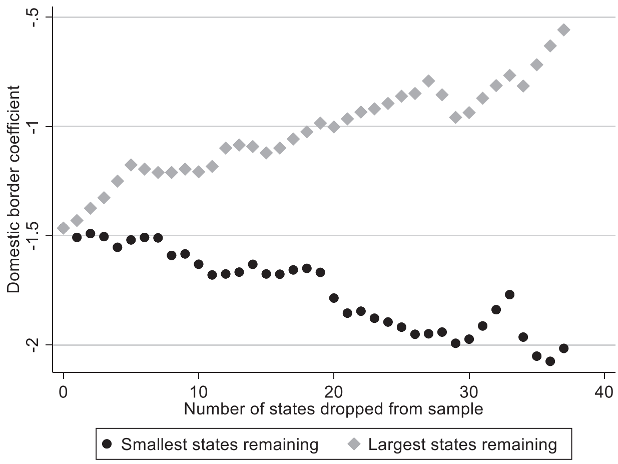
4.5. *Sample Composition Effects.* As shown above, border dummy coefficients can vary substantially across regions. They tend to be large in absolute magnitude for small states, and vice versa. It follows that when we estimate common border effects, our estimates should be sensitive to the distribution of state economic size in the sample (see the second testable implication in Section 3.2.3). We perform a simple check of this sample composition effect.

To systematically change the composition of economic size in our sample, we run rolling regressions where we keep dropping states and their associated trade flows from the sample. More specifically, we begin with the international border effect regression as in column 1 of Table 1 for the year 1993 where we obtained a coefficient on the domestic border dummy of −1.25. We then drop the largest state from the sample in terms of GDP (California) and re-estimate the border coefficient. We then drop the second largest state from the sample (New York) and re-estimate, and so on, such that the smallest states are remaining. To obtain comparable estimates, we keep the distance coefficient at its initial value but we allow the exporter and importer fixed effects to adjust freely.

The black dots in Figure 5 illustrate the international border coefficients. As predicted by our theory, we yield the following clear pattern: the more big states we drop from the sample, the larger the coefficients tend to become in absolute value. That is, the smaller the average economic size of states in the sample, the further the estimated border effect tends to get pushed away from zero. The grey diamonds in Figure 5 illustrate the coefficients obtained when we drop the smallest state first (Wyoming), then the second smallest state (Vermont), and so on. As expected, we yield the opposite pattern: the international border coefficients move upward toward zero. Overall in Figure 5, we obtain coefficients ranging from around −3.5 to 0.[39]

sure suggested by Nitsch (2000) that is based on land area (i.e., 0.56 times the square root of a state's land area). Although alternative distance measures can alter the common border effect coefficient estimate, we still obtain similar patterns of coefficient heterogeneity as in Figure 4.

[39] Balistreri and Hillberry (2007) show that the reduction of the border effect by Anderson and van Wincoop (2003) relies on the addition of trade flows between U.S. states to the sample. As U.S. states are on average considerably larger than Canadian provinces, we expect the addition of such flows to push the common border dummy estimate toward zero according to our result in Figure 5.




NOTES: When zero states are dropped, the coefficient is −1.47 as in column 3 of Table 1. Black dots plot the coefficients obtained by successively dropping the largest remaining state from the sample such that the smallest states are remaining. The grey diamonds plot the coefficients obtained by successively dropping the smallest remaining state such that the largest states are remaining. More details are provided in the main text.

FIGURE 6

PLOT OF COMMON DOMESTIC BORDER DUMMY COEFFICIENTS ESTIMATED FOR DIFFERENT SAMPLES OF U.S. STATES

In Figure 6, we repeat the rolling regressions for the domestic border effect, starting out with the same regression as in column 3 of Table 1 where we obtained a coefficient of −1.47. We find a similar pattern as in Figure 5. That is, the smaller the average economic size of states in the sample, the further the estimated border effect tends to get pushed away from zero (the downward trend is reasonably clear although not strictly monotonic). The coefficients roughly fall in the range from −2 to −0.5.

Therefore, in summary we find strong sample composition effects in Figures 5 and 6. We interpret these as further evidence corroborating the impact of state size on border effects. The figures demonstrate that this impact is quantitatively strong.

4.6. *Aggregating U.S. States.* The individual border effects illustrated in Figures 2–4 demonstrate that larger states tend to exhibit smaller border effects in absolute magnitude. We now trace this relationship between economic size and the magnitude of border effects in a different way. We aggregate U.S. states and thus enlarge the size of the underlying spatial units (see the third testable implication in Section 3.2.3).

To be specific, we aggregate the 48 contiguous U.S. states into the nine Census divisions as defined by the U.S. Census Bureau. We choose Census divisions because their borders conveniently coincide with state borders (this would not be the case with Federal Reserve Districts, for instance). But any alternative clustering of adjacent states would in principle be equally suitable for this aggregation exercise. Figure 7 provides a map of the Census divisions.

Trade flows within a division are taken to equal the sum of the internal trade flows of its states plus the flows between these states. Trade flows between divisions are given by the sum of trade flows between their respective states. Similarly, trade flows from a division to a foreign country are given as the sum of exports from the states in the division to the foreign country.

Table 2 reports regression results that correspond to Table 1. We use the simple average of distances associated with the underlying individual trade flows. The division-based international border dummy coefficients are −0.36 and −0.39 and thus considerably smaller in magnitude and significantly different from the corresponding state-based estimates of −1.25 and −1.21 in Table 1. The division-based domestic border dummy coefficients are −1.17 and

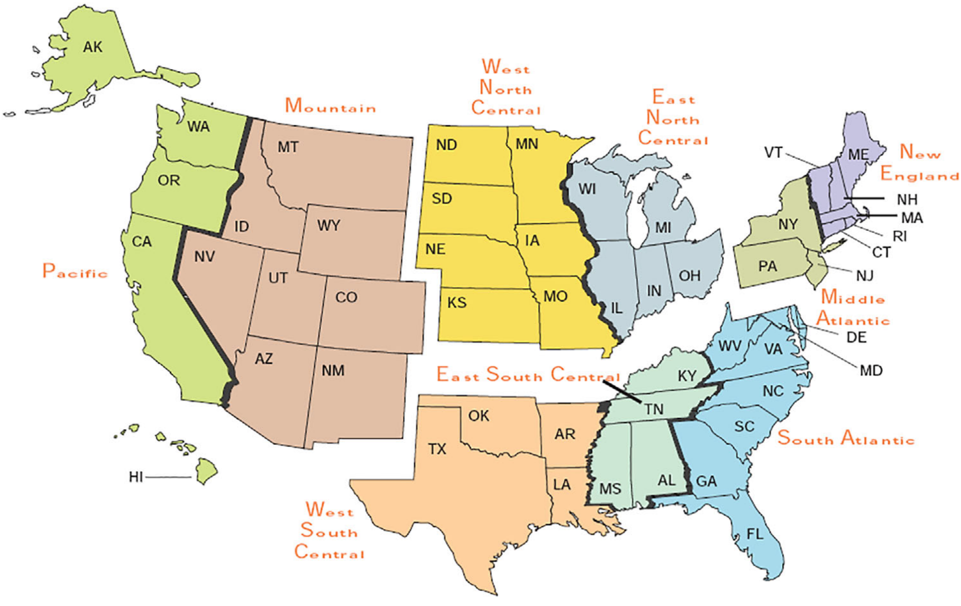



FIGURE 7

A MAP OF THE NINE U.S. CENSUS DIVISIONS (SOURCE: U.S. DEPARTMENT OF ENERGY)

