The US-Canada Border Effect:
Evidence from Online Commerce

Bo Cowgill
Columbia University

Cosmina Dorobantu
Oxford University*

August 14, 2018

Abstract

How do national borders affect trade? We examine the US-Canada border effect using a large, proprietary dataset from Google, which covers search, advertising, and e-commerce for a wide variety of economic sectors (particularly trade in services). We document a large, statistically significant US-Canada border effect, even in a setting with relatively low search costs: intranational trade is 6.7 times higher than international trade in our data. We find that a large fraction of the US-Canada border coefficient (about 1/3rd) arises from consumer purchase behavior after arriving on sellers' online storefronts. The remaining 2/3rds appears in arrival rates on sellers' websites. We also find a strong border effect in virtual goods and downloadable products which do not require shipping, as well as business-to-consumer trades in final goods (rather than intermediate goods). When we disaggregate our data by economic sector, we find widely varying border effects. The sectors with the highest US-Canada border effects feature services whose consumption is tied to particular location and goods that face large regulatory and bureaucratic hurdles at the border.

JEL Classification: F10, L86, F02.

1

1 Introduction

Many scholars, politicians and business leaders have forecast the gradual weakening of national boundaries. Former US Secretary of State John Kerry told graduates in 2016 to prepare for a "borderless world."[1] Jean-Claude Juncker, President of the European Commission, told journalists in 2016 that "borders are the worst invention ever made by politicians."[2] The consensus about borders has even inspired a backlash, as some political entrepreneurs in the United States and Europe express frustration with economic and cultural globalization.

The Internet is often portrayed in this dialogue as an agent of borderlessness. Google chairman Eric Schmidt proclaimed in 2012 that "the Web will dissolve national borders."[3] A 2005 bestseller, *The World is Flat*, by *New York Times* columnist Thomas Friedman, was a notable source of enthusiasm about the Internet's role in softening borders.

How do national borders affect trade, particularly as trade becomes digitized? This paper studies the US-Canada border in Internet commerce, using econometric methods from international trade. The relationship between national borders and commerce is familiar to trade economists: one of the most persistent findings of the trade literature (starting with McCallum, 1995) is how strongly national borders appear to suppress trade. Obstfeld and Rogoff (2001) included the "border effect" on their list of six major puzzles in international macroeconomics.

We use a rich search, advertising, and e-commerce dataset from Google to study the relevance of national borders in an online setting from 2008-2011. We find a strong border effect in online commerce, indicating that the border between US and Canada has a large and negative impact on online trade flows.[4] Our estimates indicate that online trade be-

---

[1]http://time.com/4321733/john-kerry-commencement-address-transcript-speech/, accessed September 13, 2016.

[2]https://www.thetimes.co.uk/article/juncker-borders-are-worst-of-inventions-5rgw9kb33, accessed September 13, 2016. Leaders of the US Central Intelligence Agency, National Intelligence Council, and advisors to the Director of National Intelligence have envisaged similar scenarios. See for example, "The End of the Nation-State?" http://www.nytimes.com/2013/10/13/opinion/sunday/the-end-of-the-nation-state.html, accessed September 16, 2016. Comparable ideas have been expressed by academics in business (Ohmae, 1999) and international relations (Keohane and Nye, 1998).

[3]The audience for this remark included German Chancellor Angela Merkel and Brazilian President Dilma Rousseff – "Schmidt: The Web will dissolve national barriers," http://news.cnet.com/8301-30685_3-57390831-264/schmidt-the-web-will-dissolve-national-barriers/. Accessed September 13, 2016.

[4]To our knowledge, ours is the first paper using online trade data in order to shed new light on the US-

tween two US states or two Canadian provinces is 6.7 times higher than trade between a US state and a Canadian province.[5]

We use the dataset's unique features to study the nature of border effects in trade. The e-commerce transactions in our data evade many of the traditional explanations for border effects, such as shipping costs, information costs, and multistage production. The richness of the data also allows us to measure or control for other hypotheses proposed by previous researchers for the border effect (such as data aggregation biases). Thus, our estimates contain features that complement the existing literature on border effects and international trade, both online and offline.

We have four main findings. First, we separate the US-Canada border effect on "menus" – or choices available to buyers during search – from buyers' purchase decisions conditional on menus. To achieve this, we use a unique property of our data: a consumer can transact with a foreign business only if (a) the business first makes a conscious decision to target ads to that consumer's geography, and (b) the consumer clicks on the targeted ad. We find that a large portion of the US-Canada border effect (about 33%) persists after accounting for decisions affecting consumers' final choice sets. The remaining two-thirds of the effect comes before the final purchase decision, during the process of a consumer arriving on a seller's website via browsing and advertising. We provide quantitative evidence that the border mostly affects availability – that is, buyers' menus and choice sets – but that a surprisingly large border effect persists even after accounting for these menu effects.[6]

Canada border effect. We have found three other papers in which international border effects are estimated in the statistical tables, but are not the main topic. None of these studies the US-Canada border, which is the canonical setting for the traditional (offline) literature on border effects. Einav et al. (2014) and Lendle et al. (2015) report estimates of border effects in statistical tables without much additional discussion. Hortaçsu et al. (2009) measure a subnational border effect inside the US and international border effects for nine South America countries, with some discussion of potential explanations.

[5]Although our estimates do not have a direct equivalent in the traditional (offline) estimates of border effects, the coefficient we measure is higher than many border effects in the standard literature. For example, the border effect we find in the online world is higher than the average border effect of 4.7 reported by Feenstra (2002), based on traditional trade figures from 1993.

[6]Stated differently, ad clicks are less likely to convert into sales when the transacting parties are not in the same country. The sequence of the purchase process for transactions in our data requires the firm to first make an entry decision and pay for an ad before a consumer can click and purchase. We thus interpret the final purchase decision (following the entry and click outcomes) as the consumer's choice. This indicates that consumers' reluctance to trade with foreign sellers – for price, customer support, or other reasons – is one of the causes of the US-Canada border effect.

Second, we examine the role of shipping costs in trade, using another unique feature of our data: a large portion of our sample involves the purchase of digital goods with no physical shipping costs (software upgrades, downloads, and the purchase of virtual goods). Estimating the US-Canada border effect on trade in "virtual" goods, for which shipping costs are irrelevant, we find that the border effect is even higher than in the overall sample. Thus, strong border effects can be generated even without shipping costs.

Our third finding is that border effects vary widely across sectors of online economic activity. When we estimate the border effect for individual NAICS2 industries, the highest coefficient is 7.4 times the lowest.[7] Our sectoral results also show that the US-Canada border effect is largest for service categories that need to be consumed in a particular location, such as *Administrative support and waste management*, *Real estate*, and *Public administration*.[8]

Our fourth main finding concerns trade in final goods (retail) without trade in intermediate inputs. We again find a large border coefficient, suggesting that intranational retail trade online is 6.2 times higher than international retail trade. Some scholars have suggested multi-stage production as a possible explanation for the border effect (Yi, 2010; Rossi-Hansberg, 2005 and Hillberry and Hummels, 2008). Our data on business-to-consumer transactions in the online retail sector allow us to measure border effects in a setting without cross-border trade in intermediate goods. Our results still show a strong border effect. We also see that for certain products, like motor vehicles, the regulatory and bureaucratic burdens on end-consumers contribute to the high border effect that we observe.

Throughout this paper, we provide estimates not only of US-Canada border effects, but also of distance effects and subnational border effects for a variety of online sectors. Our estimates complement the existing literature on subnational border effects, which mostly uses data on business-to-business transactions from the US Commodity Flow Survey (CFS).[9] Our sample, with consumer transactions from a rich array of online retail and

---

[7]These results are qualitatively similar to Anderson et al.'s (2014) findings in offline transactions data.

[8]By examining a wider array of services and industries, our results provide additional evidence for the intuition behind Hortaçsu et al.'s (2009) discovery of a same-city bias for opera tickets (and other merchandise on eBay tied to a location).

[9]The CFS mostly covers the movement of commodities and business-to-business transactions such as mining, manufacturing, and wholesale distribution. The creator of the CSF, the US Department of Transportation, states: "most retail and services industries are excluded from the [Commodity Flow] survey." http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012.html, accessed September 13, 2016. For an example of literature using the CFS, see Wolf (2000) and successors.

service providers, provides a useful counterpart to the CFS.

We also provide some of the first estimates of border effects in services and the only estimate in the literature of the border effect in online services.

While the transactions in this paper come from e-commerce, many of our results point to more general phenomena. The concepts of "border effects in availability" or "border effect, conditional on menus," which we measure using online browsing data, have offline equivalents. In particular, the border effect we find in "ad clicks" may have an offline equivalent in firms choices of export markets to enter, and the extent that consumers are willing to consider those firms after their entry. These offline equivalents may be very difficult to measure. However, our use of data about search, advertising and clicks allows us to measure these relationships in a rapidly growing type of trade (e-commerce).[10]

The contribution of this paper is twofold. First, we provide the first estimates of the US-Canada border effect in online commerce across a variety of economic sectors. Second, we use the unique properties of our data – its information on (a) the detailed pre-purchase marketing and search behavior of buyers and sellers, and (b) on the types and sectors of each transaction – to obtain insights into the factors that give rise to border effects.

Online commerce is a large and growing form of trade. In 2015, $343 billion worth of products were sold online to Internet users in the United States, making up 7% of the country's retail product sales.[11] In Canada, online sales reached $22 billion in 2015, with over half of the products being ordered from retailers outside Canada.[12] The online retail market is large and is one of the most rapidly growing segments, with double-digit annual growth rates predicted for both the United States and Canada over the next five years.

Academics and practitioners have raised policy-relevant questions about the properties of this market and its relationship to state boundaries. Yet empirical academic papers about cross-border Internet trade are rare. Our paper is particularly well positioned to contribute to these debates.

The remainder of this paper proceeds as follows. In Section 2, we briefly review strands

---

[10]Similarly, our results about stronger border effects in regulated and/or locally consumed products may extend to offline sectors.

[11]The figures come from the US Department of Commerce's *Quarterly Retail E-commerce Sales Report (Q1 2016)*.

