DISCUSSION PAPER SERIES

No. 8730

CONSUMER ARBITRAGE ACROSS A POROUS BORDER

Ambarish Chandra, Keith Head
and Mariano Tappata

*INTERNATIONAL MACROECONOMICS
and INTERNATIONAL TRADE AND
REGIONAL ECONOMICS*

Centre for Economic Policy Research

www.cepr.org

CONSUMER ARBITRAGE ACROSS A POROUS BORDER

Ambarish Chandra, University of Toronto
Keith Head, University of British Columbia and CEPR
Mariano Tappata, University of British Columbia

This Discussion Paper is issued under the auspices of the Centre's research programme in **INTERNATIONAL MACROECONOMICS and INTERNATIONAL TRADE AND REGIONAL ECONOMICS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

CEPR Discussion Paper No. 8730

January 2012

ABSTRACT

Consumer Arbitrage Across a Porous Border*

National borders, including the easily crossed US-Canada border, have been shown to separate markets and sustain price differences. The resulting arbitrage opportunities vary temporally with the exchange rate and cross-sectionally with travelers' distance to the border. We estimate a structural model of the border crossing decision using data on the location of Canadian crossers and their date of travel. Price differences motivate cross-border travel; a 10% exchange rate appreciation raises the average crosser's welfare by 2.1%. Distance strongly inhibits crossings, with an implied cost of $0.9 per mile. These costs prevent consumers from fully arbitraging price differences, leading to partial segmentation.

Ambarish Chandra
Rotman School of Management
University of Toronto
105 St. George Street
Toronto, ON, M5S 3E6
CANADA

Keith Head
Sauder School of Business
University of British Columbia
2053 Main Hall
Vancouver, BC V6T1Z2
CANADA

Email: ambarish.chandra
@rotman.utoronto.ca

Email: keith.head@sauder.ubc.ca

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=174815

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=154019

Mariano Tappata
Sauder School of Business
University of British Columbia
2053 Main Mall
Vancouver, BC V6T 1Z2
CANADA

Email:
mariano.tappata@sauder.ubc.ca

1   Introduction

One of the most vexing questions in economics has been how to characterize the extent of integration between markets in different countries. The challenge is to reconcile large cross-border flows of goods and people—visible at any land or sea port—with the results of statistical analyses of prices that find strong evidence of market segmentation. The Engel and Rogers (1996) study of price dispersion between cities in Canada and the US reports that crossing the border is equivalent to a distance of 1,780 miles. Presumably, this would surprise the 50 million residents of Canada and the United States who drove across the border in 2010. On average, each person living within a three hour drive of the border makes more than one cross-border car trip per year.[1] Indeed, Canadian residents travel more frequently to the US than they do to other provinces in Canada.[2] While the Engel and Rogers estimate of the border's width has been challenged by Gorodnichenko and Tesar (2009), recent studies that examine disaggregated price data for identical goods on both sides of the border confirm the market segmentation view. Notably, Gopinath et al. (2011) examine pass-through of domestic and foreign costs shocks to grocery store prices and conclude "our results strongly suggests that the US-Canada border almost perfectly segments the retail and wholesale markets..."

Market segmentation, as defined by Gopinath et al. (2011), occurs when transaction costs are high enough to deter all residents of the high price market from buying in the low price market. Investigations of price differences shed some light on the degree of market segmentation. Thus, the Gopinath et al. finding of a discontinuity in grocery prices of 24% at the Canada-US border points towards high transaction costs. Similarly, Burstein and Jaimovich (2009) find substantial amounts of pricing to market using scanner data from both sides of the border.[3] The price studies tend to infer that arbitraging activity is absent or negligible. A smaller literature provides suggestive counter-evidence: Campbell and Lapham (2004) and Baggs et al. (2010) find that exchange rate changes affect employment and exit of retail firms located near the US-Canada border. Asplund et al. (2007) infer cross-border shopping for alcohol between Sweden and Denmark by observing response of retail sales to variation in relative prices caused by taxes and exchange rate fluctuation. Neither of the main strands of work on market segmentation directly consider the actual behavior of the travelers who potentially arbitrage across markets.

This paper develops and estimates the first structural model of the decision by residents of one country to cross the border and purchase cheaper goods in the other

---

[1] 22 million Canadians and 24 million Americans reside in this region.

[2] In 2004, the most recent year for which data are available, Canadian residents made 22 million inter-provincial trips, compared to 36 million trips to the US.

[3] Goldberg and Knetter (1997), summarizing the earlier literature, point out that studies consistently find significant pricing to market. Boivin et al. (2011) show that even online book prices differ greatly between the US and Canada, and that their prices do not respond to exchange rate movements, thereby indicating a large degree of market segmentation.

country. In the model the benefits of crossing are a convex function of the real exchange rate. The convexity arises because a stronger home currency expands the set of goods that are cheaper in the foreign country. Our estimated results provide robust support for this hypothesis. Evaluated at recent exchange rates, the crossing elasticity is around two, which is approximately twice the elasticity observed when the currency is weak, and higher than the Blonigen and Wilson (1999) estimates for the responsiveness of US-Canada trade in goods. Offsetting the benefits of crossing are the observed costs related to fuel prices and distance to the border. A one percent increase in distance reduces the propensity to cross by almost the same amount as a one percent exchange rate appreciation increases it. The view of the market that emerges from our results is one of partial segmentation. Consumer arbitrage is visible in the behavior of same-day travelers but it is concentrated among residents near the border. Thus the median day tripper in Canada lives 18 miles from the border whereas the median Canadian lives 81 miles away. Even among the minority that is most likely to consider crossing for shopping purposes, the responsiveness to prices gaps is finite, leaving scope for pricing to market.

Cross-border movement is not only important as a determinant of market integration. Understanding human travel is also vital for infrastructure planning, traffic forecasting, taxation, preventing terrorism, and controlling the spread of infectious diseases. Motivated in part by these concerns, recent studies of "human dynamics" have applied ideas from statistical physics to analyze the movement of people. Using bank notes (Brockmann et al., 2006) and mobile phones (González et al., 2008) to track individuals, scientists have shown that most travel is over short ranges but the distribution of distances traveled has a very long tail. Instead of following bank notes or cell phones, our study takes advantage of the careful tracking of border crossings undertaken by the Canadian Border Services Agency. In contrast to the purely statistical models employed so far in the science literature, our model emphasizes the economic motivations that we hypothesize to underlie much of human travel.

The main economic motivation for cross-border travel is to purchase goods and services in the less expensive jurisdiction. Our paper relates to the literature on intranational border crossings. These studies generally exploit differences in taxes, since products are priced in the same currency. They also tend to examine cross-border shopping for a single good. Chiou and Muehlegger (2008) examine the circumstances under which US residents cross state lines to take advantage of tax differences on the sale of cigarettes. Similar to our paper, they have access to survey data on the residence of individuals, which allows them to calculate the distance to the nearest state border, and thus permits them to estimate the relative importance of cigarette taxes and travel costs. However, other studies generally do not have data on the location of consumers, and instead rely on sales data to infer the extent of cross-border sales. For example, Manuszak and Moul (2009) estimate how differences in gasoline taxes across US states create incentives for residents to cross state borders. Knight and Schiff (2010) exploit the varying payoffs offered by state lotteries, rather

than tax differences, to estimate the extent to which consumers cross US state borders to purchase lottery tickets. Rather than focus on the decision of where to purchase a single good, we model the endogenous decision of the range of goods that each consumer will purchase across the border.

The paper comprises three main exercises. We begin with reduced form regressions which uncover a number of stylized facts that our model will need to be able to accommodate. First, we establish that travelers respond strongly to the economic incentives created by fluctuations in the exchange rate, suggesting that cross-border shopping is an important economic phenomenon. This finding corroborates results from reduced form estimations conducted by Ford (1992), Di Matteo and Di Matteo (1993, 1996) and Ferris (2000, 2010). Second, we find that US and Canadian residents respond differently to changes in the buying power of their home currency. In particular, while residents of both countries cross the border more when their currency appreciates, Canadian residents have a higher elasticity to exchange rate changes. And finally, we find that exchange rate elasticities depend on the level of the exchange rate. In both countries, the elasticity of crossings with respect to the exchange rate increases in absolute value as the home currency strengthens.

We then develop a model to explain these patterns. Drawing on Dornbusch et al. (1977), our model assumes a continuum of goods available in both countries. Travelers who cross the border purchase the set of goods in each country that is cheaper in that country. Travelers who do not cross purchase all goods at home. The model naturally generates the prediction that as the home currency strengthens, the elasticity of crossings rises in absolute value. However, this is not because of heterogeneity in travel costs across residents, which tends to work in the opposite direction. Instead, the result is for two reasons: first, goods that were already cheaper in the foreign country are even more attractive now. Second, the set of goods that are cheaper in the foreign country expands.

Using a new dataset with information on the residence of consumers and their date of crossing, we estimate the parameters of this model. The geographic and temporal variation allows us to estimate a structural model of the international border crossing decision that can reveal the implicit trade-off between travel costs and lower prices. Our estimated coefficients imply that the median crosser requires savings of almost $30 per hour of travel time. The model also permits counterfactual experiments with respect to the key variables. We show that a 10% appreciation of the real exchange rate would increase cross-border travel frequencies by about 10% when the Canadian dollar is weak but by 24% when it is strong. On the other hand, an exogenous doubling of border wait times would lower crossing frequencies by 50–60%, depending on the province. We estimate that travel has fallen by 32% since September 11, 2001, compared with the otherwise expected level of travel given the realized values of the exchange rate, gasoline prices, income, and population. The model provides a natural way to calculate the average crosser's welfare gains in response to these changes. We find the 10% appreciation yields average crosser gains of 2.1% whereas

the consequences of 9/11 have lowered average crosser welfare by 3.4%.

In the next section we establish patterns of cross-border travel and document the differing effects of exchange rate changes across the two countries as well as over time. In Section 3 we present the model of cross-border travel. We estimate the parameters of this model, calculate the implied travel costs, and conduct counterfactual experiments with respect to the key variables in Section 4. We conclude in Section 5.

2   Stylized facts of border crossings

In this section we estimate the relationship between exchange rates and the propensity of residents of the US and Canada to cross the border. We first show that there is strong evidence that exchange rates influence travel behavior in a manner that is consistent with cross-border shopping. Additionally, we find interesting variation in the response of travelers to currency fluctuations, both across countries and over time.