TABLE 2
INTERNATIONAL AND DOMESTIC BORDER EFFECTS BASED ON U.S. CENSUS DIVISIONS

| Sample | U.S. and Foreign Countries | | U.S. only | |
|---|---|---|---|---|
| Year | 1993 | 1993, 1997, 2002, 2007 | 1993 | 1993, 1997, 2002, 2007 |
| | (1) | (2) | (3) | (4) |
| $\ln(dist_{ij})$ | −1.17*** | −1.21*** | −1.07*** | −1.04*** |
| | (0.03) | (0.03) | (0.10) | (0.08) |
| $INT_{ij}$ (international border dummy) | −0.36*** | −0.39*** | | |
| | (0.11) | (0.10) | | |
| $DOM_{ij}$ (domestic border dummy) | | | −1.17*** | −1.25*** |
| | | | (0.18) | (0.17) |
| Internal trade (within Census divisions) | no | no | yes | yes |
| Domestic trade (between Census divisions) | yes | yes | yes | yes |
| International trade (with foreign countries) | yes | yes | no | no |
| Observations | 2746 | 10,984 | 81 | 324 |
| Clusters | – | 2746 | – | 81 |
| Fixed effects | yes | yes | yes | yes |
| $R^2$ | 0.78 | 0.79 | 0.95 | 0.96 |

*Notes*: The dependent variable is ln(xij). OLS estimation. Robust standard errors are reported in parentheses, clustered around bilateral pairs ij in columns 2 and 4. Division and foreign country fixed effects in columns 1 and 2; exporter and importer fixed effects in columns 3 and 4; those fixed effects are time-varying in columns 2 and 4. *** denotes statistical significance at the 1% level.

−1.25 and thus smaller in magnitude than the corresponding state-based estimates of −1.47 and −1.48 in Table 1, albeit not statistically different. The distance coefficients are very similar between Tables 1 and 2. A common pattern arises: the border coefficients are further away from zero when states are the underlying spatial units, and the border coefficients are closer to zero when we use divisions as the larger underlying spatial units. This pattern mirrors the cross-sectional heterogeneity apparent in the individual border coefficients depicted in Figures 2–4.[40]

We briefly comment on the estimation method. Although the point estimates naturally change when we follow Santos Silva and Tenreyro (2006) and use Poisson pseudo maximum likelihood (PPML) estimation as opposed to OLS, the coefficient patterns are qualitatively the same. In particular, we find the same type of coefficient heterogeneity as in Figures 2–4, consistent with our theory. For the international border coefficients in Figures 2 and 3, we find a correlation of 75% between coefficients obtained with PPML and those obtained with OLS. The corresponding correlation for the domestic border coefficients in Figure 4 is 97%. Furthermore, the relationship between coefficients in Tables 1 and 2 is the same with PPML.

Finally, we add another aggregation exercise. Instead of using Census divisions, we merge the 48 contiguous states into hypothetical U.S. regions by sequentially merging them in a quasi-randomized way.[41] Aggregating trade flows and distances in the same way as above, we form four samples: 24 regions consisting of two states each, 12 regions consisting of four states each, eight regions consisting of six states each, and six regions consisting of eight states each.[42]

In Table 3, we present regression results for data pooled over the years 1993, 1997, 2002, and 2007. Columns 1–4 report coefficients for the international border dummy across the four samples. Although in column 1 based on 24 regions we obtain a coefficient of $-0.64$, its magnitude shrinks to $-0.21$ in column 4 based on six regions. Consistent with our theoretical prediction, we find that the larger the regions, the smaller the border dummy coefficient becomes in absolute value. Columns 5–8 report results for the domestic border effect. Here we do not see a systematic change in border dummy coefficients. However, we find a clear pattern in the distance coefficients. The further we aggregate, the weaker the distance elasticity in absolute magnitude. That is, aggregation reduces the estimated trade-impeding effect of space not through the domestic border dummy coefficient but the distance elasticity.[43]

4.7. *Multilateral Resistance Effects in General Equilibrium.* In their seminal paper, Anderson and van Wincoop (2003) highlight the role of general equilibrium. They show that small and large countries react differently to changes in international border barriers. Intuitively, removing the border leads to a reallocation of trade away from domestic towards international partners. But as a small country is more exposed to international trade and thus more exposed to the border barrier, this reallocation is relatively stronger for the small country.[44] This differential response between small and large countries is entirely driven by price index or "multilateral resistance" effects.

In our theoretical framework, however, multilateral resistance is symmetric across countries (see Appendix A.2). The differential trade response is instead driven by heterogeneity in the border effect itself due to spatial aggregation, as shown in Equation (21).

[40] As an alternative distance measure, we weight distances by individual trade flows. That is, distances associated with stronger trading activity are given proportionally larger weights. The international border dummy coefficients corresponding to columns 1 and 2 of Table 2 are $-0.31$ and $-0.34$ and thus roughly the same. The domestic border dummy coefficients corresponding to columns 3 and 4 are $-0.92$ and $-0.98$ and thus smaller in absolute magnitude. The reason is that weighted distances tend to be shorter than average distances such that the extent of internal trade seems less extreme.

[41] This procedure is quasi-random in the sense that we impose the restriction that no state must be left in isolation. For example, Maine is adjacent to only one state (New Hampshire) such that those two states never end up in separate regions.

[42] These regions are described in detail in Appendix C.5.

[43] Although the average distance for cross-region trade remains roughly stable, the average distance for within-region trade increases dramatically (from 5.5 to 6.3 in logarithmic distance between columns 5 and 8 of Table 3).

[44] To be precise, the ratio of bilateral international trade to bilateral domestic trade increases more strongly for a country consisting of smaller regions such as Canada. Anderson and van Wincoop (2003), Section IV.C, discuss "the relatively small size of the Canadian economy" in the context of their data set of trade flows between Canadian provinces and U.S. states.