[12]The figures come from Statistics Canada and eMarketer's 2015 *Canada Retail E-commerce Forecast*.

of related literature and outline our contribution to them. Section 3 presents the data and its characteristics. Section 4 describes our empirical specification and estimation method. Section 5 contains the results on aggregated data for our main specification. Section 6 disaggregates the border effect into an effect on choice sets and an effect on consumer behavior conditional on choice sets. Section 7 examines the border effect in digital goods and services. Section 8 discusses the results when we disaggregate our data by economic sectors. Section 9 examines border effects in trade in final goods (retail) without intermediate inputs. Section 10 concludes.

2   Related Literature

Border effects have been studied by trade economists since McCallum (1995). The works of Anderson and van Wincoop (2003) and Feenstra (2002) contained major theoretical and methodological developments.[13] Several strands of the subsequent literature examine hypotheses for the existence of international as well as intranational, border effects. In various ways, our paper tests or comments upon these hypotheses. We discuss four below: information costs, trade barriers, multistage production, and aggregation bias. Then, we discuss online border coefficients that appear in in the statistical tables of related papers.

Information costs: Chen (2004) provides evidence of information costs creating a border effect, showing that the international border effect is larger in industries with a high degree of product differentiation – which face bigger information costs – compared to industries with homogenous goods.

In a related strand of research, several authors suggest that migrant and business networks reduce intranational border effects by facilitating information flow. The study by Combes et al. (2005) on trade between French regions found that the combined existence of business and migrant networks reduce the intranational border effect by 53%. Millimet and Osang (2007) showed that within the US, migrant networks dampen the intranational border effect, while Garmendia et al. (2012) reached similar conclusions by examining trade patterns within Spain.

---

[13]Although we do not cover it here, there is a related strand of the literature which examines the US-Canada border effect through the lens of price differences on either side of the border. See, for example, Gorodnichenko and Tesar (2009), Gopinath et al. (2011), Boivin et al. (2012), and Broda and Weinstein (2008).

Our paper contributes to the literature emphasizing the role of search frictions in reducing cross-border transactions by uncovering border effects in a setting optimized for buyers and sellers to discover each other easily (a search engine). Even in the face of low search frictions, we find strong border effects.

Trade barriers: A related literature examines the role of tariff and non-tariff barriers to cross-border trade, such as regulatory differences or standards requirements. However, researchers that investigate this explanation (see, for example, Head and Mayer, 2000; Chen, 2004 and Evans, 2003) find limited evidence to support this rationale for border effects.

Our findings shed additional light on the role that regulation plays in border effects. In particular, our sectoral results point to several cases where border effects are highest in regulated products.

We also contribute to the literature about the role of shipping costs in international trade.[14] In recent work, Yi (2010) calculated that the trade weighted cost of shipping a good between Canada and the United States, relative to the average cost of shipping a good internally, is 14.8%.[15] We contribute to this dialogue in Section 7 of our analysis by showing that software downloads, upgrades, and other virtual goods, which require no shipment costs, still encounter high border effects.

Multi-stage production: Yi (2010) examined the multi-stage production explanation of border effects, proposing a Ricardian model of trade similar to Eaton and Kortum (2002) to explain the international border effect between the United States and Canada. While Eaton and Kortum (2002) have single stage production, Yi's (2010) model has multistage, multi-region production. Yi (2010) then showed that in a world where different stages of the same good are produced in different regions, over two-fifths of the border effect between the US and Canada can be explained. By comparison, in a world with single

---

[14]This theory is particularly salient to online commerce, which often requires door-to-door delivery to consumers. Door-to-door delivery across borders may actually be more expensive per unit than traditional bulk shipments across borders to retail outlets, which would enlarge the online border effect. Furthermore, in some cases, firms can move goods (such as alcohol, tobacco, and other regulated products) across borders in ways that individual buyers cannot. Regulations may create border effects for consumers that might not exist for shipping companies.

[15]His trade cost calculation includes tariffs, non-tariff trade barriers, transport costs and wholesale distribution costs.

stage production, only one-sixth of the border effect can be explained.[16]

Rossi-Hansberg (2005) and Hillberry and Hummels (2008) also turned to multistage production processes in order to explain the border effect. However, unlike in Yi's (2010) model, the border effect in these two papers arises as a result of the spatial clustering of economic activity. Hillberry and Hummels (2008) showed that shipments within the United States are highly localized.[17] Hillberry and Hummels (2008) noted that for the US data that they have, 88% of the intranational border effect comes from the number of shipments, rather than the average value of each shipment. In other words, the extensive, rather than the intensive, margin drives the border effect.

Our findings also speak to the literature on multi-stage production and trade in intermediate goods. We contribute to this literature by analyzing a consumer-driven dataset, which covers final goods bought by individuals – without upstream trade in intermediate goods. We document strong border effects even in the setting of trade in final goods.

Aggregation bias: Finally, some researchers suggest that aggregating data over sectors of economic activity leads to biased estimates of the border effect. Hillberry (2002) showed that across industries, the US-Canada border effect varies from .03 for fur goods to 263.7 for gum and wood chemicals. The author argues that without disaggregating the data, researchers overestimate the border effect.

More recently, Anderson and Yotov (2010) focused on the manufacturing sectors and show that the aggregate US-Canada border effect is lower than the average border effect estimated with commodity level data. Anderson et al. (2014) underlined the need to estimate disaggregated gravity models for services as well. They show that the coefficients on the US-Canada border effect dummies vary significantly across nine service sectors.

We contribute to this strand of the literature by estimating the US-Canada border effect by sector of economic activity. Similar to previous researchers, we find that border effects vary widely across sectors. When we disaggregate our online data into NAICS2 industries, the US-Canada border affects the sector with the highest estimated border effect 7.4 times more than it affects the sector with the lowest estimated effect.

---

[16]The differences between these two results is due to the fact that in a multistage model, a given trade cost gives rise to a larger border effect than in a standard model. In Yi's (2010) model, goods end up crossing the border multiple times. As they do so, they incur the trade costs multiple times.

[17]In particular, they find that shipments within a Zip code are three times higher than shipments outside the Zip code.

Other studies of online border effects: A literature at the intersection of computer science, economics and quantitative marketing has studied economic geography and the Internet. This literature has primarily focused on consumers' substitution between online and offline purchasing channels (Lewis and Reiley, 2014; Forman et al., 2009; Brynjolfsson et al., 2009; Goolsbee, 2001).

A subset of this literature has focused on online trade across borders. In particular, three previous papers report estimates for border effects using online trade data: Hortaçsu et al. (2009), Einav et al. (2014), and Lendle et al. (2015). These papers mostly mention border effects in passing, but we report them here. Table I summarizes the intranational and international online border effects estimated by previous researchers.[18] Notably, prior to our paper, no other papers have studied the US-Canada border effect using online trade data.

Table I: Online Border Effects Obtained by Previous Researchers

Study	Intranat'l Border	Internat'l Border	Country Coverage	Estimator Used	Dependent Variable
Hortaçsu et al. (2009)	1.75	-	US	OLS	Number of sales (logs)
	2.75	431.82	9 South American states	OLS	Value of sales (logs)
Einav et al. (2014)	1.75	-	US	PPML	Number of sales
Lendle et al. (2015)		4.55	61 countries	OLS	Value of sales (logs)

Using eBay data, Hortaçsu et al. (2009) and Einav et al. (2014) find the same intranational border effect: all else being equal, trade within a US state is 1.75 times larger than trade outside the state. Hortaçsu et al. (2009) and Lendle et al. (2015) are the only researchers to estimate international border effects using online trade data. Their estimates are widely divergent.[19]

---

[18]The table reports the exponents of the estimated coefficients. The values reported in the table should be interpreted as the number of times that home trade exceeds out of state or out of country trade.

[19]Hortaçsu et al. (2009) contains some discussion of the possible rationales for these online border effects.

3 Data

The primary dataset in this paper contains online trade volumes between a) US states to US states (including own-state trade), b) Canadian provinces to Canadian provinces (including own-province trade) and c) US states to Canadian provinces during the years 2008 to 2011. We do not have data at the firm level or at a finer level of geographic disaggregation than state/province. In later sections of this paper, we analyze similar datasets consisting of subsamples of our larger dataset.

All data on trade volumes in this paper come from Google's online advertising platforms,[20] which track purchases by consumers who arrive through clicks on ads served by Google. To provide the clients of its advertising program with useful intelligence, Google offers free "conversion tracking" software. This software tracks users anonymously from viewing an ad to purchasing, placing the item in a shopping cart, downloading or other forms of online "conversions" that are valuable to businesses. It allows Google's clients to measure the return on their investment in advertising in terms of transaction counts. The transaction data we use in this study come from the records generated by the conversion tracking software. We describe a concrete example of how this software works with technical details in Appendix A.

The data in our paper come from a proprietary, commercial dataset. A natural limitation of such research data is that the figures come from a sample of buyers and sellers who self-select into the researcher's dataset (eBay, MercadoLibre, or in our case, Google). Google's large market share in North America on both the consumer and advertiser side somewhat alleviates this concern.

By contrast, the classical studies of trade (for example, McCallum, 1995) use government data with fewer (or no) issues of self-selection or representativeness. One major exception – relevant to the topic of this paper – is the study of subnational border effects. Despite the relevance of intrastate trade to the US Constitutional system,[21] the US government has few comprehensive, government data sources about trade among states.

---

[20]"AdWords," http://adwords.google.com

[21]Much of the growth of the size and scope of the US federal government during the 20th century was legally supported by the "Interstate Commerce Clause" of the US Constitution (Article 1, Section 8, Clause 3), which gives Congress regulatory authority over commerce "among the several States." The Supreme Court has interpreted this clause to give Congress broad authority to regulate the economy and advance civil rights.