In Table 1, we present the commonly stated motives for crossing the border. The data are based on the International Travel Survey of visitors and returning residents to Canada. Approximately 50,000 travelers who cross the land border are asked to fill out these anonymous surveys each year; more details on the data are presented in Section 4.2. Trips for pleasure or personal reasons, which include shopping trips,

Table 1: Reasons for Crossing the Border, 1990–2010 (in percent)

Trip Duration:	Sameday		Overnight	
Residence of Travelers:	US	Canada	US	Canada
Business Affairs	7.5	7.4	7.6	7.5
Visit friends/relatives	15.2	8.8	22.8	22.2
Pleasure or personal trip	43.1	53.2	62.3	64.6
Commuting to work	2.3	6.0	-	-
Other	21.1	15.4	7.2	5.5
Not stated	10.8	9.2	0.1	0.2
Total Respondents ('000s)	304	445	226	264

Source: Authors' calculations from the International Travel Survey

are potentially the most likely to respond to exchange rates. The survey responses indicate that this is easily the largest category. Trips for the purpose of business or driving to work, which are likely to be less sensitive to the exchange rate, account for under 10% of responses.

This information suggests that the exchange rate potentially plays an important role in the decision to cross the border, for residents of both countries. We now attempt to quantify the relationship between exchange rates and cross-border travel.

2.1 Data

We obtained data on cross-border travel from Statistics Canada, using information collected by the Canadian Border Services Agency (CBSA).[4] These data consist of counts of all vehicles entering Canada at all land crossings with the United States. US residents encounter the CBSA on their outbound journey and Canadian residents on their return journey.

We use these data on vehicle counts for the 7 Canadian provinces that share a land border with the United States: British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, Quebec and New Brunswick.[5] We use monthly data for the calendar years 1972–2010. Data are available separately for passenger vehicles, commercial vehicles, trucks, motorcycles etc. We focus only on travel by passenger vehicles. The counts are separated by travelers' country of residence, which is determined by whether the vehicle has US or Canadian license plates. Finally, the data are broken down by the length of the cross-border trip. We analyze same-day and overnight trips separately.[6]

We obtained monthly average data on the spot market exchange rate between the US and Canadian currencies. Multiplying the nominal exchange rate by the ratio of monthly CPIs for both countries we construct the Real Exchange Rate (RER) for each month.[7] It is defined with US prices in the numerator such that RER increases correspond to Canadian dollar depreciations. The RER incorporates relative taxes on goods and services in the two countries because the consumer price indexes in both countries are based on after-tax prices. Thus the 1991 introduction of the 7% goods and service tax (GST) in Canada is built into the RER. We fixed the absolute level of the RER using relative price levels from OECD data.

Figure 1 shows patterns in the data over time.[8] Figure 1(a) shows monthly same-day trips by residents of the two countries from 1972 through 2010. Travel is highly seasonal, for residents of both countries. Canadian residents exhibited a sharp rise in same-day trips during the period 1988–1993. The decline in US travel in recent years appears to coincide with the period of heightened security concerns after September 2001, and stricter requirements in recent years regarding passports or other identification. Figure 1(b) shows average travel over the 38-year period for each calendar month. On average, Canadian residents make about 50% more daytrips across the border than do US residents. The number of overnight trips for the two countries is approximately the same. Cross-border travel peaks in the summer months for all groups.

The non-seasonal variation in crossings shown in Figure 1(a) can potentially be

---

[4]See Cansim Table 427-0002.

[5]Nova Scotia has a marine border with the US as it accepts ferry traffic from Maine. The Yukon Territory shares a border with Alaska. We omit these jurisdictions due to difficulties in ascertaining the corresponding US port from which vehicles enter Canada.

[6] "Overnight" is a short-hand to refer to trips spanning two or more days.

[7]Data sources and other details are provided in Appendix A.3.

[8]A table of summary statistics for these data is available in the supplementary file to this paper.

Figure 1: Annual and monthly variation in crossings

(a)

(b)
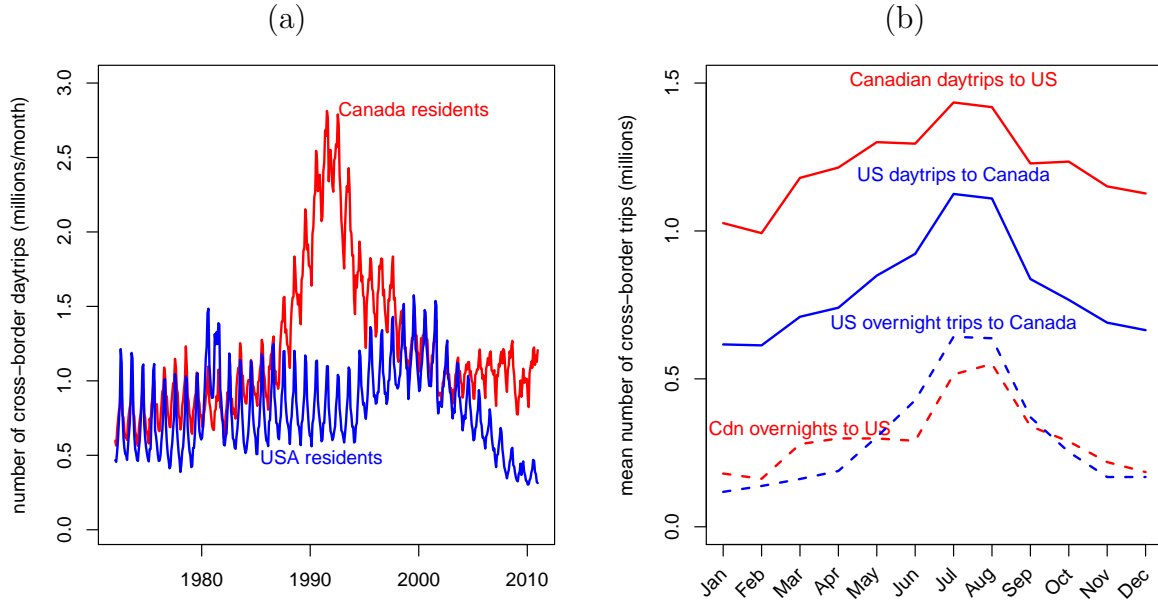
explained by the real exchange rate. The solid line in Figure 2 shows the RER, starting in January 1972 and continuing to December, 2010. The dashed line shows the monthly nominal exchange rates, expressed in the figure as an index of the July 1993 level (1.29 CAD per USD), when the RER was approximately one (that is, prices of the consumer bundle expressed in a common currency were approximately equal). Horizontal dot-dashed lines show the 25th and 75th percentiles of the real exchange rate: "strong USD" corresponds to RER> 1.09 and "strong CAD" corresponds to RER< 0.9. The main messages delivered by this figure are that there is substantial variation in the real exchange rate and, because both countries have mainly had similar inflation rates, the primary source of real variation is nominal variation.[9]

2.2 The Exchange Rate Elasticity of Cross-Border Travel

Our first regression exercise is to determine the elasticity of cross-border trips with respect to the real exchange rate. Our main goal is establish simple data relationships to motivate the development of a model in the subsequent section of the paper. We therefore work with a minimal specification. Denoting the number of cars that cross

---

[9]Put more precisely, log first differences of the nominal exchange rate can explain 94% of the variation in log first differences of the real exchange rate over the period 1972–2010. In levels the R^2 is 0.89.

Figure 2: Canada-US real and nominal exchange rates since 1972
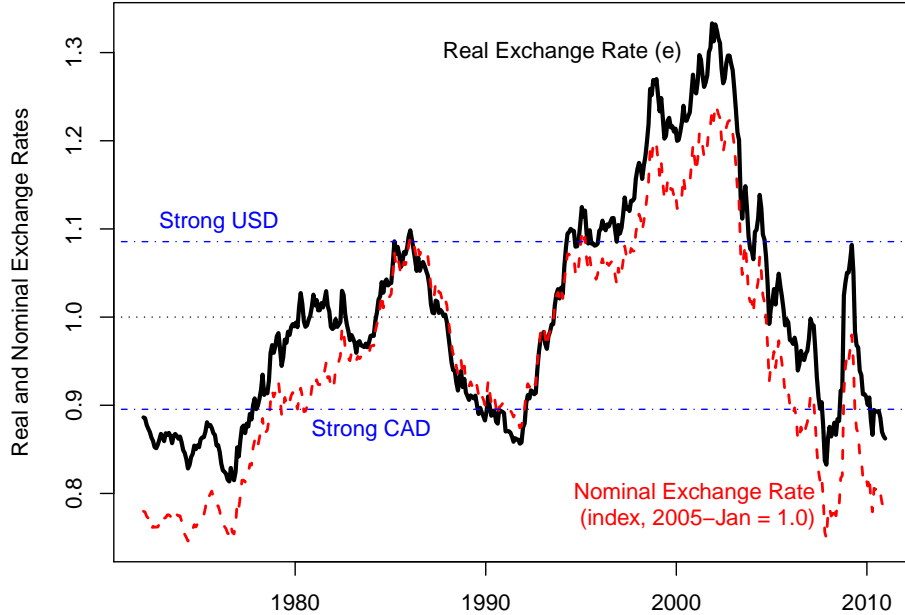
the border by n, and the real exchange rate by e, our specification is:

$$\ln n_{it} = \text{Month}_t + \text{Province}_i + \eta_1 \ln e_t + \eta_2 \text{post911}_t + \eta_3 t + \eta_4 t^2 + \varepsilon_{it}, \qquad (1)$$

where i denotes a province and t denotes time (in months since January 1972). The month effects account for the strong seasonality in travel. We add province fixed-effects, as well as an indicator variable for the period following September 11, 2001 when border security was increased. Finally, we add a linear and quadratic trend to capture secular effects such as population changes. We estimate this equation separately for residents of each country. Therefore, for Canada, this regression models the number of cars returning from the US in a given province and month. For the US, it represents the cars that enter the corresponding Canadian province.