TABLE 3
INTERNATIONAL AND DOMESTIC BORDER EFFECTS BASED ON QUASI-RANDOMLY AGGREGATED U.S. STATES

| Sample | U.S. and Foreign Countries | | | | U.S. only | | | |
|---|---|---|---|---|---|---|---|---|
| Number of hypothetical U.S. regions | 24 | 12 | 8 | 6 | 24 | 12 | 8 | 6 |
| Number of U.S. states per region | 2 | 4 | 6 | 8 | 2 | 4 | 6 | 8 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(dist_{ij})$ | −1.23*** | −1.21*** | −1.22*** | −1.21*** | −1.29*** | −1.13*** | −0.95*** | −0.70*** |
| | (0.03) | (0.03) | (0.03) | (0.03) | (0.04) | (0.06) | (0.10) | (0.12) |
| $INT_{ij}$ (international border dummy) | −0.64*** | −0.44*** | −0.39*** | −0.21* | | | | |
| | (0.08) | (0.09) | (0.11) | (0.12) | | | | |
| $DOM_{ij}$ (domestic border dummy) | | | | | −1.36*** | −1.43*** | −1.35*** | −1.47*** |
| | | | | | (0.18) | (0.20) | (0.16) | (0.22) |
| Internal trade (within U.S. regions) | no | no | no | no | yes | yes | yes | yes |
| Domestic trade (between U.S. regions) | yes | yes | yes | yes | yes | yes | yes | yes |
| International trade (with foreign countries) | yes | yes | yes | yes | no | no | no | no |
| Observations | 15,772 | 11,812 | 10,724 | 10,228 | 2232 | 576 | 256 | 144 |
| Clusters | 3943 | 2953 | 2681 | 2557 | 558 | 144 | 64 | 36 |
| Fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| $R^2$ | 0.80 | 0.79 | 0.79 | 0.78 | 0.90 | 0.94 | 0.95 | 0.96 |

*Notes*: The dependent variable is ln(xij). Data are for the years 1993, 1997, 2002, and 2007. The 48 contiguous U.S. states are aggregated quasi-randomly into hypothetical U.S. regions with 24 regions in columns 1 and 5, 12 regions in columns 2 and 6, 8 regions in columns 3 and 7, and 6 regions in columns 4 and 8. For more information, see the main text. OLS estimation. Robust standard errors are reported in parentheses, clustered around bilateral pairs ij. Time-varying U.S. region and foreign country fixed effects in columns 1–4; time-varying exporter and importer fixed effects in columns 4–8. * and *** denote statistical significance at the 10% and 1% level, respectively.

Although multilateral resistance is the same across countries in our theory, we cannot assume this to be the case with actual trade flows. In Table 4, we explore the general equilibrium counterfactuals implied by removed international border barriers, accounting for both heterogeneous border effects as well as heterogeneous multilateral resistance effects. We use the same balanced sample as for column 2 of Table 1 based on 24,996 observations for the years 1993, 1997, 2002 and 2007 (6,249 observations per year).

In panel 1, we report counterfactuals based on removing a common international border barrier as in the standard Anderson and van Wincoop (2003) model. As in column 2 of Table 1, we estimate this border barrier based on the logarithmic version of the standard gravity equation (1) with logarithmic bilateral distance and country fixed effects as additional controls. The border dummy captures the U.S. international border only.[45] We then remove the U.S. international border and recompute the associated general equilibrium.[46]

Panel 1 presents the logarithmic differences between the counterfactual and initial equilibria. Removing the U.S. border leads to an increase in bilateral trade flows by 23% on average (see the top row of panel 1). Trade would have increased by 31% just through the direct (partial equilibrium) effect of reducing bilateral trade costs.[47] This direct effect is the same for all U.S. states by construction because we impose a common border barrier. The offsetting general equilibrium effect through falling multilateral resistance is 10% on average but varies somewhat across states, while the increase in incomes pushes up trade by 2%. In sum, there is a modest degree of variation across states due to the heterogeneous general equilibrium effects. For instance, the bilateral trade of California goes up by 24% on average, whereas the trade of Wyoming goes up by 21%.

In panel 2, we report counterfactuals based on our framework with heterogeneous border barriers. We estimate state-specific border coefficients as described in Section 4.4.1. Those are plotted in Figures 2 and 3. We also account for multilateral resistance effects when computing the counterfactual equilibrium. Removing the heterogeneous border barriers leads to average effects that are almost identical (see the top row of panel 2). However, the underlying effects for individual states exhibit much more variation. The key insight is that this variation is primarily driven by the heterogeneous direct effects (see column 2b), not multilateral resistance effects. The overall differences across states can be quite substantial. For instance, here the bilateral trade of California goes up by 13% on average, whereas the trade of Wyoming goes up over four times as much (61%). Consistent with our theory, small states are more affected by the removal of the border.[48]

Overall, we conclude that heterogeneous border barriers translate into heterogeneous trade effects. Quantitatively, this form of heterogeneity is considerably more important than heterogeneity associated with multilateral resistance effects.

4.8. *Exploring Potential Determinants of Border Effects.* It is conceivable that the heterogeneity of border effects documented in Section 4.4 is related to features at the state level, for example, differences in transport infrastructure. It is not clear a priori whether better infrastructure would be associated with weaker or stronger border effects, depending on whether infrastructure primarily facilitates within-border trade or cross-border trade (see Section 3.1).

[45] The distance and border dummy coefficients are $-1.21$ and $-0.60$, respectively, both highly significant at the 1% level. As the border dummy only captures the U.S. border, its coefficient is directly comparable to the individual border coefficients for U.S. states plotted in Figures 2 and 3. Their average is $-0.64$ and thus about the same.

[46] For the initial equilibrium, we take the income data for the 48 U.S. states and 50 large foreign countries in our sample for the year 1993, thus capturing the vast majority of global economic activity. Using our estimated distance and border dummy coefficients, we use numerical methods to compute the multilateral resistance variables and construct the associated bilateral trade flows based on gravity equation (1). For the counterfactual, we set the border dummy coefficient to zero and recompute the full equilibrium, assuming that the endowment quantities are fixed.

[47] Assuming $\sigma = 5$, this corresponds to a cut in trade costs by 7.75% as $0.31/(1-\sigma) = -0.0775$.

[48] For some states, the overall trade effect shows up as slightly negative (e.g., $-7\%$ for Connecticut). This happens because some individual border coefficients were estimated to have a positive sign (see Figures 2 and 3). Most of these positive coefficients are not significant, but we report the associated results in Table 4 nevertheless.