Thus, one of the contributions of this paper is to provide estimates on intranational border effects in the US using a new dataset with attractive sample properties. Lacking a more comprehensive dataset, the incumbent literature on intranational borders utilizes the US Commodity Flow Survey, a dataset developed by the US Department of Transportation (DOT) that focuses on the movement of commodities and business-to-business transactions such as mining, manufacturing and wholesale distribution. The DOT states that "most retail and services industries are excluded from the survey."[22]

Issues of representativeness affect our study as well – we have no data about purchases that that do not involve the consumer clicking an ad served by Google at some point in her search and checkout purchase.[23] Similarly, we have no data about transactions from sellers who are not AdWords advertisers.[24]

Nonetheless, our data are much larger, more representative and diverse than previous studies focusing on online trade. This is in part because of Google's large market share in North America on both the consumer and advertising side (unlike eBay), and because of Google's status as a generic, all-purpose search engine for a wide variety of online goods and services – which also differentiates it from platforms like eBay. The overall number of transactions in our data is in excess of 10 billion, conducted by several million online sellers.

Given that our data come from an online ads platform, a few details are worth clarifying. First, Google permits Canadian and American advertisers to advertise equally on both sides of the border. The search engine's advertising placement algorithm does not give advantages to firms based on geographic location of the advertiser.

In our data, buyers' locations are determined by IP address, and sellers' locations are determined by the sellers' addresses given to Google for correspondence. Additional details about the geolocation coding in our data are discussed in Appendix B.

---

[22]http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/methodology_2012.html, accessed September 13, 2016.

[23]This includes visits directly to the seller's homepage. For purchases made after an ad click, our transaction data is limited to purchases made 30 days after the initial click of an ad.

[24]As previously noted, most transactions on this platform are with consumers – there is little business-to-business transactions online (especially in our data).

3.1 Variables in the Data

The unit of analysis in our data is a pair of subnational geographies for which we report trade flows. Our data includes all pairs of US states (including own-state trade), all pairs of Canadian provinces (including own-province trade) and all pairs of US states and Canadian provinces. As a reminder to the reader, a conversion is a transaction undertaken by the user on an advertisers website. Positive conversion counts are reported for advertisers and sellers in all 50 states, the District of Columbia (DC), the 10 Canadian provinces and the 3 Canadian territories. As with previous offline studies, our main specification excludes the three northern Canadian territories: Yukon, Nunavut, and the Northwest Territories.[25] The territories are very scarcely populated, comprising just 0.30% of Canada's population. In the US, we focus our analysis on the 48 continental states and DC. The data cover all conversions recorded over a period of four years, from 2008 to 2011.

Our data report transaction volumes (counts), and not prices.[26] The data include a rich set of covariates describing the sector of economic activity and details about the type of conversion or product purchased.

Besides conversion numbers, we are also able to obtain some data from Google regarding the number of ad clicks. The figures we obtain represent the number of times users on Google or Google's partner sites clicked on online ads placed by Canadian and American businesses. Positive click counts are reported for each US state and Canadian province. The click data cover a period of three years, from 2008 to 2010.[27] Click data are available only at the aggregate level; unlike conversion counts, we are unable to break down the figures by conversion type or sector of economic activity. Even so, we are able to use the aggregate click counts in Section 6 to separate the border effect in choice sets from the border effect in final purchase decisions.

For our distance measure d_{ij}, we follow other papers in this literature by measuring the distance between region i and region j in kilometers using the Great Circle Distance

---

[25]Similar to other researchers who examine online trade flows (for example, Lendle et al., 2015), we collapse this panel into a single cross-section by averaging across these four years.

[26]This is because of the way the data is collected – these counts represent the number of times users visited the sections of the seller's websites designated as post-transaction or post-conversion pages.

[27]Conversion data cover a period of four years, from 2008 to 2011.

Formula.[28] For same state or same province trade, we calculate the internal distance using the methodology proposed in Head and Mayer (2002).[29]

Conversion types

Online conversions can constitute a variety of advertiser-specified events. Some advertisers choose to measure conversions that do not immediately lead to money changing hands. For example, they may track mailing list signups or page views. Although they may not lead immediately to purchases, we believe all conversions tracked have economic value to the advertisers.[30]

When setting up tracking for their accounts, Google's advertisers can identify the types of conversions being tracked. This selection falls into one of the categories listed in Table II below.

Table II: Conversion Types

Conversion type	Conversion is due to:
Purchase	A purchase, sale, or "order placed" event
Sign-up	A sign-up user action
Lead	A lead-generating action
Page view	A user visiting a page
Login	A user logging into a pre-existing account
Shopping cart post	An item put into a shopping cart
Order charged	A purchase or order that was successfully charged for
Install	A software install action
Download	A software download action
Referral	A referral of new customers

---

[28] All data on latitude and longitude come from the World Gazetteer web page.

[29] Thus, we calculate the internal distance of region i using the formula $d_{ii} = 0.67\sqrt{\text{internal area}_i/\pi}$. Data on the internal area of the US states come from the United States Census Bureau, while data on the internal area of the Canadian provinces and territories come from Statistics Canada.

[30] If they did not, advertisers would not spend money to promote their sites on Google's platform in order to generate these conversions.

We make use of this labeling data in Section 7 to isolate purchases where no physical shipment was involved. Otherwise, all regressions in this paper use data pooled across all conversion types.

Sector of economic activity

Google has automatic methods to classify ads running on its platforms into categories called "verticals." These verticals are similar to NAICS sectors and they follow the same structure. Our data includes 27 top-level vertical categories, which are similar to the 2-digit NAICS sectors. These are the most generic classifications available. Under these top level verticals, there are 241 second- and third- level categories, which are similar to the 4-digit and 6-digit NAICS classifications. An algorithm assigns each Google search to the relevant verticals and subverticals. [31]

To map Google's verticals to NAICS sectors, we use a classification scheme designed by Google Chief Economist Hal Varian in Choi and Varian (2012).[32] A few examples of how the mapping is done are provided in Table III below.

Table III: Google Vertical to NAICS Sector Mapping

Google Vertical		NAICS Sector	
ID	Title	ID	Title
47	Automotive	441	Motor vehicle and parts dealers
5	Computers and Electronics	443	Electronics and appliance stores
1868	Apparel	448	Clothing and clothing accessories stores
...	...	...	...

We successfully assign a 2-digit NAICS classification to over 99% of the conversions in our data set. Only 0.2% of our data remain unclassified. Out of the classified conversions, 1.5% of them fall within the Agriculture, Mining, Construction and Manufacturing sec-

---

[31]For example, a search for [car tires] would be classified under the third level category 'Vehicle Tires,' the second level category 'Auto Parts' and the top level vertical 'Automotive.'

[32]Where Choi and Varian (2012) did not provide a NAICS/vertical mapping, we assigned an encoding.

tors,[33] leaving the vast majority of the classified conversions (over 98%) reflected within one of the service sectors.[34] Over 30% of the conversions reflected in our data set fall within "Retail trade." Due to the large number of conversions in this category, we are able to classify retail trade conversions at the NAICS3 level and analyze these retail subsectors in Section 9.

3.2 Comparability to Border Effect Estimates Obtained Using Government Data Sources

Traditionally, studies that examine the US-Canada border effect (for example, McCallum, 1995) or intranational border effects (for example, Wolf, 2000) use government data. The estimates these studies report are not directly comparable to ours. As noted in Section 3.1, over 98% of the conversions in our dataset are reflected within one of the service sectors, with more than 30% of all conversions reflected in the two-digit NAICS services sector "Retail trade."

Unfortunately, government data sources do not collect trade flow figures between US states and Canadian provinces for services. For trade flows between US states and Canadian provinces, Statistics Canada and the US Census Bureau provide figures only for the following two-digit NAICS sectors:

11: Agriculture, forestry, fishing and hunting.
21: Mining, quarrying, and oil and gas extraction.
22: Utilities, 23: Construction.
31-33: Manufacturing.

No data are available for total trade in Services or for any of the two-digit NAICS service sectors.

---

[33]The underrepresentation of Manufacturing warrants a brief explanation. Google has a vertical for Industries, the company's equivalent category to what trade economists would call manufacturing. Less than 0.2% of the conversions in our data set are classified under Industries. Google also has verticals such as 'Computers and electronics.' We can classify this vertical as NAICS 334, Computer and electronic product manufacturing, or as NAICS 443, Electronics and appliance stores. The first classification would fall under Manufacturing, the second one would fall under Services. We follow Choi and Varian (2012) and classify 'Computers and electronics' within 'Services,' under 'Retail trade.' The reason they suggest this classification and we choose to follow it, is because retailers, rather than manufacturers, are Google's usual advertising clients.

[34]The only service sector that we cannot map any conversions to is "Management of Companies and Enterprises."

Data for most service sectors are also unavailable at the US state to US state level (see discussion of the CFS in the beginning of Section 3). Given the discrepancy between our data and traditional data sources' coverage of the service sectors, our estimates should be seen as complementing – rather than providing a direct comparison to – the estimates reported by researchers who examined the US-Canada border effect using traditional, offline data figures.

4   Empirical Estimation Details

4.1   Specification

Using the theoretical framework proposed by Anderson and van Wincoop (2003), the following gravity equation can be easily derived:

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

where:

x_{ij} = region i's exports to region j

y_i = GDP of region i

y_j = GDP of region j

y_w = world GDP

t_{ij} = transport costs between i and j.