Implicit in the estimation of equation 1 is the assumption that causation runs only from the real exchange rate to crossing decisions. This assumption is defensible because demand for foreign currency created by US and Canadian cross-border shoppers is unlikely to be large enough to move the global foreign exchange markets. To gain some perspective on relative magnitudes, Canadians spent \$4.2bn in the US while Americans spent \$1.8b in Canada during the first quarter of 2010.[10] This represents

---

[10]This includes expenditures by air travelers. Source: International Travel Account Receipts and Payments (`http://statcan.gc.ca/daily-quotidien/100827/dq100827-eng.pdf`)

a mere 0.04% of the foreign exchange turnover involving the Canadian Dollar.[11]

To establish the robustness of the stylized facts, we also estimate using year-on-year differences of equation 1. That is, we subtract from each variable the value it had twelve months before. This holds constant season and province effects and also removes time-varying factors that may not have been well captured by the trend variables:

$$
\begin{aligned}
\ln n_{it} - \ln n_{i,t-12} &= \{12\eta_3 + 144\eta_4\} + \eta_1 \left[\ln e_t - \ln e_{t-12}\right] \\
&+ \eta_2 \left[\text{post911}_t - \text{post911}_{t-12}\right] + 24\eta_4 t + \varepsilon_{it} - \varepsilon_{i,t-12}.
\end{aligned} \tag{2}
$$

The 12-month differences transform the linear trend into the constant term and the quadratic trend to a linear trend.

Table 2: Regression of log crossings, 1972–2010.

Method:	Levels (contemp.)				Year-on-year diffs.			
Length of stay:	Daytrip		Overnight		Daytrip		Overnight	
Residence:	US	CA	US	CA	US	CA	US	CA
$\ln e$	1.24^a	-1.62^a	0.47^a	-1.78^a	0.38^a	-1.16^a	0.12^c	-1.36^a
(CAD/USD)	(0.17)	(0.24)	(0.17)	(0.17)	(0.09)	(0.12)	(0.07)	(0.16)
N	3276	3276	3276	3276	3192	3192	3192	3192
R^2	0.98	0.98	0.96	0.97	0.06	0.27	0.01	0.24
R^2 (excl. $\ln e$)	0.98	0.97	0.96	0.95	0.03	0.05	0.00	0.00
RMSE	0.25	0.22	0.30	0.23	0.14	0.13	0.14	0.15

Newey-West standard errors in parentheses are robust to serial correlation out to 60 months. Significance indicated by ^{c} $p < 0.1$, ^{b} $p < 0.05$, ^{a} $p < 0.01$. An observation is a province-year-month. Coefficients on month and province fixed-effects, the post 9/11 indicator, and the trend variables are not reported.

The results of estimating these equations are presented in Table 2. We treat each province in a calendar month as a separate observation.[12] Since monthly crossing data are serially correlated, we use Newey-West standard errors.[13] The first four columns present results using the contemporaneous specification described in equation 1 and the next four columns use the 12-month difference specification in equation 2. The results of both specifications indicate that travelers respond to the exchange rate, as represented in the negative elasticity of Canadian residents and the positive elasticity of US residents with respect to the real exchange rate. In addition, the elasticities of

---

[11]Source: Authors' calculations from the BIS Central Bank Survey of Foreign Exchange and Derivatives Market Activity, 2010 (http://www.bis.org/publ/rpfxf10t.htm)

[12]In Table 2 in the supplementary file we present corresponding regressions using country-level data, instead of breaking up the data by provinces. The results in that regression are similar to those presented here.

[13]There are too few provinces (7) for clustering at the province level to work.

Canadian residents are bigger than those of US residents, across both specifications and both categories of trip-length.[14]

We investigate whether the crossing elasticity with respect to exchange rates varies with the level of the exchange rate in Table 3. We find significant interactions between the log of the RER and indicators for the highest and lowest quartiles of the RER over the 38-year period. In particular, the coefficient for the period when the US dollar was strong is generally positive, for residents of both countries. This has the effect of increasing the positive elasticity of US residents, and decreasing the negative elasticity of Canadian residents. In other words, US residents become more responsive to the exchange rate in periods when the US dollar is strong, while Canadian residents become less responsive. We observe the opposite pattern during periods when the US dollar is in its lowest quartile.[15]

Table 3: Regression of log crossings using Quartiles of RER, 1972–2010.

Method:	Levels (contemp.)				Year-on-year diffs.			
Length of stay:	Daytrip		Overnight		Daytrip		Overnight	
Residence:	US	CA	US	CA	US	CA	US	CA
$\ln e$	0.93^a	-1.71^a	0.32	-2.08^a	0.55^a	-1.13^a	0.14^c	-1.42^a
(CAD/USD)	(0.28)	(0.28)	(0.23)	(0.21)	(0.11)	(0.12)	(0.08)	(0.15)
$\ln e \times [e > 1.09]$	0.90^b	0.54^c	0.83^a	0.65^b	0.14	0.25	0.26^a	0.27
(strong USD)	(0.37)	(0.33)	(0.31)	(0.29)	(0.11)	(0.18)	(0.07)	(0.16)
$\ln e \times [e < 0.90]$	-0.87^b	-0.87^a	-1.25^a	-0.31	-0.44^a	-0.25^b	-0.22^b	-0.06
(strong CAD)	(0.34)	(0.24)	(0.32)	(0.22)	(0.11)	(0.11)	(0.09)	(0.12)
N	3276	3276	3276	3276	3192	3192	3192	3192
R^2	0.98	0.98	0.97	0.97	0.08	0.28	0.02	0.24
RMSE	0.24	0.21	0.29	0.23	0.14	0.13	0.14	0.15

Newey-West standard errors in parentheses are robust to serial correlation out to 60 months. Significance indicated by ^{c} $p < 0.1$, ^{b} $p < 0.05$, ^{a} $p < 0.01$. An observation is a province-year-month. Coefficients on month and province fixed-effects, the post 9/11 indicator, and the trend variables are not reported.

Canadian residents have zero exemptions from taxes and duties on goods pur-

[14]Adding economic indicators, such as unemployment and GDP, to the regressions has a modest effect on the coefficient of interest, and does not affect the general pattern of results. See Table 4 in the supplementary file for details. We do not include these variables here in order to maintain a minimal specification.

[15]In Table 3 in the supplementary file we present corresponding regressions using country-level data. The results in that regression are similar to those presented here. We also conducted other robustness checks. Instead of using indicators for the top and bottom quartiles of the RER, we used a 10% cutoff above and below PPP values. We also included a second-order term for $\ln e$. All the results indicated the same pattern of exchange rate elasticities being sensitive to the level of the RER.

chased abroad when returning from a trip of less than 24 hours.[16] Despite this, we observe same day travel being extremely sensitive to exchange rates: we estimate the elasticity of Canadian residents as well over 1. It may well be the case that some residents do not report their purchases truthfully, or that border agents do not bother to charge taxes for small amounts.

This section has uncovered four stylized facts of cross-border travel that should be features of a quantitative model of crossing decisions. First, while there is always two-way movement across the border, there are large within- and between-year fluctuations. Second, there is a robust relationship between exchange rates and travel: the stronger the currency in the country of residence, the more trips. Third, elasticities are asymmetric: In absolute value Canadian residents have higher percentage responses to changes in the exchange rate. Fourth, exchange rate elasticities are higher (in absolute value) when the home currency is stronger.

3 Model of the crossing decision

In this section we develop a model that formalizes the trade-offs faced by border crossers. The benefits from crossings are modelled using a continuum of goods structure adapted from Dornbusch, Fischer, and Samuelson (1977), hereafter DFS, and we retain that paper's notation where applicable. This structure implies a non-linear relationship between exchange rates and the savings obtained from cross-border shopping. We model the cost of crossings by incorporating heterogeneity in consumers' distance to the border, which implies differences in time and fuel costs. This in turn explains differences in crossing elasticities between otherwise similar regions. After allowing for idiosyncratic heterogeneity, the model predicts the share of residents in each community that cross each month.

3.1 The supply side

A continuum of goods are indexed z on the interval from zero to one. Good z has price $P(z) = a(z)W$ in the home country, where W is the wage. In the DFS model $a(z)$ represents unit labor requirements. Here we can generalize $a(z)$ to be the multiple of unit labor requirement and a country-industry-specific markup over marginal costs. Thus, $a(z)$ can be thought of as the good-specific ratio of consumer prices to input prices. Analogously, prices in the foreign country are given by $P^*(z) = a^*(z)W^*$. Both prices and wages are expressed in terms of the respective local currency units.

Defining $A(z) \equiv a^*(z)/a(z)$, goods are ordered such that $A'(z) < 0$. The relative

---

[16]Under NAFTA, Canadian residents are not required to pay duties on most products that were manufactured in the US or Mexico. They are generally still required to pay taxes on these purchases. US residents generally have a $200 exemption when returning from a same-day trip to Canada.

price of *foreign goods* in local units is defined as

$$p(z) \equiv P^*(z)/P(z) = A(z)(W^*/W).$$

Since $A'(z) < 0$ and we take relative wages as given, we have $p'(z) < 0$. Whether relative prices differ across goods due to relative costs or relative markups does not matter from the consumer's point of view. The model only requires that the relative (local) price ordering, $p(z)$, remains stable as the exchange rate changes.[17]

Let the domestic currency price of foreign currency be E. The relative price of foreign goods *expressed in a common currency* is therefore $Ep(z)$. Let $\tilde{z}$, the borderline good for which prices are equal (after converting currency), be defined implicitly by

$$P(\tilde{z}) = EP^*(\tilde{z}) \quad \Leftrightarrow \quad Ep(\tilde{z}) = 1.$$

For $0 \leq z \leq \tilde{z}$, goods are cheaper at home and the remaining goods $\tilde{z} \leq z \leq 1$ are cheaper abroad. Given wages and the exchange rate, we can solve for the borderline good as

$$\tilde{z} = A^{-1}[W/(EW^*)] = p^{-1}(1/E). \tag{3}$$

Since $A' < 0$, $\tilde{z}$ is an increasing function of EW^*/W (the relative wage of foreign workers expressed in a common currency). Using relative price notation, the expression following the second equality shows that $\tilde{z}$ is increasing in E. Thus a nominal appreciation of the foreign currency, holding prices constant in local currency units, contracts the range of goods that are cheaper in the foreign country.