TABLE 4
GENERAL EQUILIBRIUM EFFECTS IN RESPONSE TO REMOVING THE U.S. INTERNATIONAL BORDER

| | **Panel 1: Common Border Effect** | | | | | | | **Panel 2: Heterogeneous Border Effects** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total Effect | | Direct Effect | | Indirect GE Effects | | | Total Effect | | Direct Effect | | Indirect GE Effects | | |
| U.S. State | $\Delta\ \ln(x_{ij})$ (1a) | = | $(1-\sigma)\ \Delta\ \ln(t_{ij})$ (1b) | + | $(\sigma-1)\ \Delta\ \ln(P_iP_j)$ (1c) | + | $\Delta\ \ln(y_iy_j/y^W)$ (1d) | $\Delta\ \ln(x_{ij})$ (2a) | = | $(1-\sigma)\ \Delta\ \ln(t_{ij})$ (2b) | + | $(\sigma-1)\ \Delta\ \ln(P_iP_j)$ (2c) | + | $\Delta\ \ln(y_iy_j/y^W)$ (2d) |
| **Average** | **0.23** | **=** | **0.31** | **+** | **−0.10** | **+** | **0.02** | **0.24** | **=** | **0.33** | **+** | **−0.11** | **+** | **0.02** |
| AL | 0.24 | = | 0.31 | + | −0.08 | + | 0.01 | 0.11 | = | 0.18 | + | −0.10 | + | 0.02 |
| AR | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | 0.45 | = | 0.60 | + | −0.19 | + | 0.04 |
| AZ | 0.21 | = | 0.31 | + | −0.12 | + | 0.02 | 0.32 | = | 0.45 | + | −0.17 | + | 0.03 |
| CA | 0.24 | = | 0.31 | + | −0.08 | + | 0.01 | 0.13 | = | 0.18 | + | −0.06 | + | 0.01 |
| CO | 0.23 | = | 0.31 | + | −0.10 | + | 0.02 | 0.40 | = | 0.51 | + | −0.14 | + | 0.03 |
| CT | 0.25 | = | 0.31 | + | −0.06 | + | 0.01 | −0.07 | = | −0.04 | + | −0.03 | + | 0.00 |
| DE | 0.25 | = | 0.31 | + | −0.06 | + | 0.01 | 0.01 | = | 0.05 | + | −0.05 | + | 0.01 |
| FL | 0.21 | = | 0.31 | + | −0.12 | + | 0.02 | −0.13 | = | −0.18 | + | 0.08 | + | −0.02 |
| GA | 0.23 | = | 0.31 | + | −0.09 | + | 0.01 | 0.01 | = | 0.05 | + | −0.05 | + | 0.01 |
| IA | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | 0.30 | = | 0.40 | + | −0.13 | + | 0.03 |
| ID | 0.23 | = | 0.31 | + | −0.09 | + | 0.01 | 0.58 | = | 0.71 | + | −0.16 | + | 0.03 |
| IL | 0.25 | = | 0.31 | + | −0.07 | + | 0.01 | 0.07 | = | 0.11 | + | −0.04 | + | 0.01 |
| IN | 0.24 | = | 0.31 | + | −0.09 | + | 0.01 | 0.23 | = | 0.31 | + | −0.10 | + | 0.02 |
| KS | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | 0.16 | = | 0.23 | + | −0.09 | + | 0.02 |
| KY | 0.24 | = | 0.31 | + | −0.08 | + | 0.01 | 0.18 | = | 0.25 | + | −0.09 | + | 0.02 |
| LA | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | −0.27 | = | −0.45 | + | 0.23 | + | −0.05 |
| MA | 0.23 | = | 0.31 | + | −0.09 | + | 0.01 | 0.03 | = | 0.08 | + | −0.06 | + | 0.01 |
| MD | 0.25 | = | 0.31 | + | −0.07 | + | 0.01 | 0.12 | = | 0.17 | + | −0.06 | + | 0.01 |
| ME | 0.20 | = | 0.31 | + | −0.13 | + | 0.02 | 0.44 | = | 0.66 | + | −0.28 | + | 0.05 |
| MI | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | 0.18 | = | 0.27 | + | −0.12 | + | 0.02 |
| MN | 0.25 | = | 0.31 | + | −0.06 | + | 0.01 | 0.20 | = | 0.25 | + | −0.06 | + | 0.01 |
| MO | 0.23 | = | 0.31 | + | −0.10 | + | 0.02 | 0.37 | = | 0.49 | + | −0.15 | + | 0.03 |
| MS | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | 0.20 | = | 0.32 | + | −0.15 | + | 0.03 |
| MT | 0.18 | = | 0.31 | + | −0.15 | + | 0.03 | 1.08 | = | 1.37 | + | −0.36 | + | 0.07 |
| NC | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | 0.03 | = | 0.08 | + | −0.06 | + | 0.01 |
| ND | 0.19 | = | 0.31 | + | −0.15 | + | 0.03 | 0.77 | = | 1.03 | + | −0.32 | + | 0.06 |
| NE | 0.23 | = | 0.31 | + | −0.10 | + | 0.02 | 0.48 | = | 0.61 | + | −0.16 | + | 0.03 |
| NH | 0.23 | = | 0.31 | + | −0.09 | + | 0.01 | 0.28 | = | 0.42 | + | −0.17 | + | 0.03 |
| NJ | 0.26 | = | 0.31 | + | −0.05 | + | 0.01 | 0.03 | = | 0.06 | + | −0.04 | + | 0.01 |

*(Continued)*

TABLE 4
(CONTINUED)