P_i = consumer price index of region i, also called a multilateral resistance term

P_j = consumer price index of region j

σ = elasticity of substitution between all goods

Taking the log of each side of equation 1, we obtain:

$$\ln x_{ij} = \ln y_i + \ln y_j - \ln y_w + (1-\sigma) \ln t_{ij} - (1-\sigma) \ln P_i - (1-\sigma) \ln P_j \tag{2}$$

16

The transport costs, t_{ij}, are hypothesized to be a loglinear function of the bilateral distance between the two regions, d_{ij}, as well as all the observable intranational and international border barriers, $b_{ij}(h)$, where h indexes these barriers.

$$\ln t_{ij} = \rho \ln d_{ij} + \sum_h \gamma_h b_{ij}(h) \tag{3}$$

We use equation 3 to substitute for $\ln t_{ij}$ in equation 2[35] and include importer and exporter fixed effects.[36] This produces the following operational econometric model:

$$\begin{aligned}
\ln x_{ij} = \beta_0 &+ \beta_1 \ln d_{ij} \\
&+ \beta_2 \text{ contig US-CA}_{ij} + \beta_3 \text{ contig CA}_{ij} + \beta_4 \text{ contig US}_{ij} + \beta_5 \text{ internal}_{ij} \\
&+ \beta_6 \text{ border US} \leftrightarrow \text{CA}_{ij} + \lambda_i + \chi_j + \varepsilon_{ij}
\end{aligned} \tag{4}$$

where $Contig\ US-CA_{ij}$, $Contig\ CA_{ij}$, $Contig\ US_{ij}$, $Internal_{ij}$, and $Border\ US \leftrightarrow CA_{ij}$ capture the observable intranational and international border barriers. We define the following dummy variables:

Contig US-CA: 1 if a Canadian province and a US state share a land border, 0 otherwise.
Contig CA: 1 if two Canadian provinces share a land border, 0 otherwise.[37]
Contig US: 1 if two US states share a land border, 0 otherwise.
Internal: 1 for a Canadian province or a US state's trade with itself, 0 otherwise. This permits us to estimate a home market effect.[38]

---

[35]Note that $\ln y_w$ is a constant term

[36]These fixed effects capture the effects of the regions' GDPs and unobserved multilateral resistance terms. They also pick up other state- or province-specific unobservables. For example, Quebec is home to 6.3 out of the total of 7.3 million native francophones in Canada. Among all the states and provinces in this study, it is the only region where English is not an official language and where the population is predominantly French speaking. Being the only region with these linguistic characteristics, the fixed effects pick up this individual variation.

[37]Our specification allows for asymmetric contiguity effects. Brown and Anderson (2002) hypothesize that contiguity between two US states and contiguity between a province and a state might have a different effect on trade. Their regressions are on a per-industry basis. They find that contiguity between two US states has a positive and significant effect on trade for all industries considered, while contiguity between a US state and a Canadian province does not have a statistically significant effect for every industry. Examining this issue from a Canadian prospective, Anderson and Yotov (2010) find that at the aggregate level, contiguity between two Canadian provinces does not have a statistically significant effect on trade, while contiguity between a Canadian province and a US state has a positive and significant effect. Although the literature has not reached a consensus on the direction of these effects, the results of the papers mentioned above point to the importance of allowing contiguity to affect trade asymmetrically.

[38]Including same-state trade and dummies has a clear advantage. Hillberry and Hummels (2008) argue

Border US↔CA: 1 for exports between states and provinces and 0 otherwise.

In the econometric model, λ_i and χ_j are exporter and importer fixed effects, respectively, and ε_{ij} is the error term.

Note that in equation 4, $\beta_1 = (1 - \sigma) \times \rho$. In other words, the coefficient in front of the bilateral distance variable is the product of the elasticity of substitution, $(1 - \sigma)$, and the elasticity of total trade costs with respect to distance (ρ). Similarly, the coefficients in front of the observable intranational and international trade barriers variables (β_2, β_3, β_4, β_5, and β_6) are also the product of the elasticity of substitution, $(1 - \sigma)$, and the elasticity of total trade costs with respect to each trade barrier variable. This is due to the fact that in equation 2, $\ln t_{ij}$ is preceded by $(1 - \sigma)$.

4.2   Estimation Methods

The prevalence of zero trade flows in the offline world gave rise to a substantial literature advocating against the use of linear estimators, which ignore the zeros in the data, in gravity exercises. Over the past decade, researchers such as Helpman et al. (2008) and Silva and Tenreyro (2006) proposed alternative estimation methods that account for the missing flows.

A particularity of our data set, however, is that we have very few missing trade flows. Our aggregated US-Canada data contain no zeros at all: all states and provinces export to all locations. When we disaggregate our data by sector of activity, we do encounter some zero trade flows. Even so, every industry we consider features at least 45 states and provinces that export to all locations.

The low incidence of zero trade flows in our data affects our choice of estimation methods. We follow Head and Mayer's (2014) recommendation and choose an appropriate estimator for our US-Canada border effect analysis from the OLS, the Poisson Pseudo

---

that trade is extremely localized and high national border effects might be due to researchers failing to control for this. In the absence of finer distance measures, as is the case here, one way to ensure that extremely localized trade does not drive the high national border effect is by including same state trade and same state dummy variables. We were unable to obtain more geographically granular data from Google at the time of writing.

Maximum Likelihood (PPML),[39] and the Gamma Pseudo Maximum likelihood (GPML) estimators.[40] Even in the absence of zero trade flows, the PML estimators have a potential advantage over standard OLS: they do a better job at handling heteroskedasticity.

We perform the RESET test proposed by Silva and Tenreyro (2006) to ascertain whether the OLS or the two PML estimators do a better job at addressing potential heteroskedasticity problems in our data.[41] We find that the PPML estimator has the best performance in the RESET test out of the three estimators considered, so we use the PPML estimator for our analysis. For comparison purposes, we also provide the results we obtain by using OLS for our main specification.

5 Aggregate Results

5.1 Main Specification

Table IV provides aggregated results for our main specification described in Section 4. We include estimates from both PPML and OLS for comparison. We also include results including non-continental US states and the Northern Canadian territories. If we estimate our regressions with all 50 US states, DC, the 10 Canadian provinces, and the 3 Canadian territories, our results are virtually unchanged. All regressions have importer and exporter fixed effects. We report robust standard errors, clustered by region pair.

We highlight five main results from Table IV. First, the p-values for the RESET test (in the last row of Table IV) clearly show the errors of the OLS model are heteroskedastic. By contrast, the p-values for the RESET test for the PPML model show that the regressions estimated using this method are adequate.

---

[39]Suggested by Silva and Tenreyro (2006).

[40]Note that we cannot employ a methodology similar to that suggested by Helpman et al. (2008), as it relies on a first stage regression that models the probability of two locations trading. With a data set where there are many zero trade flows, estimating a first stage Probit regression is a reasonable proposition. With our data set, however, estimating a first stage Probit regression is not feasible: in the absence of zero trade flows, the probability of exporting is always 1.

[41]The RESET test (Ramsey Regression Equation Specification Error Test, Ramsey, 1969) is in essence a test for the correct specification of the conditional expectation. It is a special case of White's test for heteroskedasticity. The test is conducted by examining the significance of an additional regressor, $x'\hat{\beta}$, where $\hat{\beta}$ is the vector of estimated parameters.

Second: The US/CA border coefficients indicate a strong border effect in e-commerce. In an Internet setting, where information costs and business-to-business transactions in intermediate inputs are largely absent, this is a surprising finding. The coefficient of the US↔CA dummy variable in column (2) of Table IV is large, negative and statistically significant. Taking the exponent of this coefficient gives an average border effect of 6.7, which indicates that online trade between two US states or two Canadian provinces is 6.7 times higher than trade between a US state and a Canadian province.

Table IV: Results at the Aggregate Level for the Main Specification

	PPML			OLS
	48 US states, DC, 10 CA provinces		50 US states, DC, 10 CA provinces, 3 CA territories	48 US states, DC, 10 CA provinces
	(1)	(2)	(3)	(4)
$\ln d_{ij}$	-0.019	0.037	0.032	-0.152***
	(0.016)	(0.024)	(0.023)	(0.022)
contig US-CA$_{ij}$		0.223**	0.218**	0.169**
		(0.106)	(0.105)	(0.085)
contig CA$_{ij}$		0.028	0.021	-0.242
		(0.194)	(0.194)	(0.306)
contig US$_{ij}$		0.219***	0.212***	0.086**
		(0.042)	(0.041)	(0.037)
internal$_{ij}$	0.555***	0.706***	0.696***	1.261***
	(0.087)	(0.095)	(0.094)	(0.159)
border US ↔ CA$_{ij}$	-1.912***	-1.903***	-1.903***	-1.858***
	(0.067)	(0.080)	(0.080)	(0.097)
Implied US↔CA Border Effect	6.76	6.71	6.71	6.41
Observations	3,481	3,481	4,096	3,481
RESET p-values	0.134	0.136	0.113	0.000

 Importer and exporter fixed effects included. The dependent variable is the level of conversion counts aggregated across all conversion types and sectors of economic activity for the PPML specification and the log of conversion counts similarly aggregated for OLS. The data cover four years, 2008 to 2011. Robust standard errors (clustering by region pair).
 *** Significant at 1%, ** significant at 5%, * significant at 10%

Third, we find no statistically detectable relationship between trade and distance in this

dataset. We discuss this result further in Section 5.2.

Fourth: We find a positive "home state bias" in online trade. The coefficient of the *Internal* dummy variable in column (2) of Table IV is economically and statistically significant.[42] The coefficient implies that trade within a state or a province is 2.03 times larger than intranational trade.

Lastly: Contiguity does not appear to have a large influence in online trade. The coefficients of the dummy variables measuring the effect of sharing a land border between a US state and a Canadian province, as well as between two US states, are positive, but their magnitudes are small.[43]

5.2 Discussion of Distance Coefficients

There are several reasons why distance might matter less online. In particular, the information frictions that hinder trade in the offline world are mostly absent online. In the virtual world, search costs are close to zero: buyers and sellers find each other with a simple internet search, the cost of which is low and uncorrelated with the distance between the transacting parties. Communication costs online are also close to zero and independent of distance: buyers and sellers can easily communicate via email or web-based forms. Finally, contracting costs are also likely to be lower online, with transacting parties generally able to rate each other's behavior after a sale.

Previous researchers that examined online trade flows in a North American context found that, indeed, distance plays a smaller role online. Authors like Einav et al. (2014) and Hortaçsu et al. (2009) used data on the number of eBay transactions conducted within the United States to estimate distance coefficients of -0.10, and -0.07 respectively.[44]

---

[42]This finding agrees with previous work. Einav et al. (2014) report a coefficient of 0.56 for their same state dummy. Hortaçsu et al. (2009) also report a coefficient of 0.56 when studying the number of eBay sales within the United States. At 0.71, our estimated coefficient is slightly higher.