We illustrate the model in Figure 3 using data from Porter (2009). The author reports prices for 19 goods available on both sides of the border. Calculating $p(z)$ as the ratio of the US price (in USD) to the Canadian price (in CAD), we sort z in decreasing order and plot relative prices. With the lone exception of ice cream obtained at a Cold Stone Creamery and a 32G iTouch, Canadian prices expressed in local units were higher. At the time the article was written the exchange rate was 1.09 CAD/USD. With such a strong Canadian dollar, it is not surprising that 15 out of 19 goods were less expensive in the US after converting prices to a common currency. The figure shows that dramatic changes would arise if the USD were to revert to the 1972–2010 mean and appreciate by about 15% to 1.25 CAD/USD. With sticky prices, this would lead to a rise in $\tilde{z}$, i.e. a contraction in the set of goods that are cheaper in the US.
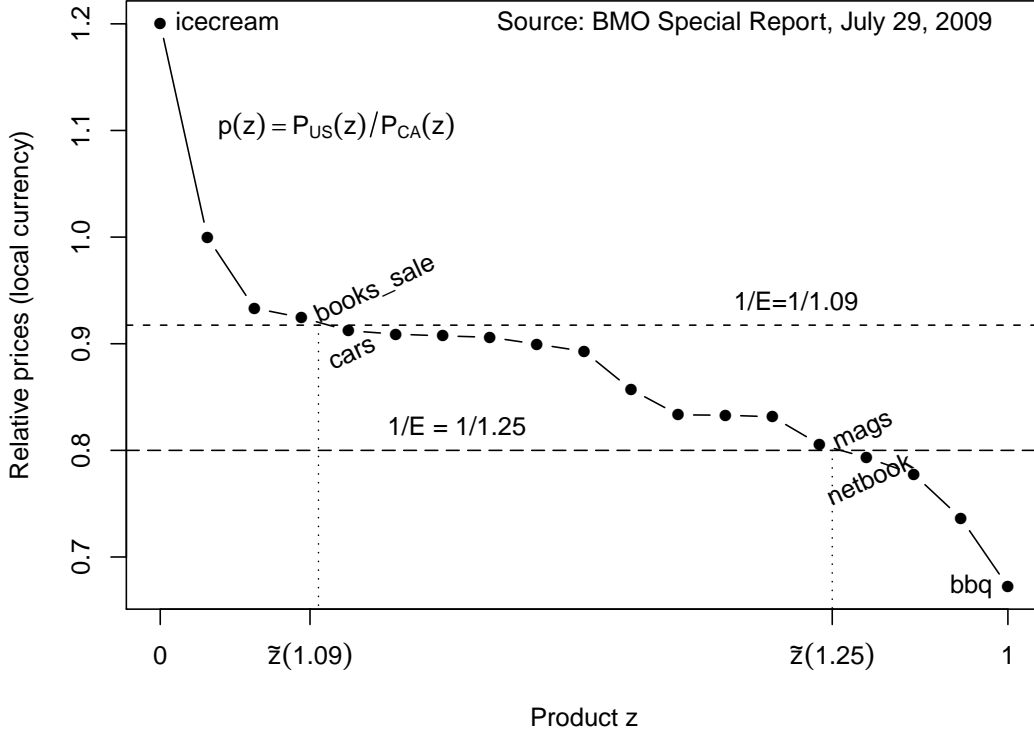
3.2 Consumer problem

Consumers have Cobb-Douglas utility with expenditure share parameters $b(z)$. Utility (subject to a monotone transformation) can be expressed as

$$\ln U = \int_0^1 b(z) \ln C(z) dz,$$

---

[17]One market structure that would generate this result would be Dixit-Stiglitz monopolistic competition.

Figure 3: Exchange rates and relative prices: 19 products

where $C(z)$ denotes consumption of good z.

We now depart from the DFS model by assuming that the purchase of foreign products requires the consumer to engage in cross-border shopping. This does not rule out trade by wholesalers but, due to pricing-to-market by home retailers, consumers can only pay the foreign price if they cross the border. Individuals decide whether to stay at home or cross by comparing the indirect utility associated with each option. Stayers spend their entire income, W, in the home country and have indirect utility, v_S, given by

$$v_S = \ln W - \int_0^1 b(z) \ln P(z) dz.$$

Crossers buy goods $\tilde{z} \leq z \leq 1$ in the foreign country.[18] We assume that the home government does not collect duties or taxes on the goods travelers bring back with them so the price paid for foreign goods is $EP^*(z)$ in domestic currency.[19] The cost

---

[18]The model should not be taken literally since cars cannot physically accommodate all the products that are cheaper in the foreign country. A more realistic approach would be to consider a model of random replacement of durable goods. The $b(z)$ would measure the probability that a particular good needed to be replaced. The v_S and v_X would become expected utilities.

[19]This assumption is grounded in anecdotal evidence on smuggling and *de facto* exemptions for

of the cross-border trip consists of the sum of the opportunity cost of driving time and fuel costs. Letting parameters ψ equal speed (kilometer per hour), ϕ equal fuel efficiency (kilometers per liter), and H equal the endowment of hours, the crossing cost is $D[W/(\psi H) + P(g)/\phi]$, where $P(g)$ is the price of gasoline (per liter) and D is driving distance (in kilometers). Net income of crossers, i.e. W minus the crossing cost, can be expressed as W/τ, with

$$\tau \equiv \left[1 - D\left(\frac{1}{\psi H} + \frac{P(g)}{\phi W}\right)\right]^{-1}. \tag{4}$$

The indirect utility of crossers is given by

$$v_X = \ln W/\tau - \int_0^{\tilde{z}} b(z) \ln P(z)dz - \int_{\tilde{z}}^1 b(z) \ln EP^*(z)dz.$$

Ignoring fuel costs, $1/\tau$ is the constant fraction of income that "melts away" in the trip across the border, analogous to Samuelson's iceberg form of transport costs. In our empirical exercise we will estimate the model with and without fuel costs.

The gross benefits of crossing, denoted B, are given by the savings from buying goods in the foreign country instead of domestically:

$$B \equiv \int_{\tilde{z}}^1 b(z)[\ln P(z) - \ln EP^*(z)]dz.$$

The net benefit of crossing is therefore

$$v_X - v_S = B - \ln \tau(D, P(g)/W).$$

For any interior value of $\tilde{z}$, B is positive since $P(z) > EP^*(z)$ for all $z > \tilde{z}$.

To specify benefits of crossing in terms of the real exchange rate, we now make use of the supply-side assumptions from subsection 3.1. Replacing $\ln P(z)$ with $\ln a(z) + \ln W$ and $\ln P^*(z)$ with $\ln a^*(z) + \ln W^*$, recalling that $A(z) \equiv a^*(z)/a(z)$, and rearranging we obtain

$$B = -\int_{\tilde{z}}^1 b(z) \ln A(z)dz - \ln \frac{EW^*}{W} \int_{\tilde{z}}^1 b(z)dz. \tag{5}$$

While EW^*/W would be one way to define the real exchange rate, it is more customary to do so in terms of price indexes. The model implies a simple relationship between relative price indexes and relative wages. With Cobb-Douglas preferences the natural definition of the price indexes are $\mathbf{P} = \exp(\int_0^1 b(z) \ln P(z)dz)$ and $\mathbf{P}^* = \exp(\int_0^1 b^*(z) \ln P^*(z)dz)$. Substituting in the expressions for prices we obtain

$$\ln \mathbf{P}^*/\mathbf{P} = \ln W^*/W + \ln \kappa,$$

---

small amounts of declared spending. Adding a tax would not change the specification because it would just be a scalar multiplying the real exchange rate

where $\kappa \equiv \exp\left(\int_0^1 [b^*(z)\ln a^*(z) - b(z)\ln a(z)]dz\right)$ is a constant if budget shares and relative productivities across goods do not change over time. We can now express the real exchange rate as a function of the relative wage:

$$e \equiv \frac{E\mathbf{P}^*}{\mathbf{P}} = \kappa \frac{EW^*}{W}.$$

It is also useful to follow DFS in defining $\vartheta(\tilde{z}) = \int_0^{\tilde{z}} b(z)dz$ as the share of expenditures on goods for which the home country is the low-price supplier. Making these substitutions in equation 5, we can express the benefits of crossing as a function of the log real exchange rate:

$$B(\ln e) = -\int_{\tilde{z}}^1 b(z)\ln A(z)dz - (1 - \vartheta(\tilde{z}))(\ln e - \ln \kappa). \qquad (6)$$

Taking the derivative of (6) with respect to $\ln e$ we obtain[20]

$$B' = b(\tilde{z})\ln A(\tilde{z})\frac{\partial \tilde{z}}{\partial \ln e} - (1 - \vartheta(\tilde{z})) + (\ln e - \ln \kappa)\frac{\partial \vartheta(\tilde{z})}{\partial \tilde{z}}\frac{\partial \tilde{z}}{\partial \ln e}$$

Noting that $\frac{\partial \vartheta(\tilde{z})}{\partial \tilde{z}} = b(\tilde{z})$, we can express this as

$$B' = [\ln A(\tilde{z}) + \ln e - \ln \kappa]\, b(\tilde{z})\frac{\partial \tilde{z}}{\partial \ln e} - (1 - \vartheta(\tilde{z})) = -(1 - \vartheta(\tilde{z})) < 0 \qquad (7)$$

The second equality in (7) is obtained by noting that $A(\tilde{z}) = A\left(A^{-1}\left(\frac{W}{EW^*}\right)\right) = \frac{W}{EW^*} = \frac{\kappa}{e}$, which implies $\ln A(\tilde{z}) + \ln e - \ln \kappa = 0$. Thus, the term involving the derivative of $\tilde{z}$ with respect to the exchange rate drops out and the first derivative of benefits is just minus the share of consumer income spent on the goods that are cheaper in the foreign country. The benefits of crossing the border respond more to a given percentage change in the exchange rate when this budget share is high. Foreign appreciation contracts the basket of goods that are cheaper abroad (i.e. rising e decreases $1 - \vartheta(\tilde{z})$). This leads the benefit function to be convex in the real exchange rate:

$$B'' \equiv \frac{\partial^2 B}{\partial \ln e^2} = b(\tilde{z})\frac{\partial \tilde{z}}{\partial \ln e} = -b(\tilde{z})\frac{A(\tilde{z})}{A'(\tilde{z})} > 0. \qquad (8)$$

The convexity of the $B(\ln e)$ function arises under general functional form assumptions for preferences, $b(z)$, and the relative price ordering $A(z)$. However, it is also useful to consider a special case where the integrals have closed form solutions. Suppose equal expenditure on all goods, i.e. $b(z) = 1 \forall z$ and $\ln A(z) = \alpha_0 - \alpha_1 z$ with $\alpha_0, \alpha_1 > 0$.

---

[20]Consistent with our model, we assume that changes in $\ln e$ are generated by either the nominal exchange rate E or the ratio of relative wages (through its effect on relative prices $\mathbf{P}^*/\mathbf{P}$). We do not consider the changes in $\ln e$ generated by adjustments in $A(z)$ as it would require a reordering of goods in the $[0, 1]$ interval.