| | Panel 1: Common Border Effect | | | | | | | Panel 2: Heterogeneous Border Effects | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total Effect | | Direct Effect | | Indirect GE Effects | | | Total Effect | | Direct Effect | | Indirect GE Effects | | |
| U.S. State | $\Delta \ln(x_{ij})$ (1a) | = | $(1-\sigma)\ \Delta \ln(t_{ij})$ (1b) | + | $(\sigma-1)\ \Delta \ln(P_i P_j)$ (1c) | + | $\Delta \ln(y_i y_j / y^W)$ (1d) | $\Delta \ln(x_{ij})$ (2a) | = | $(1-\sigma)\ \Delta \ln(t_{ij})$ (2b) | + | $(\sigma-1)\ \Delta \ln(P_i P_j)$ (2c) | + | $\Delta \ln(y_i y_j / y^W)$ (2d) |
| NM | 0.20 | = | 0.31 | + | −0.13 | + | 0.02 | 0.51 | = | 0.70 | + | −0.24 | + | 0.05 |
| NV | 0.25 | = | 0.31 | + | −0.06 | + | 0.01 | 0.83 | = | 0.91 | + | −0.11 | + | 0.02 |
| NY | 0.18 | = | 0.31 | + | −0.15 | + | 0.03 | −0.05 | = | −0.10 | + | 0.06 | + | −0.01 |
| OH | 0.23 | = | 0.31 | + | −0.09 | + | 0.01 | 0.14 | = | 0.21 | + | −0.09 | + | 0.02 |
| OK | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | 0.35 | = | 0.48 | + | −0.17 | + | 0.03 |
| OR | 0.23 | = | 0.31 | + | −0.09 | + | 0.01 | 0.27 | = | 0.37 | + | −0.12 | + | 0.02 |
| PA | 0.22 | = | 0.31 | + | −0.11 | + | 0.02 | 0.14 | = | 0.25 | + | −0.13 | + | 0.02 |
| RI | 0.25 | = | 0.31 | + | −0.07 | + | 0.01 | 0.19 | = | 0.28 | + | −0.10 | + | 0.02 |
| SC | 0.23 | = | 0.31 | + | −0.10 | + | 0.01 | 0.00 | = | 0.04 | + | −0.04 | + | 0.01 |
| SD | 0.20 | = | 0.31 | + | −0.14 | + | 0.02 | 0.86 | = | 1.11 | + | −0.30 | + | 0.06 |
| TN | 0.23 | = | 0.31 | + | −0.09 | + | 0.01 | 0.18 | = | 0.26 | + | −0.11 | + | 0.02 |
| TX | 0.21 | = | 0.31 | + | −0.11 | + | 0.02 | −0.09 | = | −0.10 | + | 0.02 | + | 0.00 |
| UT | 0.22 | = | 0.31 | + | −0.10 | + | 0.02 | 0.42 | = | 0.54 | + | −0.15 | + | 0.03 |
| VA | 0.25 | = | 0.31 | + | −0.07 | + | 0.01 | −0.03 | = | −0.01 | + | −0.03 | + | 0.00 |
| VT | 0.19 | = | 0.31 | + | −0.14 | + | 0.02 | 0.21 | = | 0.42 | + | −0.26 | + | 0.05 |
| WA | 0.20 | = | 0.31 | + | −0.13 | + | 0.02 | −0.09 | = | −0.12 | + | 0.05 | + | −0.01 |
| WI | 0.23 | = | 0.31 | + | −0.10 | + | 0.01 | 0.11 | = | 0.17 | + | −0.07 | + | 0.01 |
| WV | 0.23 | = | 0.31 | + | −0.09 | + | 0.01 | 0.20 | = | 0.28 | + | −0.10 | + | 0.02 |
| WY | 0.21 | = | 0.31 | + | −0.11 | + | 0.02 | 0.61 | = | 0.78 | + | −0.21 | + | 0.04 |

*Notes*: This table reports logarithmic differences of variables between an initial equilibrium with international border barriers and a counterfactual equilibrium where these border barriers are removed. Two scenarios are considered. The first scenario in panel 1 is based on a common international border barrier for all 48 U.S. states in the sample. The second scenario in panel 2 is based on heterogeneous international border barriers across U.S. states. The sample is balanced over the years 1993, 1997, 2002, and 2007 with 24,996 observations in total (6,249 for each year). Apart from the international border dummies, the underlying regressions include log distance and time-varying state and country fixed effects. Columns 1a and 2a: average change in bilateral trade (total effect); columns 1b and 2b: change in bilateral trade costs scaled by the substitution elasticity due to the removal of the international border; columns 1c and 2c: average change in multilateral resistances scaled by the substitution elasticity; columns 1d and 2d: average change in incomes. The first row reports the simple average across all states. The reported numbers are rounded off to two decimal digits. For more information see the main text.

Intuitively, if infrastructure facilitates all trade flows by the same degree, then it should not be systematically related to estimated border dummy coefficients.

To explore the role of transport infrastructure at the state level, we collect data on two time-varying measures: the length of public roads per capita and the number of airports per capita. As an additional potential determinant of border effects, we also collect data on personal income per capita as a measure of productivity. We describe these measures in more detail in Appendix C.4. We consider our panel data from Table 1, and we interact the border dummy variables with the above measures. We present regression results in Table 5. Columns 1–3 are for the international border effect, and columns 4–6 are for the domestic border effect.[49]

Column 1 shows that our measures of transport infrastructure have a negative interaction effect with the international border dummy. This means that the more roads and airports per capita there are in a state, the more negative (i.e., the larger in absolute magnitude) is the impeding effect associated with an international border. The interpretation would be that in relative terms, this transport infrastructure appears to predominantly facilitate domestic trade between U.S. states instead of international trade between U.S. states and foreign countries. Income per capita exhibits the same qualitative effect. As a comparison, in column 2 we interact the international border dummy with state GDP. The positive interaction coefficient confirms our theoretical prediction and the pattern in Figure 3: the larger a U.S. state, the smaller is the associated border effect in absolute magnitude. In column 3, we combine all regressors. The GDP interaction coefficient remains positive and significant but is reduced in magnitude.

Columns 4–6 show the corresponding interaction effects for the domestic border dummy. In column 4, only the interaction term of road length per capita is significant. The interpretation here would be that the road network facilitates trade within U.S. states relatively more than trade between U.S. states. The GDP interaction effect in column 5 again confirms our theoretical prediction: larger U.S. states are associated with weaker border effects in absolute magnitude. But this coefficient becomes insignificant in column 6 when we include all regressors.

We perform an additional check in Table 6. We take the individual border effect coefficients from Sections 4.4.1 and 4.4.2 (also see Figures 2–4) and regress them on the infrastructure and productivity measures at the U.S. state level. In column 1, only the airport variable is significantly related to the 48 international border coefficients. Consistent with column 1 of Table 5, we find that states with more extensive airport infrastructure are associated with stronger (i.e., more negative) dummy coefficients. In column 2, we regress on GDP and again find that larger states tend to have weaker border coefficients, although the GDP coefficient becomes insignificant in column 3 when we include all regressors. The latter finding is mirrored in columns 5 and 6 for the domestic border effect. The infrastructure measures in columns 4 and 6 do not appear systematically related to the size of domestic border coefficients, with the (marginal) exception of road length per capita.

In summary, we continue to find evidence that border effects are systematically related to state size. Although Table 5 shows evidence of a systematic relationship between transport infrastructure and the size of border effects, Table 6 is less clear about this link. Further evidence from different data sets would be helpful to clarify the role of transport infrastructure in determining border effect estimates.

4.9. *Evidence from Industry-Level Trade.* One potential explanation for our finding of heterogeneous border effects could be differences in the industry mix across states. To account

[49] As we have data of the measures for U.S. states only and not for foreign countries, we interact with the logarithmic value of the exporter whenever it is a U.S. state in columns 1–3 of Table 5, and we set values to 0 whenever U.S. states are not involved. As we have trade flows within the U.S. only in columns 4–6, we interact with the logarithm of the product of the exporter and importer values.