[43]Our results imply that trade between two contiguous US states or a contiguous US state and Canadian province is approximately 25% higher than trade between non-contiguous locations. The coefficient of the dummy variable measuring the effect of sharing a land border between two Canadian provinces is not significant at any conventional level. These contiguity results are similar to the ones reported by Anderson and Yotov (2010), who find that contiguity has a positive and significant effect on trade between US states and Canadian provinces. Like us, they also find that sharing a land border makes no difference for trade between Canadian provinces.

[44]Like us, Einav et al. (2014) use the PPML estimator. Hortaçsu et al. (2009), on the other hand, estimate

These coefficients are lower than those traditionally found in offline data, but are significantly different from zero.

By contrast, we find no statistically significant results of distance on online conversions in our PPML estimations in Table IV. The differences between our estimates and those reported by Einav et al. (2014) and Hortaçsu et al. (2009) might be due to the fact that our datasets cover different sets of transactions. The previous studies use data from eBay, which is a website dedicated to e-commerce transactions. As discussed earlier in this paper, our data cover much wider variety of economic activity.

The differences we notice between our estimates and those reported by Einav et al. (2014) and Hortaçsu et al. (2009) may also be due to the fact that the average service or good traded on Google and its partner sites is easier to provide at a distance than the average product sold on eBay. For example, while many eBay sellers must physically ship the product sold to the buyer, some service providers on Google or Google's partner sites might be able to transact electronically with their clients, without needing any physical shipments.

Note that in column (4) of Table IV, where we report the estimates we obtain by using the OLS, rather than the PPML estimator, the coefficient on distance is negative and statistically significant. The difference between the results we obtain by using PPML and the results we obtain by using OLS are likely due to heteroskedasticity. With a p-value of zero, the RESET test we perform for the OLS estimation clearly rejects the null hypothesis of homoskedastic errors. In the presence of heteroskedastic errors, Silva and Tenreyro (2006) conclusively show that a log linear model cannot be expected to produce consistent estimates of the parameters of interest. By contrast, the PPML specifications in Table IV all pass the RESET test.

6 The Border Effect in Ad Clicks versus Purchases

Before we observe a transaction in our data, the buyers and sellers must take several strategic steps. A buyer must have expressed demand or interest by typing a search term (also known as a keyword) into Google. The seller must have created an ad campaign that

---

their regressions using OLS. Both papers include in their estimations same-state trade.

responds to a buyer typing this keyword from the buyer's country. Then, the buyer must click on the ad, visit the website and make a purchase.

A border effect may arise at any stage in the purchase funnel – keyword searches, displays of ads, clicks on ads (visits to seller websites) and final purchases. Up to now in this paper, we have examined the US-Canada border effect only in final purchases. Most of the incumbent border effects literature similarly studies final purchases (unconditional on prior actions).

Conventional border effects could arise as a result of differences at the very top of the purchase funnel: Domestic consumers might very rarely search for keywords describing foreign products. While we express this idea in terms of online search, the concept has an offline equivalent: A border effect may arise offline if consumers simply have no demand or interest for foreign goods at all.

Alternatively, buyers may have demand for foreign goods but foreign sellers may refuse to meet it (perhaps because of costs of serving faraway markets). A conventional border effect may arise if consumers demand foreign products – but foreign sellers refuse to meet the demand. Servicing foreign demand may be undesirable for businesses for many well-documented reasons. Exporting often has extra financial, logistic, customer-support and regulatory hurdles. These costs may also generate a border effect.

In this section, we examine a setting where both of these explanations are off the table. We study consumer purchase behavior after a click on an ad. At this point in the purchase process, the consumer has already expressed interest in the seller's product by entering a related query. The seller has expressed interest by permitting an advertisement to be shown on that keyword to the user's location.[45] The consumer has then re-affirmed interest by clicking on the ad.

We begin by examining the border effects in ad clicks – the event immediately preceding a purchase in our data.[46] A priori, we expect to find the US-Canada border to have a dampening effect on ad clicks because of the aforementioned selection. For an ad click to occur, three things must have already taken place:

---

[45]A user cannot see an ad unless the advertiser that placed the ad agreed to target the user's geographic location.

[46]Our dataset does not contain variables for queries and ad impressions that did not result in a purchase, but we do know the number of clicks on ads that fail to result in a purchase.

1. First, a consumer must have typed in a relevant keyword. There may be a border effect in choices of keywords and the underlying product being requested.

2. Second, a seller must have arranged an advertisement to be displayed to users at the consumer's location when this keyword was typed. A border effect may arise in firms' choices of where to advertise.[47]

3. Third, a consumer must click on the ad. The choices of ad clicking may also be influenced by the borders.[48]

In column (1) of Table V, we run our specification with the number of clicks as the dependent variable and obtain a negative and statistically significant coefficient on the US↔CA border dummy. Taking the exponent of this coefficient gives a border effect of 4.06, which indicates that the number of clicks that users and businesses located within the same country generate are 4.06 times higher than the number of clicks generated by users and businesses located on opposite sides of the US-Canada border.

In columns (2) and (3) of Table V, we decompose the total number of clicks into clicks that do not convert and clicks that do convert for the advertisers. In column (2), the dependent variable is the difference between clicks and conversions. This variable measures the number of times that users clicked on ads but did not purchase anything. In column (3), the dependent variable is conversions.[49] This variable measures the number of times that users clicked on ads and purchased.

The results reported in columns (2) and (3) show a discrepancy between the coefficient estimated for the US↔CA border dummy when using the difference between clicks and conversions as the dependent variable and the coefficient estimated when using conversions on the left hand side. In particular, the US-Canada border seems to have a more severe dampening effect on conversions than on clicks that do not lead to conversions.

---

[47] A firm may choose to conduct business near home because of an easier or more familiar environment for regulation, shipping or customer-support. Alternatively, the firm may prefer to leverage pre-existing infrastructure already built near its home.

[48] For example, some ad text may explicitly reference the firm's location.

[49] Table V's results on conversions are estimated from 2008-2010 only because we were not able to obtain click outcomes for 2011.

Table V: Results for Clicks versus Conversions

	Clicks	Clicks minus Conversions	Conversions	Conversion to click ratio
	(1)	(2)	(3)	(4)
$\ln d_{ij}$	-0.001	-0.006	0.035	-0.002
	(0.013)	(0.012)	(0.023)	(0.002)
contig US-CA$_{ij}$	-0.009	-0.024	0.079	0.005
	(0.068)	(0.071)	(0.111)	(0.006)
contig CA$_{ij}$	0.002	0.018	-0.044	0.011
	(0.215)	(0.219)	(0.232)	(0.030)
contig US$_{ij}$	0.120***	0.107***	0.231***	0.003
	(0.023)	(0.022)	(0.043)	(0.003)
internal$_{ij}$	0.503***	0.484***	0.686***	0.014
	(0.057)	(0.055)	(0.099)	(0.012)
border US↔CA$_{ij}$	-1.400***	-1.347***	-1.805***	-0.023***
	(0.069)	(0.071)	(0.091)	(0.006)
Implied US↔CA Border Effect	4.06	3.85	6.08	n/a[†]
Observations	3,481	3,481	3,481	3,481

Importer and exporter fixed effects included. The dependent variable is the number of clicks in column (1), the difference between the number of clicks and the number of conversions in column (2), the number of conversions in column (3), and the ratio of conversions to clicks in column (4). All dependent variables are aggregated over all sectors of economic activity and over all conversion types and they cover three years, 2008 to 2010. Estimations use PPML in columns (1) to (3) and fractional Probit in column (4). The coefficients reported in column (4) represent average marginal effects. Robust standard errors (clustering by region pair).

[†] Not available, as the estimation is performed using a fractional Probit.

*** Significant at 1%, ** significant at 5%, * significant at 10%

Because of the sequence of the purchase and advertising decisions, we interpret a de-clined purchase as the buyer's choice. If a seller were reluctant to transact with a foreign buyer, she would have not paid for her products to be advertised abroad.[50] The reluctance

[50]As an example, suppose we have a California-based bicycle seller that places ads in Ontario. Now sup-pose that following an ad click, no transaction takes place between the seller and the user. This can happen if the seller is unwilling to sell to an Ontario-based user or if the buyer is unwilling to buy from a California-based seller. Because of the sequence of the purchase process, we interpret a declined purchase as the buyer's choice. By making the ad visible within Ontario and by paying for the ad click, the bicycle seller has already signaled her willingness and ability to trade with the Ontario-based user. The user, however, upon clicking on the ad, might decide that she is unwilling to trade with a California-based seller.

can happen, for example, if the user learns that the seller offers limited warranty coverage to foreign customers or if the seller is unable to offer next-day deliveries to international addresses.[51]

We further investigate consumers' reluctance to trade with foreign sellers in column (4). In this column, the dependent variable is the ratio of conversions to clicks. Since this variable lives on the [0,1] interval, we estimate our regression using a fractional Probit. The reported coefficients represent average marginal effects.

If consumers are reluctant to trade with foreign sellers (even after clicking), we would expect to see a negative and significant US-Canada border effect in column (4). This is, indeed, the case. The estimated coefficient for the US↔CA border dummy is -0.023, and it is significant at the 1% level. This indicates that an ad click has a lower probability to convert into a sale when the transacting parties are not located within the same country. Interestingly, we do not notice the same bias for same-state transactions. The coefficient on the same-state dummy variable is insignificant. Our results show that the only border that has a negative impact on the conversion to click ratio is the international border. This is the first indication we have that consumers' reluctance to trade with foreign sellers might be one of the factors behind the US-Canada border effects observed in the data.

While our results are obtained from an online search setting, they have a more general offline interpretation: Much of the border effect appears well before consumers are faced with a menu of concrete choices. For example: In an offline setting, a "click" might be conceived as a customer walking into a storefront to inquire (and possibly purchase) a physical product. For this to take place, the product must be available in the consumer's country, and the consumer must be willing to appear at the store to consider the product. Our estimates are akin to separating the border effect of appearing at stores from the border effect of purchases after appearing at stores.

To our knowledge, our paper is the first to provide separate quantitative estimates of these two types of border effects. We show that most of the border effect appears as part of process of two-sided strategic search. The border mostly affects availability – that is, on buyers' menus and choice sets. The border effect on choices, conditional on equilibrium menus, still exists but is smaller.