Parameter α_0 shifts the relative prices of the foreign country in all goods whereas α_1 measures the extent of cross-good heterogeneity in relative prices. Imposing these functional forms simplifies $\ln \kappa$ to $\alpha_0 - \alpha_1/2$ and $\vartheta(\tilde{z})$ to $\tilde{z}$. Solving for the critical good defined in equation 3 yields

$$\tilde{z} = \frac{1}{2} + \frac{1}{\alpha_1} \ln e.$$

To obtain an interior solution the real exchange rate must lie in the range $\exp(-\alpha_1/2) < e < \exp(\alpha_1/2)$.

Plugging in these equations and integrating the benefit function shown in equation 6 yields a quadratic function in which the coefficients are simple functions of α_1, the key cost heterogeneity parameter,

$$B(\ln e) = \beta_0 + \beta_1 \ln e + \beta_2 [\ln e]^2, \tag{9}$$

where $\beta_0 \equiv \alpha_1/8$, $\beta_1 \equiv -1/2 < 0$, and $\beta_2 \equiv \frac{1}{2\alpha_1} > 0$. Note that α_0 does not appear, because it only matters through changes in the real exchange rate. Increasing α_0 is equivalent to increasing $\ln W^*/W$. On the other hand, α_1, the measure of dispersion in productivities, determines both the size of savings for a given basket of goods to be bought in the foreign country and the extent of that shopping basket.

We use the quadratic form shown in equation 9 in our empirical specification. It can be thought of either as a second-order approximation of a general B or as the exact solution under the strong assumptions of uniform budgeting and log-linear $A(z)$.

Foreign crossings into the home country depend on a similar benefit function:

$$
\begin{aligned}
B^*(\ln e) & \equiv \int_0^{\tilde{z}} b^*(z)[\ln EP^*(z) - \ln P(z)]dz \\
& = \int_0^{\tilde{z}} b^*(z) \ln A(z)\, dz + \vartheta^*(\tilde{z})(\ln e - \ln \kappa),
\end{aligned}
$$

where $\vartheta^*(\tilde{z}) = \int_0^{\tilde{z}} b^*(z)dz$ is the share of expenditures that foreign consumers allocate to goods that are less expensive in the home country. The derivative of B^* with respect to the log real exchange rate is given by

$$\frac{\partial B^*}{\partial \ln e} = \vartheta^*(\tilde{z}).$$

The derivatives of the benefits of crossing can only be equal in absolute value if $1 - \vartheta(\tilde{z}) = \vartheta^*(\tilde{z})$.[21] There is only a single value of the real exchange rate that meets this condition. This potentially helps us understand why Canadian elasticities with respect to the RER are larger (in magnitude) than US elasticities. If the set of goods that is cheaper in Canada is generally smaller (i.e. $\tilde{z} < 1/2$) then Canadians will be more responsive to changes to the RER than Americans, since their benefits of crossing can be realized over a larger set of goods.

---

[21]With identical preferences, $b(z) = b^*(z)$, this requires $\vartheta^*(\tilde{z}) = 1/2$.

3.3 Individual Heteregeneity

To account for the fact that only a subset of individuals cross, we make two departures from the representative-consumer DFS. First, communities are located at different distances from the border and their gas price to income ratios differ. Therefore we expect major differences in crossing propensities based on location. To capture this idea we add community c subscripts to the determinants of transport costs shown in equation 4: distance, fuel price, and income.

A second departure from the representative consumer argument is needed to explain why some people in a given community cross but others do not. In particular we introduce $\zeta(i)$ distributed with a CDF denoted $F(\zeta)$ as the individual-specific heterogeneity in the unobserved net benefits of crossing. Combining this with the benefit and transport cost function, net benefits for individual i in community c are given by

$$v_X - v_S = B(\ln e) - \ln \tau_c + \zeta(i).$$

Within each community c there is a marginal individual who is indifferent between crossing and staying. This ζ_c^* is defined by setting $v_X - v_S = 0$, yielding

$$\zeta_c^* = -B(\ln e) + \ln \tau_c.$$

Thus, residents of distant communities (high τ_c) require a higher idiosyncratic shock to benefit from crossing the border. For a community with mass N_c of potential crossers, the aggregate net gains of crossers are given by integrating across the set of individuals for whom $\zeta(i) > \zeta_c^*$:

$$G_c = N_c \int_{\zeta_c^*}^{\infty} [B - \ln \tau_c + \zeta] dF(\zeta). \tag{10}$$

In the results section, we show how G_c can be calculated up to a scalar using the econometric estimates of the B and $\ln \tau$ functions. This permits computation of how changes in the real exchange rate or determinants of τ_c affect the aggregate welfare gains of crossers from all communities.

The next step is to determine the model's predictions for travel elasticities with respect to the real exchange rate and distance. Let x_c denote the probability of crossing for residents of community c. With a continuum of individuals, x_c also measures the fraction who cross:

$$x_c = \mathbb{P}(\zeta(i) < -\zeta_c^*) = F(B(\ln e) - \ln \tau_c). \tag{11}$$

The elasticity of the crossing probability with respect to the real exchange rate is

$$\frac{\partial \ln x_c}{\partial \ln e} = \frac{F'}{F} B' < 0. \tag{12}$$

Since crossing costs increase with distance to the border, $\partial \ln \tau_c / \partial \ln D_c > 0$, the elasticity of crossing with respect to distance is negative:

$$\frac{\partial \ln x_c}{\partial \ln D_c} = -\frac{F'}{F}\frac{\partial \ln \tau_c}{\partial \ln D_c} < 0.$$

While these elasticities can be signed for general distributions on individual heterogeneity, the second derivative with respect to $\ln e$ and the cross-partial of how distance affects the exchange rate elasticity cannot be signed without determining the shape of F'/F. A change in the real exchange rate level determines (through B) a shift in the location of the marginal crosser, ζ_c^*, in the $F()$ domain. The rate of such change depends on the initial location of the marginal crosser as well as the curvature of F.

Differentiating equation 12 we obtain

$$\frac{\partial^2 \ln x_c}{\partial \ln e^2} = \frac{[FF'' - (F')^2]}{F^2}(B')^2 + \frac{F'}{F}B''.$$

Examination of this expression leads to two important results. First, once heterogeneity is added into the model, the positive second derivative of the individual benefit function (B'') shown in (8) will not translate into a positive second derivative for aggregate log crossings if the term in brackets is sufficiently negative. Second, we see that in models with a constant elasticity at the individual level ($B'' = 0$), convexity of log crossings requires the term in square brackets to be positive. For commonly used distributions of individual heterogeneity, the factor in brackets has a negative sign.[22]

Consider the consequences of aggregating multiple communities c, of size N_c into a single region R of size $N_R = \sum_{c \in R} N_c$ which could be a province, state or country.

$$x_R = \sum_{c \in R} \frac{N_c}{N_R}x_c$$

The elasticity of crossings of this region with respect to the RER is given by

$$\frac{\partial \ln x_R}{\partial \ln e} = \sum_{c \in R} \frac{N_c}{N_R}\frac{x_c}{x_R}\frac{\partial \ln x_c}{\partial \ln e}. \tag{13}$$

The cross-partial effect of distance and the log exchange rate on crossing is

$$\frac{\partial^2 \ln x_c}{\partial \ln e \partial D_c} = -\frac{[FF'' - (F')^2]}{F^2}B'\frac{\partial \ln \tau_c}{\partial \ln D_c}. \tag{14}$$

---

[22]F'/F is globally decreasing for uniform, normal, logit, gumbel. Even the highly convex power distributions, $F(\zeta) = (\zeta/\bar{\zeta})^\lambda$ for $0 < \zeta < \bar{\zeta}$ has F'/F decreasing. Although certain parameterizations of beta distributions can have upward sloping regions in the right tail, our numerical analysis suggests F'/F is decreasing over most of the support.

Figure 4: Population and Distance to the Border
(a)   (b)
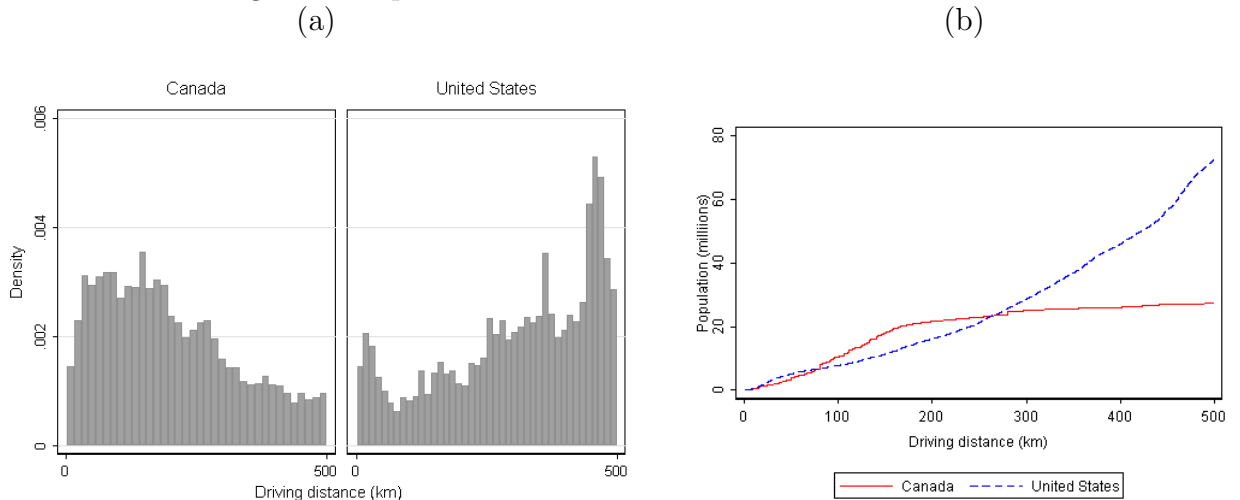
Since commonly used $F()$ distributions imply that the term in brackets is negative, equation 14 leads to a somewhat counterintuitive prediction: As distance to the border increases, exchange rate responsiveness—as measured by the absolute value of the elasticity—becomes stronger.[23] Considering both equations 13 and 14 we can infer that a region with population clustered far from the border should have a high (in absolute value) estimated crossing elasticity, in large part because it has low crossing rates. This implies that differences in the geographic distribution of the population have the potential to explain differences in crossing elasticities.