TABLE 5
INTERNATIONAL AND DOMESTIC BORDER EFFECTS INTERACTED WITH POTENTIAL DETERMINANTS

| Sample | U.S. and Foreign Countries | | | U.S. only | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\ln(dist_{ij})$ | −1.22*** | −1.20*** | −1.21*** | −1.09*** | −1.08*** | −1.09*** |
| | (0.02) | (0.02) | (0.02) | (0.03) | (0.03) | (0.03) |
| $INT_{ij}$ | −12.42*** | −11.00*** | −9.55*** | | | |
| | (1.38) | (0.52) | (1.52) | | | |
| $INT_{ij} * \ln(\text{road length per capita}_i)$ | −0.43*** | | −0.40*** | | | |
| | (0.05) | | (0.05) | | | |
| $INT_{ij} * \ln(\text{airports per capita}_i)$ | −0.31*** | | −0.12* | | | |
| | (0.06) | | (0.07) | | | |
| $\text{INT}_{\text{ij}} * \ln(\text{income per capita}_i)$ | −0.24*** | | −0.39*** | | | |
| | (0.09) | | (0.09) | | | |
| $INT_{ij} * \ln(GDP_i)$ | | 0.41*** | 0.19*** | | | |
| | | (0.02) | (0.03) | | | |
| $DOM_{ij}$ | | | | −10.74** | −8.16*** | −9.62* |
| | | | | (4.77) | (1.89) | (5.34) |
| $DOM_{ij} * \ln(\text{road length per capita}_{ij})$ | | | | −0.41*** | | −0.41*** |
| | | | | (0.14) | | (0.14) |
| $DOM_{ij} * \ln(\text{airports per capita}_{ij})$ | | | | −0.12 | | −0.07 |
| | | | | (0.17) | | (0.19) |
| $DOM_{ij} * \ln(\text{income per capita}_{ij})$ | | | | 0.14 | | 0.09 |
| | | | | (0.13) | | (0.17) |
| $DOM_{ij} * \ln(GDP_{ij})$ | | | | | 0.30*** | 0.05 |
| | | | | | (0.08) | (0.11) |
| Internal trade (within U.S. states) | no | no | no | yes | yes | yes |
| Domestic trade (between U.S. states) | yes | yes | yes | yes | yes | yes |
| International trade (with foreign countries) | yes | yes | yes | no | no | no |
| Observations | 24,996 | 24,996 | 24,996 | 6904 | 6904 | 6904 |
| Clusters | 6249 | 6249 | 6249 | 1726 | 1726 | 1726 |
| Fixed effects | yes | yes | yes | yes | yes | yes |
| $R^2$ | 0.83 | 0.83 | 0.83 | 0.90 | 0.90 | 0.90 |

*Notes*: The dependent variable is ln(xij). Data are for the years 1993, 1997, 2002, and 2007. The interacted determinants in columns 1–3 are the logarithmic values of exporting U.S. states only; their main effects are included but not reported; the interacted determinants in columns 4–6 are the logarithmic values of the products of the exporter and the importer values of U.S. states; their main effects are collinear with the fixed effects and thus drop out. OLS estimation. Robust standard errors are reported in parentheses, clustered around bilateral pairs ij. Time-varying state and foreign country fixed effects in columns 1–3; time-varying exporter and importer fixed effects in columns 4–6. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

TABLE 6
REGRESSING INDIVIDUAL BORDER EFFECTS ON POTENTIAL DETERMINANTS

| | International Border Dummy Coefficients | | | Domestic Border Dummy Coefficients | | |
|---|---|---|---|---|---|---|
| Dependent Variable | ß$_r$ (1) | ß$_r$ (2) | ß$_r$ (3) | $\gamma_r$ (4) | $\gamma_r$ (5) | $\gamma_r$ (6) |
| ln(road length per capita$_r$) | −0.20 | | −0.19 | −0.73* | | −0.71* |
| | (0.15) | | (0.15) | (0.40) | | (0.41) |
| ln(airports per capita$_r$) | −0.63*** | | −0.53*** | −0.31 | | −0.17 |
| | (0.16) | | (0.19) | (0.44) | | (0.51) |
| ln(income per capita$_r$) | −0.25 | | −0.24 | 1.88 | | 1.88 |
| | (0.48) | | (0.48) | (1.30) | | (1.31) |
| ln($GDP_r$) | | 0.44*** | 0.10 | | 0.66*** | 0.13 |
| | | (0.07) | (0.09) | | (0.18) | (0.25) |
| Observations | 48 | 48 | 48 | 48 | 48 | 48 |
| $R^2$ | 0.64 | 0.46 | 0.65 | 0.41 | 0.23 | 0.42 |

*Notes*: The dependent variables are the international border dummy coefficients ß$_r$ in columns 1–3 and the domestic border dummy coefficients $\gamma_r$ in columns 4–6 (one for each of the 48 contiguous U.S. states). The regressor values are for the year 2007. OLS estimation. Robust standard errors are reported in parentheses (not bootstrapped). A constant is included but not reported. * and *** denote statistical significance at the 10% and 1% level, respectively.

for such composition effects, we estimate heterogeneous border effects based on trade flows at the industry level. If the spatial attenuation effect we document in previous sections is not related to industry composition effects, then we would expect border effect patterns at the industry level to qualitatively resemble those at the aggregate level.

For this purpose, we examine trade flows at the industry level from Crafts and Klein (2015) for the year 2007, sourced from the Commodity Flow Survey. By "industry," we mean "commodity" in Commodity Flow Survey terminology. We have data on domestic flows within the United States but not with foreign countries. As in previous sections, we focus on the 48 contiguous states. For aggregate data ("All Commodities" ), we have 2,125 observations (compared to a maximum possible of $48\times48 = 2,304$ in a fully balanced data set). We naturally have fewer observations at the industry level due to missing trade flows or observations withheld due to data concerns. We describe the data in more detail in Appendix C.6. For the estimation, we drop industries with a very small number of observations. Specifically, our criterion is to drop industries with fewer than a quarter of the observations available for aggregate data (i.e., fewer than 532), although our results are not sensitive to this particular criterion. This leaves us with 22 industries, listed in Appendix C.6.

We first estimate the standard common domestic border effect in exactly the same way as in column 3 of Table 1. For aggregate data, the coefficient stands at −1.51 with a standard error of 0.20. It is thus virtually the same as in Table 1. We then run the same specification for each industry. In all 22 industries, we estimate a negative border coefficient, and in 21 industries it is statistically significant at the 1% level (the average coefficient value is −1.42 with an average standard error of 0.21 ). This evidence is consistent with Crafts and Klein (2015) who also find statistically significant domestic border dummy coefficients.

Second and most importantly, we estimate individual domestic border effects by allowing them to vary by U.S. state, following our approach in Section 4.4.2. We start by estimating these coefficients for aggregate data. The average of the 48 individual coefficients is −1.37 with an average standard error of 0.14. This is very close to the results in Section 4.4.2 where we report an average coefficient of −1.32 with an average standard error of 0.13. A plot of the individual coefficients produces a heterogeneity pattern that looks very similar to Figure 4, with values spanning a range of more than six log points (the smallest being −4.40 and the largest 2.46).