---

[51] Buyers may also see a higher price or shipping costs from foreign sellers.

The ratio we present is taken from an equilibrium. From these data, we cannot say how the ratio would change if the equilibrium were perturbed. It is possible that the seller's upstream entry decisions are completely driven by expectations about consumer's underlying preferences downstream. In particular, if equilibrium menus expanded for some reason – for example, a drastic drop in advertising prices – the border effect conditional on those menus might rise.

While our online data enables us to quantify these ratios, our results could be biased estimates of their offline counterparts. In particular, the online environment allows firms to enter export markets relatively easily because no physical presence is necessary. Because entry would be more difficult offline, we may expect the border effect in "menus" to be higher in offline data where physical presence may be necessary.

7   Downloads, Digital Goods, and Transactions without Shipment

Our findings in Section 5 show a large and significant US-Canada border effect in e-commerce in aggregate. In this section, we use the data we have on purchase types to generate insights into the factors that give rise to this border effect. Yi (2010) calculates that the trade weighted cost of shipping a good between Canada and the United States, relative to shipping a good internally, is 14.5%. While we cannot directly calculate the cost of shipping goods for our data, we can use our data to understand how much the difficulty of shipping goods across the US-Canada border is the factor that drives the border effect.

As noted in Section 3.1, Google records ten different types of conversions. Only three of them describe a sale where a product physically changes hands between a buyer and seller. These three types are *Purchase* (a purchase, sale or "order placed" event), *Order charged* (a purchase or order that was successfully charged for), and *Shopping cart post* (an item put into a shopping cart). The other seven conversion types (*Page view*, *Install*, *Download*, *Login*, *Signup*, *Referral*, *Lead*) involve an online event only, where no physical good is being shipped from the buyer to the seller.

Table VI: Results for Conversions Where No Product is Shipped

	All Conversions	Conversions Where No Product Is Shipped
	(1)	(2)
$\ln d_{ij}$	0.037	0.033
	(0.024)	(0.026)
contig US-CA$_{ij}$	0.223**	0.173**
	(0.106)	(0.085)
contig CA$_{ij}$	0.028	-0.009
	(0.194)	(0.158)
contig US$_{ij}$	0.219***	0.229***
	(0.042)	(0.046)
internal$_{ij}$	0.706***	0.717***
	(0.095)	(0.107)
border US $\leftrightarrow$ CA$_{ij}$	-1.903***	-2.071***
	(0.080)	(0.063)
Implied US↔CA Border Effect	6.71	7.93
Observations	3,481	3,481

 Importer and exporter fixed effects included. The dependent variable is the levels of conversion counts aggregated over all sectors of economic activity and over all conversion types (column 1) and aggregated over all sectors of economic activity and the 7 conversion types that refer purely to online events (column 2). The data cover four years, 2008 to 2011. Estimations use PPML. Robust standard errors (clustering by region pair).
 *** Significant at 1%, ** significant at 5%, * significant at 10%

In column (2) of Table VI, we report the results we obtain by limiting our data set to the conversion types that refer to online events only. In column (1) of Table VI, we replicate from Table IV the results we obtain when using our full data set for comparison. If the difficulty of shipping a good across the US-Canada border is one of the factors which drive the international border effect in our estimations, the reported coefficient on the US↔CA border dummy should be smaller for the subset of our data that excludes the 3 types of conversion which might describe goods being physically shipped.

For the subset of the data that describes online events only, the US-Canada border effect appears to be even higher than the border effect for all conversion types. The coefficient

for the US↔CA border dummy in column (2) of Table VI is -2.071, larger in magitude than the one we report in column (1) of the same table for all conversion types.[52]

The comparison between the two estimated US↔CA border effects in Table VI shows that when no product is physically shipped, the border effect is equally large as for all conversion types. This result suggets that the difficulty of shipping goods across the US-Canada border does not explain the large international border effect that we find in the data.

8   Sector Level Variation in Border Coefficients

Several other researchers suggest that aggregating data over sectors of economic activity leads to biased estimates for the border effect (Hillberry, 2002; Anderson and Yotov, 2010; Anderson et al., 2014). In this section, we use subsets of our data to estimate per-sector border coefficients.

We find widely varying effects. In our data, the lowest coefficient that we find for the US↔CA dummy variable is -1.278, while the highest is -3.282. This large difference points to the importance of using disaggregated data when studying border effects in a gravity setting, and provides justification for earlier suggestions to use disaggregated data.

We summarize our sectoral level results in Table VII.[53] A pattern that emerges in our results is that border coefficients are largest for services that need to be consumed in a particular location.[54] As discussed in Section 4, we cannot disentangle the effect of the elasticity of substitution from the effect of the elasticity of the total trade costs with respect to each trade barrier.[55] Nonetheless, the ranking of the sectors in Table VII provides some

---

[52]A t-test reveals that the difference between these two coefficients is not statistically significant.

[53]The complete set of results, which reflect the values of all coefficients included in our regressions and their corresponding standard errors, is presented in Appendix C.

[54]This is a hypothesis proposed by Hortaçsu et al. (2009), as well, to explain high levels of same city sales. For similar reasons, many of these sectors also have high intranational border effects, also reported in Table VII.

[55]As noted in Section 4, the estimated coefficients are the product of the elasticity of substitution, $(1 - \sigma)$, and the elasticity of total trade costs with respect to each trade barrier variable. Prior research provides substantial evidence that σ, the elasticity of substitution, varies greatly across industries. See, for example, Broda and Weinstein (2006), Chen and Novy (2011) or Imbs and Mejean (2015). Since we cannot assume that σ is constant across industries, we cannot attribute the differences between the coefficients reported in Table VII solely to differences in trade frictions across sectors. In other words, if we estimate a smaller coefficient

evidence that locally-consumed services face larger border effects as we explain below.

Table VII: Results at the NAICS2 Level

NAICS2 Sector:	$\ln d_{ij}$	$internal_{ij}$ (exponent)	$border\ US \leftrightarrow CA_{ij}$ (exponent)
	(1)	(2)	(3)
Administrative, support, and waste management	0.056º	2.01	26.63
Public administration	-0.117	1.74º	21.48
Real estate and rental and leasing	-0.221	1.48º	19.83
Health care and social assistance	0.102º	2.47	13.21
Accommodation and food services	-0.119º	2.74	12.88
Transportation and warehousing	-0.039º	1.70	10.86
Finance and insurance	0.023º	1.85	9.78
Education services	0.057º	3.17	8.92
Retail trade	0.008º	1.67	6.20
Information and cultural industries	0.059º	2.02	5.04
Arts, entertainment, and recreation	0.064º	6.22	4.90
Wholesale trade	-0.070	1.37	4.59
Other services	0.007º	1.97	4.18
Professional, scientific, and technical services	-0.102	1.31	4.05
Utilities	-0.124º	4.21	3.59

The first column presents the coefficient for distance obtained by applying our main specification to each NAICS2 sector in our data set. The second column reports the exponent of the coefficient of the same-state dummy variable, which gives the home state bias. The third column contains the exponent of the coefficient of the US-Canada border dummy, which gives the average border effect. We do not report standard errors in order to make the table easier to read, but we do gray out and mark with an º the coefficients that are not statistically significant at the 1% level. The complete set of results, which reflect the values of all coefficients included in our regressions and their corresponding standard errors, is presented in Appendix C. We order the NAICS2 sectors by the size of the estimated US-Canada border effect. The double line differentiates between the sectors for which the border effect is higher, and the sectors for which it is lower, than the 6.71 estimate we obtain when we use aggregate data.

Importer and exporter fixed effects included. The dependent variable is the levels of conversion counts aggregated over all conversion types. Estimations performed using PPML. The data cover four years, 2008 to 2011. Robust standard errors (clustering by region pair). Sample size is 3,481 for all regressions.

º not significant at the 1% level.

---

on distance or the border dummies, for one sector, we cannot conclude that trade frictions are necessarily smaller for that sector. The estimated coefficient might be smaller simply because the bundle of products in that sector has a lower elasticity of substitution than the bundle of products in the other sectors.

In our data, the sector with the largest US↔CA dummy coefficient is NAICS 56: "Administrative, support, and waste management services," which covers businesses performing waste disposal, office cleaning, and administration, security and surveillance. This sector also exhibits strong intranational border coefficients.

The second sector ranked in Table VII is "Public administration" – a sector tied to governments oriented in particular nations. We find a large international border effect in these services, but a smaller home-state bias in part because the sector contains many mostly national (rather than state- or provincial-) government programs.

The sectors ranked third and fourth in Table VII are "Real estate, rental and leasing," and "Health care and social assistance." The services provided under both of these classifications tend to be heavily localized. Real estate services are particularly bound to a particular physical location. In our regressions, this sector is one of only four industries for which the coefficient on distance is statistically different from zero.

The fact that all four service sectors ranked at the top of Table VII describe services that must be consumed locally make it less likely that the large coefficients estimated for the US↔CA dummy variable for each of these sectors are merely the result of these industries having larger elasticities of substitution. The local or national character of the services described by these four classifications is likely to be one of the drivers behind the large border effects for these industries – and the large estimates we obtain for the US↔CA dummy variable coefficients.

9   Trade in Final Goods

Almost a third of all conversions in our data set fall within retail trade – a sector covering business-to-consumer or consumer-to-consumer transactions. This is a particularly interesting sector for understanding the border effect because it comprises trade in final goods. Insofar as the border effect is driven by patterns in intermediate goods or multi-stage production, these factors should not affect sectoral results in retail trade.

The estimates reported in Table VII show a strong US-Canada border effect even in retail trade of final goods. By comparison, intranational trade for this sector is 6.2 times higher than international trade. In Table VIII, we further study retail trade through its NAICS3

subsectors by applying our specification to the subsector data.[56] Although this further disaggregation does not provide a complete explanation for the existence of the border effect, it is helpful in elucidating possible factors that give rise to it.