To understand the mechanisms at work in the model, we analyze the case of no heterogeneity within a community: i.e. $\zeta(i) = 0, \forall i$. Thus, communities can be divided into two groups: those where all residents cross and those where no resident crosses. The marginal community is located at a critical distance D^* such that $B(\ln e) = \ln \tau(D^*, \cdot)$. However, the elasticities of travel cannot be signed since the result of a change in $\ln e$ depends on the population size of the new communities that now cross the border with respect to the total number of residents that were engaged in cross-border shopping. In other words, information regarding the distribution of population and distance to the border are required.

Consider now the spatial distribution of residents in Canada and the US. Figure 4 shows the differences in terms of population density and distance to the border.[24] Panel (a) shows that a higher proportion of Canadians live near the border relative to the United States. Panel (b) shows the accumulated population as we move farther from the border. The argument above suggests that, ignoring idiosyncratic shocks,

---

[23]The response in *levels* of crossings shrinks with distance from the border.

[24]The figures were constructed by calculating the driving distance from each census tract to the closest land border. Details are provided in Appendix A.2.

the crossing elasticity would be similar for the two regions if the critical distance D^* falls within 100km from the border (both countries have roughly the same total population at any given point in the range). The elasticity for Canadians would be higher than for Americans in the 100–170km range since, as the critical distance increases, Canadian communities that start crossing are more populated than those in the US. The opposite is true for distances larger than 200km, as the number of American crossers expands significantly while it barely increases for Canada.

Considering these results along with those of Section 3.2, we conclude that the asymmetric elasticities of Canadian and US residents with respect to the RER, that we established in Section 2, may be attributable to two factors. These are the geographic distribution of the population in the two countries, as well as differences in the expenditure shares for goods that are cheaper in either country. Since we have established that distance of travelers to the border affects the exchange rate elasticity of travel, the model should be estimated using geographically disaggregated data. In particular, it is important to use data on the distance to the border from each community from which travelers originate.

4 Estimation of the model

In this section we take the model of the previous section to the data. We use our estimates to calculate implied travel costs and to conduct counterfactual welfare analysis.

4.1 Regression Specification

In order to estimate the crossing fraction equation shown in (11), we need to parameterize the crossing benefit and cost functions (B and $\ln \tau_c$) as well as specify the distribution of individual heterogeneity ($F(\zeta)$). We make use of the quadratic form for $B(\ln e)$ shown in equation 9. It can be thought of either as the exact solution under uniform budgeting and log-linear $A(z)$ or can be considered a second order approximation for $B(\ln e)$. This is the simplest form that allows us to test for the convexity which is a distinguishing feature of our model of cross-border shopping.

Next, we parameterize the border crossing costs that apply to all individuals using a linear-in-logs approximation of the τ_c function shown in equation 4:

$$\ln \tau_c = \gamma_0 + \gamma_1 \ln D_c + \gamma_2 \ln \left[P(g)_c / W_c \right]. \tag{15}$$

The γ_0 parameter shifts travel costs at all distances. One such shifter would be border formality compliance costs.[25] The γ_1 parameter represents the elasticity of

---

[25]Since these costs are thought to have risen following September 11th, 2001, we include a Post-9/11 dummy in most specifications.

travel costs with respect to distance.[26]

Substituting the B and $\ln \tau$ functions into equation 11, we can express the crossing fraction as

$$x_c = F[\beta_0 - \gamma_0 + \beta_1 \ln e + \beta_2 (\ln e)^2 - \gamma_1 \ln D_c - \gamma_2 \ln (P(g)_c / W_c)]. \qquad (16)$$

Next, we need to impose a particular functional form for $F(\zeta)$. Idiosyncratic crossing shocks $\zeta(i)$ are likely to depend on the sum of a large number of at least partially independent factors. The central limit theorem would therefore lead ζ to be distributed normally. Assuming ζ has expectation μ and variance σ^2, $F(\zeta) = \Phi([\zeta - \mu]/\sigma)$, where $\Phi()$ denotes the standard normal CDF. Substituting these parameterizations into equation 16 and adding time subscripts we obtain

$$x_{ct} = \Phi[\theta_0 + \theta_1 \ln e_t + \theta_2 [\ln e_t]^2 + \theta_3 \ln D_c + \theta_4 \ln (P(g)_{ct} / W_{ct})], \qquad (17)$$

where Table 4 shows the mapping between the θ and the structural parameters as well as the expected signs for each coefficient.

Table 4: Interpretation of coefficients

Parameter	Covariate	Structure	Sign
θ_0	constant,	$(\beta_0 - \gamma_0 + \mu)/\sigma = (\alpha_1/8 - \gamma_0 + \mu)/\sigma$	$+$ or $-$
θ_1	$\ln e_t$ (RER)	$\beta_1/\sigma = -1/(2\sigma)$	$-$
θ_2	$(\ln e_t)^2$	$\beta_2/\sigma = 1/(2\alpha_1\sigma)$	$+$
θ_3	$\ln D_c$	$-\gamma_1/\sigma$	$-$
θ_4	$\ln [P(g)_{ct}/W_{ct}]$	$-\gamma_2/\sigma$	$-$

Equation 17 is not yet suitable for estimation purposes because it does not allow for deviations between observed crossing fractions and those predicted by the model. Such deviations would arise from at least two sources. First, the continuum assumption is only an approximation, so the actual crossing share would only be equal to the crossing probability in expectation. Second, our data are based on a survey given out to a subset of the actual population of crossers. We elaborate on this point in section 4.2 and Appendix A. We restate equation 17 in the form of a conditional expectation:

$$\mathbb{E}[x_{ct} \mid e_t, D_c, P(g)_{ct}, W_{ct}] = \Phi[\theta_0 + \theta_1 \ln e_t + \theta_2 [\ln e_t]^2 + \theta_3 \ln D_c + \theta_4 \ln (P(g)_{ct} / W_{ct})]. \qquad (18)$$

---

[26]The empirical trade literature routinely assumes a constant elasticity of distance. We report estimation results using quadratic distance functions in the supplementary file. We also estimated a semi-parametric step function. Neither generalization improves the fit enough to justify the loss in parsimony.

Quasi-likelihood estimation of this fractional probit model yields consistent estimates of the model parameters so long as the conditional expectation shown in (18) is correctly specified (Papke and Wooldridge, 1996). Standard errors are clustered at the census division (c) level to allow for arbitrary serial correlation within divisions.

We view the fractional probit as superior to other commonly used methods for handling fractional dependent variables. The simplest alternative is to assume that F is uniform which makes the crossing share, x_{ct}, linear in the parameters and therefore estimable using OLS. The problem is that the linear model can predict negative crossing fractions, which renders it inappropriate for counterfactual analysis. A second method assumes F is logistic. Applying the log odds transformation yields an equation that is linear in the parameters:

$$\ln\left(\frac{x_{ct}}{1 - x_{ct}}\right) = \theta_0 + \theta_1 \ln e_t + \theta_2 [\ln e_t]^2 + \theta_3 \ln D_c + \theta_4 \ln \left[P(g)_{ct}/W_{ct}\right] + \varepsilon_{ct}, \quad (19)$$

where ε is an error term added after the transformation. The log-odds method is often preferred because it ensures that predicted values of x_{ct} lie between zero and one. This method has the virtue of being estimable using linear methods. Among other benefits, this allows two-way clustering of the standard errors to account for the fact that each census division in month t sees the same real exchange rate. However, Papke and Wooldridge (1996) identify two critical defects. First, the dependent variable is undefined for $x_{ct} = 0$ and $x_{ct} = 1$. As we discuss in section 4.2, over half the ct combinations in our data have $x_{ct} = 0$ and these tend to occur in divisions that are far from the border, implying that the log-odds procedure is likely to induce selection bias. A second problem with the log odds specification is that it yields the conditional expectation of the log odds ratio, a variable that is not of direct interest. As Papke and Wooldridge (1996) show, one cannot simply plug the estimated θ estimated using (19) into the logistic function to recover the conditional expectation of x_{ct}.[27] Based on these arguments, we use the fractional probit estimation of equation 18 as our main specification and only report estimates from the log-odds method as a robustness check.

4.2 Data

The dependent variable is the crossing fraction, x_{ct}, which is defined as the number of car crossings, n_{ct}, from Census Division (CD) c in month t, divided by the number of potential crossings, N_{ct}. Potential trips, N_{ct}, are approximated as the population of the census division (Pop), multiplied by the number of cars per capita (CPC) in the province multiplied by the number of days in the month. Thus, the crossing fraction is given by

$$x_{ct} = \frac{n_{ct}}{N_{ct}} \approx \frac{\hat{n}_{ct}}{\text{Pop}_{ct} \times \text{CPC}_c \times 30}. \quad (20)$$

---

[27]Intuitively, this is because the log of the expectation is not equal to the expectation of the log.

We estimate $\hat{n}_{ct}$ using data from the International Travel Survey (ITS), which is filled out by travelers returning to Canada from trips abroad. The data were derived from questionnaires distributed from 1990 to 2010 that collected information on the nature and purpose of the trip, the dates on which travelers exited and entered Canada, and information on the Census Division in which the travelers reside and ports used to cross to the US. We retain the data on Canadian residents returning from the United States by car.[28] Appendix A.1 shows the sources for the variables in equation 20 and details how we construct $\hat{n}_{ct}$ by weighting the ITS responses using the port-level counts of all crossers, so as to make the sample representative at the monthly level as well as representative at each port of entry.

We measure D_c, the distance from census division c to the border, in two ways described in Appendix A.2. Our preferred form is the population-weighted median of the driving distances of all the subdivisions within a given CD.[29] In robustness checks we also measure D_c as the median driving time to these ports and as the average of driving distances to the five most-used ports. Gas prices, $P(g)$, are obtained for the largest city in each province. Median household income, our proxy for W_c, is available at the CD-level from the Canadian census in five year intervals.[30]

We present summary statistics of the data in Table 5. There are 63000 observations, each corresponding to a census division in a given month. The first column presents variable means across all observations, while the second column does so only for the subset of observations (39088) in which there was at least one car trip across the border in the given month. Conditioning on positive trips, Census Divisions tend to be closer to the border, and more populated. The large standard deviation for gas prices is mainly driven by temporal variation, whereas there is substantial cross-CD variation in household incomes, with the richest having incomes that are several times larger than the poorest.