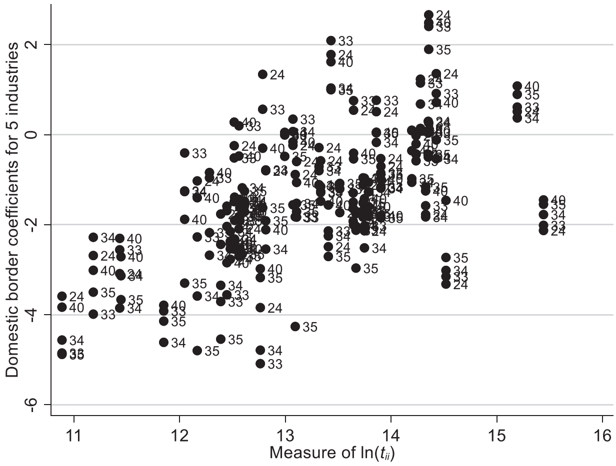
We then estimate analogous individual coefficients based on industry-level data, obtained from separate regressions for each industry. To convey a visual impression, in Figure 8 we illustrate the coefficients for the five industries with the largest number of observations. We plot them against the same measure of logarithmic internal trade costs as in Figure 4. The correlation stands at 56% and the corresponding correlation with logarithmic state GDP is 49%. This relationship becomes noisier for the full set of 22 industries, but we still find a clear positive relationship (the correlation is 35% with internal trade costs and 28% with GDP).

Overall, we find strong additional evidence based on industry-level trade flows that is consistent with the theoretical prediction of spatial attenuation. That is, individual border dummy coefficients at the industry level display the same pattern of heterogeneity as in aggregate data: larger states are systematically associated with weaker border effects.

## 5. DISCUSSION: ARE BORDER EFFECTS STATISTICAL ARTEFACTS?

The aim of much of the empirical literature on border effects is to identify the parameters $\beta$ (for the international border effect) and $\gamma$ (for the domestic border effect). However, as we have shown in the context of equations (22) and (45), these parameters cannot be identified empirically in standard gravity regressions.[50] The reason is that their estimates are subject to spatial attenuation effects.

[50] This insight is similar in spirit to the result by Gorodnichenko and Tesar (2009) who show that based on price data, border effects cannot be identified by comparing price dispersion across countries.




NOTES: The five industries are Plastics and rubber (code 24), Articles of base metal (code 33), Machinery (code 34), Electronic and other electrical equipment and components and office equipment (code 35), and Miscellaneous manufactured products (code 40). The coefficients, although plotted jointly, are obtained from separate regressions for each industry. The mean of the coefficients is −1.40. The average standard error is 0.27 (not plotted). More details are provided in the main text.

FIGURE 8

PLOT OF DOMESTIC BORDER DUMMY COEFFICIENTS BASED ON INDUSTRY-LEVEL TRADE FLOWS FOR THE 48 CONTIGUOUS U.S. STATES AGAINST THE LOGARITHM OF A MODEL-CONSISTENT PROXY FOR AGGREGATE INTERNAL TRADE COSTS $t_{ii}$

We add a note on the interpretation of domestic border effects. Although there is a friction $t^S$ between states in our model of the domestic border effect, this friction is not specific to state borders in our model. Rather, it is a general spatial friction that appears between micro regions regardless of whether they happen to be in different states or not. This general spatial friction nevertheless leads to significant domestic border dummy coefficients. That is, even if no specific frictions exist at the border, traditional gravity estimation can still indicate significant border effects, and these can be very large. It would therefore be wrong to interpret such border coefficients as solely reflecting frictions associated with state borders. Through the lens of our model the domestic border effect can therefore be seen as a statistical artefact (in the sense that it identifies frictions that are not specific to domestic borders).

In actual data, however, those coefficients might reflect a combination of general spatial frictions and—to the extent that they exist—frictions that specifically accrue at state borders. For economically less relevant entities (such as ZIP codes or U.S. Census divisions), it is hard to see domestic border effects as anything else than statistical artefacts. But for economically more meaningful entities such as important administrative units, domestic border effects are plausible. For instance, Nitsch and Wolf (2013) argue that domestic border effects might stem from informational frictions in the form of separate social and business networks. Similarly, Wrona (2018) provides a network-based explanation for domestic border effects, induced by long-lasting shocks throughout Japanese history. Using Commodity Flow Survey data, Felbermayr and Gröschl (2014) identify a friction between U.S. states that is associated with the divisive history of the Union-Confederacy border.

For the international border effect, there is a friction specific to crossing an international border in our model as long as we have $\beta < 0$.[51] This friction generally cannot be identified from traditional gravity estimation. But overall, we have little doubt that international border effects exist given the real counterparts in terms of tariffs, customs checks, regulatory differences in product standards, etc. Nevertheless, spatial attenuation effects also occur in the in-

[51] If $\beta = 0$, the international border friction does not exist. But as Equation (21) demonstrates, estimation based on aggregate data would still yield heterogeneous coefficients. In our model, those would be *positive*.

ternational context. As soon as we aggregate across space, international border dummy coefficients are pushed upwards toward zero.

## 6. CONCLUSION

We build a model of spatial aggregation that yields precise analytical results for border effects. Symmetric micro regions are aggregated into larger macro regions. Our theory shows how spatial aggregation affects the internal and bilateral trade costs of aggregated regions, and in turn estimated border effects. The main theoretical result is that aggregation leads to border effect heterogeneity: larger regions or countries are associated with border effects closer to zero, and vice versa. We call this the spatial attenuation effect. The intuition is that due to spatial frictions, aggregation across space increases the cost of trading within borders relative to trading across borders.

As an empirical test of the implications of our model, 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) as well as trade flows from an individual U.S. state to a foreign country (Minnesota–France). This data set allows us to estimate the trade effects of crossing the domestic state border and the U.S. international border. It also allows us to estimate these effects individually by state.

As predicted by our theoretical framework, we find that the larger the state, the weaker its domestic border effect and the weaker its international border effect. In addition, both border effects decline in magnitude when we aggregate states into larger units. We also find substantial sample composition effects when small and large states are systematically dropped from the sample. Overall, we conclude that border effects are inherently heterogeneous. This underlying heterogeneity drives the magnitude of standard common border dummy coefficients.

How should researchers proceed in practice? Of course, many researchers might still want to estimate standard border effects, not least because of the established convention in the literature. But in our view it would be useful to follow four steps. First, researchers can estimate heterogeneous border effects and report their range. Second, they can verify that their results are not driven by sample composition effects that depend on the size of spatial units. Third, researchers can replicate their results at different levels of aggregation. And fourth, researchers can exercise more caution when it comes to the interpretation of their results. They need to take into account that the magnitude of border effects might in part be a statistical artefact of aggregation stemming from spatial attenuation.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table C.1: Aggregation and the domestic border effect (symmetric regions)
Figure C.1: Plot of domestic border dummy coefficients for the 48 contiguous U.S. states based on actual data (as in Figure 4) against domestic border dummy coefficients based on simulated data.
Figure C.2: Plot of a simulated measure of logarithmic trade costs $t_{ii}$ for U.S. states against population shares of U.S. states (both in logarithms).