As with the NAICS2-level sectoral results above, we find stronger border effects for retail subsectors featuring locally consumed goods and regulation. "Motor vehicle and parts dealers" has the largest estimated coefficient for the US↔CA dummy variable. Purchasing a vehicle from abroad is very difficult for regulatory and customs reasons and requires extensive hassle to import the vehicle.[57]

Given the bureaucratic and regulatory procedures that end consumers must face when purchasing a car online, it is not surprising that for Motor vehicle and parts dealers, trade across the US-Canada border is 21.8 times less than trade within these countries' borders. This is an indication that regulation may hinder online trade and that policy aimed at lessening the regulatory burden faced by end consumers would be desirable.

In Table VIII, "Food and beverage" stores also have a high estimated border coefficient for the US↔CA dummy variable, implying an estimated border effect of 14.40.[58] One possible reason for this border effect may be the details of NAFTA. Some of these "Food and Beverage" sales may be classified as trade in agricultural products. Although NAFTA substantially liberalized agricultural trade, the Canada-US agricultural portion

---

[56]The complete set of results is presented in Appendix D.

[57]We have identified the set of steps necessary for a Canadian customer wishing to import a car from a US owner or dealership. The steps described here are provided in detail on Transport Canada's website, at www.tc.gc.ca, accessed on September 13, 2016. We summarize the procedures for importing a car from the US into Canada, rather than the other way around, as cars in the United States are usually cheaper than cars in Canada, so trade in motor vehicles is more likely to initiate in the US.

After extensive customs paperwork, the Canadian customer must submit proof to the Canadian Registrar of Imported Vehicles (RIV) that the vehicle has no outstanding recalls and that it suffered no modifications, such as being adapted for disabled access. After submitting the required paperwork, the Canadian customer must pay a RIV fee and present the vehicle for a federal standards inspection. To pass the inspection, she must ensure that the US vehicle she is bringing into the country is in compliance with Canadian requirements such as bilingual and metric labeling. Vehicles that pass the RIV inspection can then be registered and licensed in Canada. Vehicles that do not pass the RIV inspection have to be exported back to the US or destroyed under the supervision of Canadian customs officials.

Even once their vehicles are licensed, the long list of bureaucratic and regulatory hurdles that Canadian customers face when purchasing cars across the border does not end. If their cars purchased from the United States break down, Canadian customers might end up having to pay for all the repairs, as many Canadian companies do not honor U.S. factory warranties.

[58]Intranational trade is more than 14 times higher than international trade.

of the NAFTA agreement is significantly less liberal than the Mexico-US agreement.[59] Although our results on "Food and Beverage" have a strong US-Canada border effect, the intranational border effects appear much smaller (i.e., food and beverages travel well inside these countries).

Table VIII: Results at the NAICS3 Level: Retail Trade

NAICS3 Sector:	$\ln d_{ij}$ (1)	$internal_{ij}$ (exponent) (2)	$border\ US \leftrightarrow CA_{ij}$ (exponent) (3)
Motor vehicle and parts dealers	-0.051°	2.74	21.80
Furniture and home furnishings	0.013°	3.42	17.89
Food and beverage stores	-0.021°	1.21°	14.40
Clothing and clothing accessory stores	-0.015°	1.31	9.53
Building material and garden equipment	-0.038°	3.11	4.73
Electronics and appliance stores	-0.008°	1.13°	3.86
Sporting goods, hobby, book, and music	0.006°	1.41	3.82
Health and personal care stores	-0.031°	1.52	3.67

Importer and exporter fixed effects included. The dependent variable is the levels of conversion counts aggregated over all conversion types. Estimations performed using PPML. The data cover four years, 2008 to 2011. Robust standard errors (clustering by region pair). Sample size is 3,481 for all regressions. The complete set of results is presented in Appendix D.
° not significant at the 1% level.

At the other end of the table, we find that Electronics and appliance stores, Sporting goods, hobby, books, and music products, as well as Health and personal care stores have fairly low estimated coefficients for the US↔CA dummy variable. Products such as iPhones, book, or Crest toothpaste are essentially commodities that are identical whether purchased in the US or Canada. "Clothing and clothing accessory" stores have an estimated border effect that is much higher (clothing sizes differ across the borders).

---

[59]Agriculture was a rare aspect of NAFTA that was not negotiated trilaterally. Instead, three bilateral agreements were signed between Mexico, Canada and the United States.

10 Conclusion

In this study, we use a proprietary data set from Google to examine the US-Canada border effect. We add to the existing literature on two important dimensions. First, we are the only study to quantify the effect that the US-Canada border has on online trade. Second, our data allows us to obtain insights into the factors that drive the border effect in e-commerce.

We find that the US-Canada border influences online trade flows negatively and significantly: All else being equal, online trade between two US states or two Canadian provinces is 6.7 times higher than trade between a US state and a Canadian province. In an online setting, where information costs are very low, this is a surprising finding.

We also utilize our unique data about consumer search, advertising and clicks to isolate the border effect in purchase decisions from the border effect in availability and consideration. A large portion of the border effect (about 33%) persists after accounting for decisions affecting consumers' final menu of choices. The remaining 2/3rds of the effect comes prior to the final purchase decision, during the process of consumers' arrival on sellers' websites through browsing and advertising.

We also find strong border effects in products that do not need to be shipped. A large portion of our sample involves the purchase of digital goods with no physical shipping costs (software upgrades, downloads, and the purchase of virtual goods). The border effects for these transactions are even higher than in the overall sample.

Our study also shows that border effects vary widely across sectors of online economic activity. When we disaggregate our online data into NAICS2 industries, the US-Canada border affects the sector with the highest estimated border effect 7.4 times more than it affects the sector with the lowest estimated effect. We show some evidence that online trade is affected, to some extent, by regulatory barriers. Disaggregating retail at the NAICS3 level shows that goods, like motor vehicles, that face heavy regulatory and bureaucratic hurdles when crossing the border between the United States and Canada encounter very high international border effects. This result suggests that there is a role for policy to reduce the online border effect for these products.

Lastly, we examine trade in online retail – a category composed mostly of trade in final

outputs, and excluding trade in intermediate input goods. We again find a large border coefficient, suggesting that intranational retail trade online is 6.2 times higher than international retail trade. Our data on business-to-consumer transactions in the online retail allow us to measure border effects in a setting where multi-stage production is largely absent. Our results show a strong border effect in this sector, even in the absence of B2B transactions in intermediate inputs.

In the past decade, international trade has substantially digitized. Technology has created new opportunities for buyers and sellers in different countries to reduce their search and transaction costs. Digitization has also created new opportunities and challenges for understanding international economics. The digitization of trade will produce not only online equivalents of offline data, but will also produce a new abundance of detail. These new set of covariates can be useful for better understanding classic patterns in international economics and for exploring new phenomena.

References

Anderson, James E and Eric van Wincoop, "Gravity with gravitas: A solution to the border puzzle," *American Economic Review*, 2003, *93* (1), 170–192.

_ **and Yoto V Yotov**, "The changing incidence of geography," *American Economic Review*, 2010, *100* (5), 2157–2186.

_ **, Catherine A Milot, and Yoto V Yotov**, "How much does geography deflect services trade? Canadian answers," *International Economic Review*, 2014, *55* (3), 791–818.

Boivin, Jean, Robert Clark, and Nicolas Vincent, "Virtual borders," *Journal of International Economics*, 2012, *86* (2), 327–335.

Broda, Christian and David E Weinstein, "Globalization and the gains from variety," *Quarterly Journal of Economics*, 2006, *121* (2), 541–585.

_ **and** _ , "Understanding international price differences using barcode data," Technical Report, National Bureau of Economic Research 2008.

Brown, W Mark and William P Anderson, "Spatial markets and the potential for economic integration between Canadian and US regions," *Papers in Regional Science*, 2002, *81* (1), 99–120.

Brynjolfsson, Erik, Yu Hu, and Mohammad S Rahman, "Battle of the retail channels: How product selection and geography drive cross-channel competition," *Management Science*, 2009, *55* (11), 1755–1765.

Chen, Natalie, "Intra-national versus international trade in the European Union: Why do national borders matter?," *Journal of International Economics*, 2004, *63* (1), 93–118.

_ **and Dennis Novy**, "Gravity, trade integration, and heterogeneity across industries," *Journal of International Economics*, 2011, *85* (2), 206–221.

Choi, Hyunyoung and Hal Varian, "Predicting the present with Google Trends," *Economic Record*, 2012, *88* (s1), 2–9.

Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer, "The trade-creating effects of business and social networks: Evidence from France," *Journal of International Economics*, 2005, *66* (1), 1–29.

Eaton, Jonathan and Samuel Kortum, "Technology, geography, and trade," *Econometrica*, 2002, *70* (2), 1741–1779.

Einav, Liran, Dan Knoepfle, Jonathan Levin, and Neel Sundaresan, "Sales taxes and internet commerce," *American Economic Review*, 2014, *104* (1), 1–26.

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

Feenstra, Robert C, "Border effects and the gravity equation: Consistent methods for estimation," *Scottish Journal of Political Economy*, 2002, *49* (5), 491–506.

Forman, Chris, Anindya Ghose, and Avi Goldfarb, "Competition between local and electronic markets: How the benefit of buying online depends on where you live," *Management Science*, 2009, *55* (1), 47–57.

Garmendia, Aitor, Carlos Llano, Asier Minondo, and Francisco Requena, "Networks and the disappearance of the intranational home bias," *Economics Letters*, 2012, *116* (2), 178–182.

Goolsbee, Austan, "Competition in the computer industry: Online versus retail," *Journal of Industrial Economics*, 2001, *49* (4), 487–499.

Gopinath, Gita, Pierre-Olivier Gourinchas, Chang-Tai Hsieh, and Nicholas Li, "International prices, costs, and markup differences," *American Economic Review*, 2011, *101* (6), 2450–2486.

Gorodnichenko, Yuriy and Linda L Tesar, "Border effect or country effect? Seattle may not be so far from Vancouver after all," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 219–241.

Head, Keith and Thierry Mayer, "Non-Europe: The magnitude and causes of market fragmentation in the EU," *Weltwirtschaftliches Archiv*, 2000, *136* (2), 284–314.

_ **and** _ , "Illusory border effects: Distance mismeasurement inflates estimates of home bias in trade," Technical Report, CEPII Research Center 2002.