4.3 Baseline Estimation

In this section we estimate the structural model implied by equation 18. The results using the fractional probit method of estimation are presented in Table 6. The first three columns use daytrips to construct the dependent variable, while the next three use overnight trips (defined as stays of two or more days). All estimated specifications include (unreported) month dummies to allow shocks to the mean of the $\zeta(i)$ distri-

---

[28]The survey began in 1990. We do not use information on US residents since the only information on their place of residence within the US is the state in which they live. This level of aggregation is too coarse to provide meaningful information on their distance to the border.

[29]Figure B.1 contains a map of a few CDs in south-eastern Ontario and shows the subdivisions (with black borders) within each CD (with red borders). Note the importance of using driving distance, as opposed to, say, great circle distance, given that there are a number of large lakes near the US–Canada border, as well as given the actual network of highways. Using a Euclidean distance metric would greatly understate the distance of a city such as Toronto from the border.

[30]Data details and sources are provided in the Appendix.

Table 5: Summary Statistics: 63000 Census Divisions-months

Variable	Mean	Mean|trips>0[a]	SD	Median	Min	Max
Driving Distance (km)	263.0	187.0	281.2	161.9	6.8	1877.1
Driving time (hrs)	3.7	2.6	3.9	2.2	0.2	26.7
Population (1000)	116.2	165.8	273.8	40.8	1.2	2667.9
Gasoline Price (c/L)	73.5	72.5	21.1	66.5	39.5	146.6
Median HH Income ($1000)	42.8	44.1	11.3	41.2	15.2	157.7
Cross-border trips (cars):						
Same-day	4093	6597	20229	0	0	456542
Overnight	1319	2126	4146	80	0	90662

[a] 39088 CD-months with at least one car trips across the border.

bution reflecting the seasonal pattern shown in Figure 1(b). The initial specification, shown in columns 1 and 4, assumes that travel costs are constant across time and depend only on the distance of the traveler's origin to the border. Columns 2 and 5 estimate the influence of gas prices and incomes. We do not report the specification imposing equal and opposite coefficients on $\ln P(g)$ and $\ln W$ because we found that the same day travel data strongly reject this constraint. Our preferred specification, shown in columns 3 and 6, adds fixed effects (FE) for each province and a dummy for travel after September 11, 2001. The province FEs capture differences in $B(\ln e)$ that result from unmeasured cross-state differences in product prices.[31] The post 9/11 dummy is designed to capture real and perceived increases in the cost of crossing the border following heightened security measures.

The results show that driving distance creates a strong disincentive to cross the border. This is especially the case for daytrips; distance is a weaker disincentive for those planning trips of a longer duration. The coefficient on the exchange rate variables indicate that a higher value of the real exchange rate (implying a weaker CAD) reduces the probability of cross-border trips. The coefficient on the second order term is positive for daytrips, implying that travelers' responsiveness to the real exchange rate decreases as its level rises. This is in accordance with the predictions of our model and is also consistent with the reduced form results of Table 3. Residents making daytrips are more likely to expand the bundle of goods that they purchase in the US when the exchange rate becomes more favorable. We do not observe the same behavior by overnight travelers: the coefficients on $[\ln e_t]^2$ are small and statistically insignificant in columns 4–6. This may be because overnight travelers purchase a standard bundle of goods in the US (hotel stays, vacations, restaurant meals etc.) without adjusting the scope of the bundle in accordance with relative prices. In other words, day trips are consistent with an explicit shopping motive, whereas overnight

---

[31]They can also account for differences in the mean idiosyncratic shocks due to different population densities on the US side of the border which affect the likelihood of visiting friends and relatives.

Table 6: Fractional Probit estimation of crossing fractions (x_{ct})

Length of stay:	Daytrip			Overnight		
θ_0: constant	-0.23	9.80^a	4.42^a	-2.68^a	-4.59^a	-5.20^a
	(0.31)	(2.94)	(1.52)	(0.07)	(0.57)	(0.99)
θ_1: $\ln e_t$ [RER]	-0.44^a	-0.77^a	-0.65^a	-0.61^a	-0.92^a	-0.75^a
	(0.10)	(0.14)	(0.13)	(0.12)	(0.13)	(0.12)
θ_2: $(\ln e_t)^2$	0.39	1.24^a	0.82^b	-0.09	0.27	-0.17
	(0.34)	(0.33)	(0.33)	(0.30)	(0.28)	(0.24)
θ_3: $\ln D_c$ [distance]	-0.58^a	-0.58^a	-0.52^a	-0.14^a	-0.14^a	-0.12^a
	(0.06)	(0.06)	(0.04)	(0.01)	(0.01)	(0.01)
$\ln P(g)_{ct}$ [gas price]		-0.35^a	-0.07		-0.56^a	-0.13^a
		(0.09)	(0.05)		(0.04)	(0.02)
$\ln W_{ct}$ [income]		-0.80^a	-0.42^a		0.40^a	0.29^a
		(0.27)	(0.14)		(0.06)	(0.09)
New Brunswick			0.40^a			0.00
			(0.14)			(0.06)
Quebec			-0.46^a			-0.15^b
			(0.08)			(0.07)
Ontario			-0.23^b			0.07^b
			(0.11)			(0.03)
Manitoba			-0.33^a			0.03
			(0.13)			(0.04)
Saskatchewan			-0.45^a			-0.15^a
			(0.10)			(0.04)
Alberta			-0.48^a			-0.18^a
			(0.11)			(0.05)
Post-911			-0.14^a			-0.14^a
			(0.03)			(0.03)
R^2	0.24	0.29	0.53	0.05	0.07	0.08
AIC	1935.18	1908.66	1778.11	629.59	626.92	636.59

Standard errors clustered by census-division. British Columbia is the omitted province. Regressions include month fixed-effects. ^{c} p<0.1, ^{b} p<0.05, ^{a} p<0.01. $N = 63000$

24

trips are not.

Gas prices do not have a significant effect on day trips but have a negative effect on overnight trips. One explanation may be that, while higher fuel costs discourage long car trips, they may also encourage some travelers to make same day trips explicitly for the purpose of purchasing gasoline, since the absolute savings on gas purchases in the US increase as gas prices rise. This effect would tend to offset the higher cost of driving for same-day trips, but it is unlikely to do so for overnight trips.

The coefficient in column 6 is about the same as the distance coefficient.[32] Income effects are strongly negative for day-trippers. This runs counter to what would be expected if income mattered just because it affects the fuel cost to income ratio in τ. Our model assumes a constant marginal utility of income across all individuals. One interpretation of the results is that richer households are less motivated by the savings to be had from cross-border shopping. For overnight trips income effects are positive. In column 6 the regression does not reject the restriction of equal and opposite effects for gas prices and incomes that is predicted by the transport cost function shown in equation 4.

The province fixed effects capture the underlying propensity of travelers from each province to cross the border, after accounting for exchange rate, distance, and income effects. They may reflect the presence of large cities, and the provision of goods and services that may be sought by Canadians, such as gasoline, outlet malls, casinos, airports etc. It is not surprising that British Columbia (the omitted category) and Ontario have higher fixed effects than Alberta and Saskatchewan. There are population centers near the border in Washington, Michigan, and New York but not in Montana and North Dakota.

The downward shift in travel to the US following September 2011 corroborates the finding of Ferris (2010) who estimates a linear reduced form regression using aggregate monthly same-day travel for Canada from 1972 to April 2009. The distance equivalent of 9/11 is given by $\exp(0.14/0.52) - 1 = 0.31$. Thus, the extra costs of crossing the border in the years since 9/11 corresponds to a 31% increase in distance. Alternatively, using a counterfactual calculation of the kind described in Section 4.6, we find a total reduction of 32% in travel attributable to 9/11. Remarkably, given the many differences in method, Ferris (2010) reports a 29% annual reduction.

We illustrate the magnitudes of the effects we have estimated by showing how predicted crossing shares respond to changes in our key explanatory variables. This is important since the estimated coefficients are scaled by the unobserved σ parameter. Moreover, the effects of the RER and distance have to pass through the nonlinear $\Phi()$ function to determine the predicted crossing share.

We show the relationship between the crossing fraction and the real exchange rate for specific distances from the border in Figure 5. This figure is based on the specification in column 3 for Table 6 (adjusting using the coefficients on the Ontario, post 9/11, and April dummy variables). Each curve corresponds to a census division

---

[32]The larger negative effect in column 5 is mainly attributable to the absence of the 9/11 dummy.

in Southern Ontario.[33] The curves show that the convexity in the B function carries over to the log crossing function. Thus, the elasticity of crossing is larger in absolute value when the home currency is strong. Furthermore, the elasticity of crossing implied by the model is larger at greater distances from the border. We can see this in the figure as the curve for Toronto is steeper (which corresponds to greater elasticity since both axes are drawn on a log scale) than that for Niagara.

The main determinant of travel costs is distance to the border. Figure 6 shows the steep decline of crossing fractions associated with increased driving distances. The curve graphs the average of the predicted shares (in percent) that would cross from each Ontario census division during the sample period (1990–2010). The circles show actual crossing fractions averaged over the same period. The model fits the data well, further supporting the validity of the linear-in-logs approximation of the travel cost function.

Divisions further from the border than Toronto (about 90 miles) have predicted and actual crossing rates below 0.1%. This means that on any given day there is a less than 1 in 1000 chance for a car to be driven across the US border on a daytrip. In contrast, communities closer than Niagara (15 miles) have crossing rates that are more than an order of magnitude higher. The evidence of porous borders is consistent with market segmentation because of the combination of strong distance effects and the fact that the majority live more than 80 miles from the border.

4.4   Robustness to specification changes

In this section we examine the robustness of our results to different specifications and variable definitions. The results are in Table 7. We use the set of controls corresponding to columns 3 and 6 of Table 6. We prefer this specification since adding province fixed-effects improves the fit of the model considerably, compared with columns 2 and 5.

The first two columns of Table 7 present results using the log-odds model depicted in equation 19. The remaining columns use the fractional probit model, but use different measures of the costs of travel. In columns 3 and 4 we use the driving *time* to the border from each Census Division, instead of the driving distance. This exploits the information Google keeps about differences in average driving speeds relevant for different subdivisions.[34] We add 26 minutes to the driving time to account (very roughly) for border wait times.[35] In columns 5 and 6 we use our secondary measure of distance (detailed in the Appendix). Relative to the primary measure used in

---

[33]These census divisions—Niagara, Hamilton and Toronto—are CDs 26, 25 and 20 respectively in Figure B.1. The nearest border crossings are on the Niagara river at the bottom right of the map.