## REFERENCES

AGNOSTEVA, D., J. E. ANDERSON, AND V. YOTOV, "Intra-National Trade Costs: Assaying Regional Frictions," *European Economic Review* 112 (2019), 32–50.

ALLEN, T., AND C. ARKOLAKIS, "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics* 129 (2014), 1085–140.

ANDERSON, J., AND E. VAN WINCOOP, "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review* 93 (2003), 170–92.

———, "Trade Costs," *Journal of Economic Literature* 42 (2004), 691–751.

———, AND Y. YOTOV, "The Changing Incidence of Geography," *American Economic Review* 100 (2010), 2157–87.

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

BALISTRERI, E., AND R. HILLBERRY, "Structural Estimation and the Border Puzzle," *Journal of International Economics* 72 (2007), 451–63.

BEHRENS, K., C. ERTUR, AND W. KOCH, "Dual Gravity: Using Spatial Econometrics to Control for Multilateral Resistance," *Journal of Applied Econometrics* 27 (2012), 773–94.

BEMROSE, R., W. BROWN, AND J. TWEEDLE, 'Going the Distance: Estimating the Effect of Provincial Borders on Trade When Geography (and Everything Else) Matters," *Canadian Journal of Economics* 53 (2020), 1098–131.

BRIANT, A., P. COMBES, AND M. LAFOURCADE, "Dots to Boxes: Do the Size and Shape of Spatial Units Jeopardize Economic Geography Estimations?," *Journal of Urban Economics* 67 (2010), 287–302.

CASSEY, A., "State Export Data: Origin of Movement vs. Origin of Production," *Journal of Economic and Social Measurement* 34 (2009), 241–68.

CHANEY, T., "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review* 98 (2008), 1707–21.

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

COUGHLIN, C., "The Increasing Importance of Proximity for Exports from U.S. States," *Federal Reserve Bank of St. Louis Review* 86 (2004), 1–18.

———, AND T. MANDELBAUM, "Measuring State Exports: Is There a Better Way?," *Federal Reserve Bank of St. Louis Review* 73 (1991), 65–79.

———, AND D. NOVY, "Is the International Border Effect Larger than the Domestic Border Effect? Evidence from U.S. Trade," *CESifo Economic Studies* 59 (2013), 249–76.

———, AND H. WALL, "NAFTA and the Changing Pattern of State Exports," *Papers in Regional Science* 82 (2003), 427–50.

CRAFTS, N., AND A. KLEIN, "Geography and Intra-National Home Bias: U.S. Domestic Trade in 1949 and 2007," *Journal of Economic Geography* 15 (2015), 477–97.

CRONOVICH, R., AND R. GAZEL, "How Reliable Are the MISER Foreign Trade Data?" Unpublished manuscript, 1999.

EATON, J., AND S. KORTUM, "Technology, Geography and Trade," *Econometrica* 70 (2002), 1741–79.

ERLBAUM, N., AND J. HOLGUIN-VERAS, "Some Suggestions for Improving CFS Data Products," Commodity Flow Survey Conference, July 8–9, 2005, Transportation Research Circular (2006), Number E-C088, 77–97.

FALLY, T., R. PAILLACAR, AND C. TERRA, "Economic Geography and Wages in Brazil: Evidence from Micro-Data," *Journal of Development Economics* 91 (2010), 155–68.

FELBERMAYR, G., AND J. GRÖSCHL, "Within U.S. Trade and the Long Shadow of the American Secession," *Economic Inquiry* 52 (2014), 382–404.

FOTHERINGHAM, A., AND D. WONG, "The Modifiable Areal Unit Problem in Multivariate Statistical Analysis," *Environment and Planning A* 23 (1991), 1025–44.

GORODNICHENKO, Y., AND L. TESAR, "Border Effect or Country Effect? Seattle May Not Be So Far from Vancouver After All," *American Economic Journal: Macroeconomics* 1 (2009), 219–41.

HAVRÁNEK, T., AND Z. IRŠOVÁ, "Do Borders Really Slash Trade? A Meta-Analysis," *IMF Economic Review* 65 (2017), 365–96.

HEAD, K., AND T. MAYER, "Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade," in P. Bergeijk, and S. Brakman, eds., *The Gravity Model in International Trade: Advances and Applications* (Cambridge University Press, Cambridge, 2009), 165–92.

———, "Gravity Equations: Workhorse, Toolkit, and Cookbook," in G. Gopinath, E. Helpman, and K. Rogoff, eds., *Handbook of International Economics*, Chapter 3 (Elsevier, Amsterdam, 2014), 131–95.

HELLIWELL, J., "National Borders, Trade and Migration," *Pacific Economic Review* 2 (1997), 165–85.

———, "How Much Do National Borders Matter?" Brookings Institution, Washington, DC, 1998.

HILLBERRY, R., "Aggregation Bias, Compositional Change, and the Border Effect," *Canadian Journal of Economics* 35 (2002), 517–30.

———, AND HUMMELS, D., "Intranational Home Bias: Some Explanations," *Review of Economics and Statistics* 85 (2003), 1089-92.

———, "Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data," *European Economic Review* 52 (2008), 527–50.

HINZ, J., "The View from Space: Theory-Based Time-Varying Distances in the Gravity Model," Working Paper, Paris School of Economics, 2016.

MCCALLUM, J., "National Borders Matter: Canada-U.S. Regional Trade Patterns," *American Economic Review* 85 (1995), 615–23.

MELITZ, M., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica* 71 (2003), 1695–725.

MILLIMET, D., AND T. OSANG, "Do State Borders Matter for U.S. Intranational Trade? The Role of History and Internal Migration," *Canadian Journal of Economics* 40 (2007), 93–126.

NITSCH, V., "National Borders and International Trade: Evidence from the European Union," *Canadian Journal of Economics* 33 (2000), 1091–105.

———, AND N. WOLF, "Tear Down this Wall: On the Persistence of Borders in Trade," *Canadian Journal of Economics* 46 (2013), 154–79.

RAMONDO, N., A. RODRÍGUEZ-CLARE, AND M. SABORÍO-RODRÍGUEZ, "Trade, Domestic Frictions, and Scale Effects," *American Economic Review* 106 (2016), 3159–84.

SANTOS, SILVA, J., AND S. TENREYRO, "The Log of Gravity," *Review of Economics and Statistics* 88 (2006), 641–58.

SMITH, P., "Are Weak Patent Rights a Barrier to U.S. Exports?," *Journal of International Economics* 48 (1999), 151–77.

WEI, S., "Intra-National Versus International Trade: How Stubborn are Nations in Global Integration?," NBER Working Paper No. 5531, 1996.

WOLF, H., "Intranational Home Bias in Trade," *Review of Economics and Statistics* 82 (2000), 555–63.

WRONA, J., "Border Effects without Borders: What Divides Japan's Internal Trade?," *International Economic Review* 59 (2018), 1209–62.