_ **and** _ , "Gravity equations: Workhorse, toolkit, and cookbook," *Handbook of International Economics*, 2014, *4*, 131–195.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, "Estimating trade flows: Trading partners and trading volumes," *Quarterly Journal of Economics*, 2008, *123* (2), 441–487.

Hillberry, Russell and David Hummels, "Trade responses to geographic frictions: A decomposition using micro-data," *European Economic Review*, 2008, *52* (3), 527–550.

Hillberry, Russell H, "Aggregation bias, compositional change, and the border effect," *Canadian Journal of Economics*, 2002, *35* (3), 517–530.

Hortaçsu, Ali, F Asís Martínez-Jerez, and Jason Douglas, "The geography of trade in online transactions: Evidence from eBay and MercadoLibre," *American Economic Journal: Microeconomics*, 2009, *1* (1), 53–74.

Imbs, Jean and Isabelle Mejean, "Elasticity optimism," *American Economic Journal: Macroeconomics*, 2015, *7* (3), 43–83.

Keohane, Robert O and Joseph S Nye Jr., "Power and interdependence in the information age," *Foreign affairs*, 1998, pp. 81–94.

Lendle, Andreas, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vézina, "There goes gravity: eBay and the death of distance," *Economic Journal*, 2015.

Lewis, Randall A and David H Reiley, "Online ads and offline sales: measuring the effect of retail advertising via a controlled experiment on Yahoo!," *Quantitative Marketing and Economics*, 2014, *12* (3), 235–266.

McCallum, John, "National borders matter: Canada-US regional trade patterns," *American Economic Review*, 1995, *85* (3), 615–623.

Millimet, Daniel L and Thomas Osang, "Do state borders matter for US intranational trade? The role of history and internal migration," *Canadian Journal of Economics*, 2007, *40* (1), 93–126.

Obstfeld, Maurice and Kenneth Rogoff, "The six major puzzles in international macroeconomics: Is there a common cause?," in "NBER Macroeconomics Annual 2000, Volume 15," MIT press, 2001, pp. 339–412.

Ohmae, Kenichi, *The Borderless World, rev ed: Power and Strategy in the Interlinked Economy*, Harper Collins, 1999.

Ramsey, James Bernard, "Tests for specification errors in classical linear least-squares regression analysis," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1969, pp. 350–371.

Rossi-Hansberg, Esteban, "A spatial theory of trade," *American Economic Review*, 2005, *95* (5), 1464–1491.

Silva, JMC Santos and Silvana Tenreyro, "The log of gravity," *Review of Economics and statistics*, 2006, *88* (4), 641–658.

Wolf, Holger C, "Intranational home bias in trade," *Review of Economics and Statistics*, 2000, *82* (4), 555–563.

Yi, Kei-Mu, "Can multistage production explain the home bias in trade?," *American Economic Review*, 2010, *100* (1), 364–393.

Appendix: For Online Publication Only

A   Illustrative Example of Conversion Tracking

The following example illustrates how the conversion tracking software works. Suppose a bicycle producer wishes to advertise her bikes online, so she signs up for a Google AdWords account. Thanks to her new account, when users search on Google – or on millions of other websites that partner with Google – for bicycles, an ad for her business appears. Every time a user clicks on her ad, the bicycle producer pays Google for that click.

The question our bicycle producer faces is whether the clicks on her ad convert into sales. Google's conversion tracking software answers this question. This software consists of code the bicycle producer places on her "thanks for your order" page – the part of her site that users see only after completing a purchase. Once placed, the code records for the advertiser the number of times the users who click on ads reach the "thank you for your order" page. In order words, the conversion tracking code records the number of sales our bicycle producer makes thanks to her ads.

B   Additional Details About Geolocation Data

For buyers, we use estimates based on IP (Internet Protocol) address. For sellers, we use the self-reported address data required of businesses who sign up to be Google advertisers. Although we believe that this information allows us to place most buyers and sellers in the correct region, it is possible that we are attributing a small part of the buyers or sellers to the wrong location. On the buyers' side, the IP address might indicate an incorrect physical location if the user is accessing the Internet using a virtual private network (VPN) connection or if the user is, for whatever reason, using software to mask his or her actual IP address. While we cannot identify these users, the level of sophistication needed to set up a VPN network makes us believe that the users that mask their IP address represent a small enough percentage of the total Internet users to make our location identification on the buyers' side reliable. On the seller side, we have access to two address fields. The first

one is a general mailing address that is self-reported by the sellers. This is the address field that we choose to use. The second one is the address where the credit card used to pay for the ads is registered. A data check reveals that the general mailing address reported by the sellers very rarely differs from the billing address.

C   NAICS2: Full Set of Results

Table IX: Results at the NAICS2 Level

	ln d_{ij}	contig US-CA$_{ij}$	contig CA$_{ij}$	contig US$_{ij}$	internal$_{ij}$	border US ↔ CA$_{ij}$
Admin., support,	0.056	0.989***	0.064	0.411***	0.699***	-3.282***
and waste mgmt.	(0.037)	(0.197)	(0.219)	(0.096)	(0.117)	(0.098)
Public	-0.117***	0.161	-0.665**	-0.002	0.553**	-3.067***
administration	(0.041)	(0.366)	(0.265)	(0.094)	(0.238)	(0.160)
Real estate and	-0.221***	0.684*	-1.004***	-0.083	0.391*	-2.987***
rental and leasing	(0.079)	(0.366)	(0.340)	(0.147)	(0.213)	(0.156)
Health care and	0.102**	1.324***	-0.404	0.210**	0.904***	-2.581***
social assistance	(0.051)	(0.282)	(0.350)	(0.102)	(0.168)	(0.157)
Accomodation and	-0.119	-1.740**	-0.280	-0.087	1.007***	-2.556***
food services	(0.077)	(0.791)	(0.609)	(0.131)	(0.253)	(0.199)
Transportation and	-0.039	0.792***	1.044***	0.074	0.533***	-2.385***
warehousing	(0.032)	(0.205)	(0.384)	(0.070)	(0.108)	(0.127)
Finance and	0.023	-0.156	0.212	0.142***	0.614***	-2.280***
insurance	(0.015)	(0.121)	(0.261)	(0.030)	(0.068)	(0.105)
Education	0.057*	0.021	-1.102**	0.263***	1.152***	-2.188***
services	(0.034)	(0.335)	(0.448)	(0.068)	(0.116)	(0.160)
Retail trade	0.008	0.011	-0.030	0.186***	0.514***	-1.825***
	(0.022)	(0.149)	(0.275)	(0.046)	(0.116)	(0.084)
Info. and cultural	0.059**	-0.128	0.261	0.165***	0.701***	-1.618***
industries	(0.027)	(0.179)	(0.413)	(0.049)	(0.121)	(0.121)
Arts, entertainment,	0.064	0.020	-0.827**	0.453***	1.828***	-1.589***
and recreation	(0.075)	(0.365)	(0.379)	(0.151)	(0.343)	(0.122)
Wholesale trade	-0.070***	0.233	0.526**	0.031	0.317***	-1.523***
	(0.016)	(0.293)	(0.218)	(0.032)	(0.060)	(0.109)
Other services	0.007	0.053	-0.542**	0.089*	0.676***	-1.431***
	(0.020)	(0.107)	(0.254)	(0.049)	(0.104)	(0.088)
Prof., scientific, and	-0.102***	-0.014	-0.091	-0.042	0.268***	-1.399***
technical services	(0.029)	(0.201)	(0.642)	(0.051)	(0.103)	(0.137)
Utilities	-0.124**	-0.297	-0.600	0.381***	1.438***	-1.278***
	(0.052)	(0.198)	(0.407)	(0.103)	(0.199)	(0.140)

Importer and exporter fixed effects included. The dependent variable is the levels of conversion counts aggregated over all conversion types. Estimations performed using PPML. The data cover four years, 2008 to 2011. Robust standard errors (clustering by region pair). Sample size is 3,481 for all regressions.

*** Significant at 1%, ** significant at 5%, * significant at 10%

D NAICS3 – Retail Trade: Full Set of Results

Table X: Results at the NAICS3 Level: Retail Trade

	$\ln d_{ij}$	contig US-CA$_{ij}$	contig CA$_{ij}$	contig US$_{ij}$	internal$_{ij}$	border US $\leftrightarrow$ CA$_{ij}$
Motor vehicle and parts dealers	-0.051 (0.055)	0.164 (0.497)	-0.515 (0.337)	0.187* (0.110)	1.006*** (0.184)	-3.082*** (0.169)
Furniture and home furnishings	0.013 (0.093)	-0.091 (0.744)	0.470 (0.318)	0.486*** (0.146)	1.230*** (0.467)	-2.884*** (0.166)
Food and beverage stores	-0.021 (0.088)	-2.243*** (0.615)	0.233 (0.260)	0.286 (0.196)	0.191 (0.351)	-2.667*** (0.189)
Clothing and clothing accessory stores	-0.015 (0.017)	0.197 (0.157)	-0.651*** (0.129)	0.079** (0.040)	0.270*** (0.068)	-2.254*** (0.054)
Building material and garden equipment	-0.038 (0.029)	0.428 (0.444)	-0.037 (0.396)	0.185*** (0.055)	1.134*** (0.099)	-1.553*** (0.110)
Electronics and appliance stores	-0.008 (0.020)	-0.020 (0.178)	-0.352 (0.258)	0.049 (0.036)	0.125 (0.082)	-1.351*** (0.115)
Sporting goods, hobby, book, and music	0.006 (0.017)	-0.392** (0.169)	-0.773*** (0.283)	0.116*** (0.038)	0.340*** (0.075)	-1.341*** (0.119)
Health and personal care stores	-0.031 (0.022)	-0.069 (0.068)	-0.221 (0.300)	0.017 (0.045)	0.420*** (0.078)	-1.299*** (0.101)

Importer and exporter fixed effects included. The dependent variable is the levels of conversion counts aggregated over all conversion types. Estimations performed using PPML. The data cover four years, 2008 to 2011. Robust standard errors (clustering by region pair). Sample size is 3,481 for all regressions.

*** Significant at 1%, ** significant at 5%, * significant at 10%