[34]The average speed is 70 km/hour with a 5%–95% range of 51–84 km/hour.

[35]Wait time data is not systematically available across Canada during our estimation period. The 26 minutes figure is the median wait for all travelers entering the United States during the hours of 7 AM and 12 PM at the two largest ports in British Columbia, using daily data from 2006 to 2010. Data on wait times were obtained from the Whatcom Council of Governments.

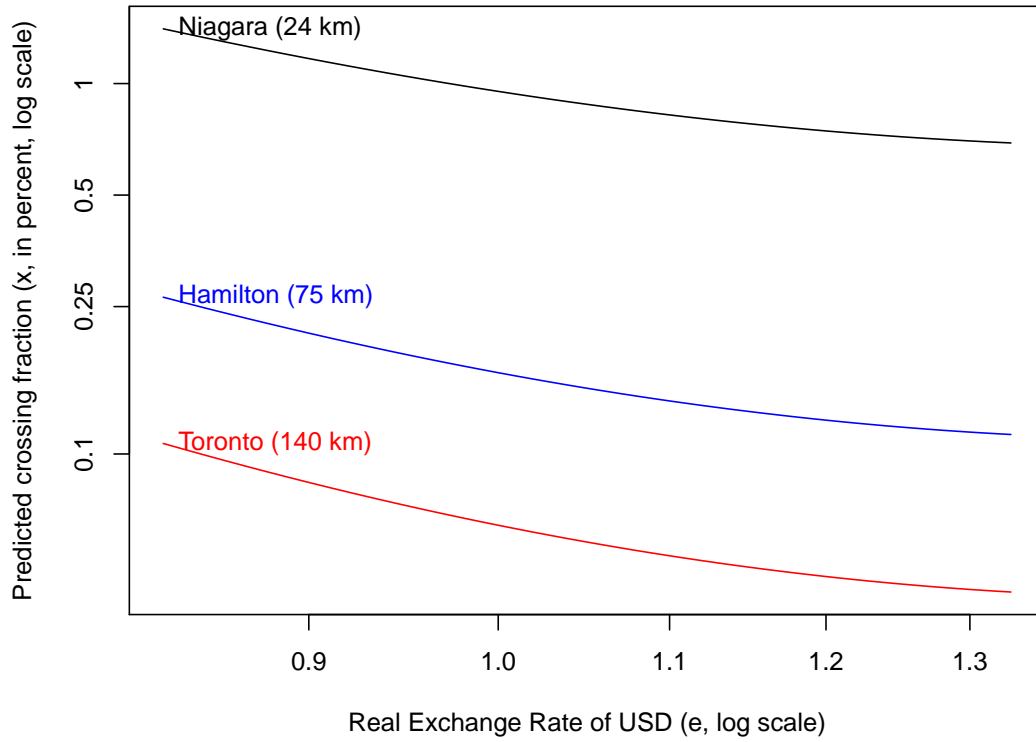Figure 5: The crossing fraction declines with strength of foreign currency

Niagara (24 km)

Hamilton (75 km)

Toronto (140 km)

Predicted crossing fraction (x, in percent, log scale)

Real Exchange Rate of USD (e, log scale)

Figure 6: The crossing fraction declines with distance to the border

Crossing fraction (in %, log scale)

Niagara

Hamilton

Toronto

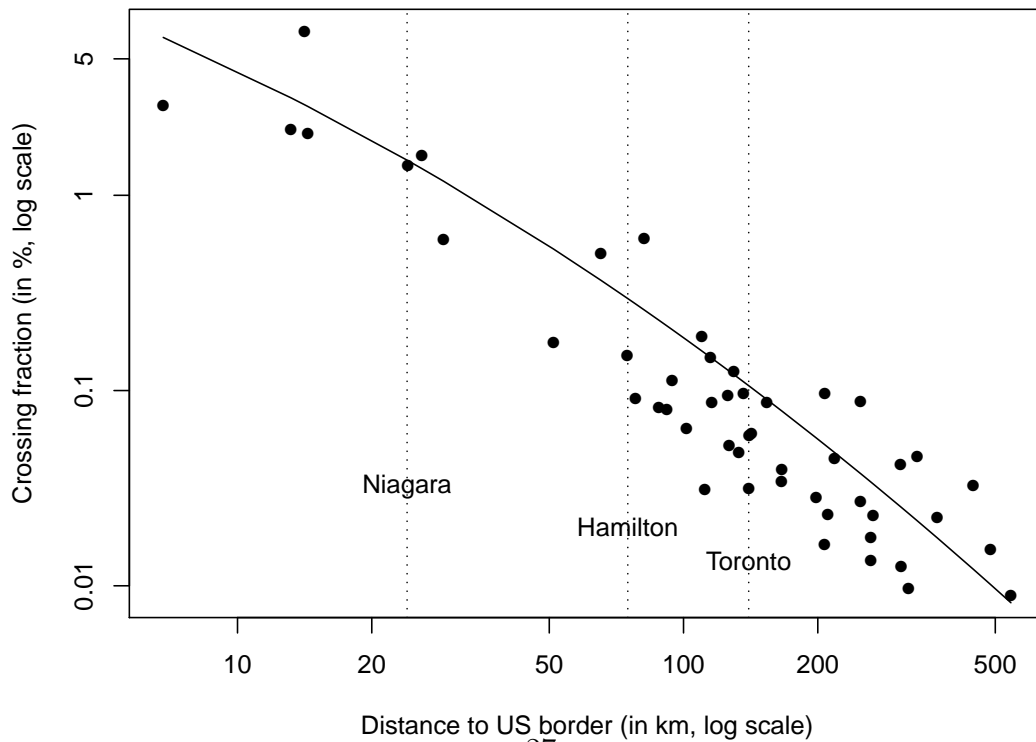Distance to US border (in km, log scale)

Table 6, it has the advantage of taking into account not just the nearest port but the five ports that residents of the CD use most frequently. It has the disadvantage of using the geographic center of the CD as the origin point, which exaggerate distances severely for some large Divisions.

Table 7: Alternative specifications of regression and travel costs

Method:	Log Odds (OLS)		Fractional Probit			
Stay:	Daytrip	Overnight	Daytrip	Overnight	Daytrip	Overnight
θ_0: constant	25.40^a	-2.28	5.07^a	-5.22^a	10.33^a	-4.61^a
	(3.14)	(1.87)	(1.82)	(1.01)	(2.47)	(1.08)
θ_1: $\ln e_t$	-1.55^a	-2.00^a	-0.65^a	-0.75^a	-0.65^a	-0.75^a
	(0.25)	(0.15)	(0.13)	(0.12)	(0.13)	(0.12)
θ_2: $(\ln e_t)^2$	3.73^a	0.20	0.93^a	-0.15	1.03^a	-0.16
	(0.73)	(0.63)	(0.33)	(0.24)	(0.32)	(0.24)
θ_3: $\ln$ dist. or time	-1.14^a	-0.28^a	-0.89^a	-0.19^a	-0.56^a	-0.14^a
	(0.07)	(0.04)	(0.06)	(0.02)	(0.05)	(0.02)
$\ln P(g)_{ct}$	-0.15	-0.42^a	-0.05	-0.13^a	-0.03	-0.13^a
	(0.13)	(0.08)	(0.05)	(0.02)	(0.05)	(0.02)
$\ln W_{ct}$	-2.41^a	-0.19	-0.64^a	0.26^a	-0.94^a	0.25^b
	(0.30)	(0.17)	(0.18)	(0.10)	(0.24)	(0.10)
Post-911	-0.25^a	-0.18^a	-0.13^a	-0.14^a	-0.12^a	-0.14^a
	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Observations	24232	33771	63000	63000	63000	63000
R^2	0.51	0.28	0.57	0.08	0.51	0.08
AIC	83374.53	93686.21	1792.09	635.49	1873.42	636.59

Standard errors clustered by census-division except cols. (1)–(2) where SEs also clustered by month-year. Regressions include month, province FEs. ^{c} p<0.1, ^{b} p<0.05, ^{a} p<0.01. Driving time in cols. (3)–(4); port-use weighted average distances in cols. (5)–(6).

Our chief results on exchange rate and distance effects hold in all specifications. The positive second-order effect for exchange rates continues to hold for daytrips and is insignificant for overnight trips. The cost of traveling to the border, whether measured in terms of distance or time, has a negative and strongly significant effect on the probability of crossing the border; much more so for daytrips than overnight ones.

There are a number of other robustnesss checks that we conducted, the results of which are contained in the supplementary file (see Table 5 in that document). We

28

included a quadratic term for distance but it was not statistically significant nor did it contribute significantly to the fit of the model. We also dropped observations where the drive times were extraordinarily long (more than 12 hours in one specification and more than 3 in another). The coefficients of the variables of interest in these specifications hardly change.

We examined whether commuters—residents of Canada who work in the United States—impact our results, since these travelers cross the border daily regardless of the exchange rate, and therefore are not the type of travelers that the model considers. Although commuters constitute less than 6% of travelers (as can be seen in Table 1), they make up a disproportionate share of travelers in certain census divisions.[36] We re-estimated the regressions dropping the census divisions where commuters made up 10% or more of travelers and found very similar results.

The real exchange rate and distance terms enter the crossing equation 16 additively. This suggests a simple falsification test. If the model is correctly specified, there should be no significant interaction between exchange rates and distance. When we add such an interaction term to the estimating equation, it is not statistically significant and it does not improve the R^2 relative to the equation implied by our model.

4.5 Implied travel cost estimates

One very useful way to evaluate our coefficients is to determine what they imply about travelers' willingness to trade off savings from cross-border shopping versus travel costs. Re-expressing the net benefits of crossing, $v_X - v_S$, using the parametric forms for $B(\ln e)$ and $\ln \tau(D)$ we obtain

$$v_X - v_S = B - \ln \tau = \beta_0 + \beta_1 \ln e + \beta_2 [\ln e]^2 - \gamma_0 - \gamma_1 \ln(D) - \gamma_2 \ln(P(g_c)/W_c).$$

Totally differentiating by e and D yields

$$d(v_X - v_S) = \frac{\partial(v_X - v_S)}{\partial e} de + \frac{\partial(v_X - v_S)}{\partial D} dD = 0.$$

Rearranging,

$$\frac{de/e}{dD/D} = \frac{\gamma_1}{\beta_1 + 2\beta_2 \ln e}$$

We do not observe β_1, β_2, or γ_1 but we do estimate $\theta_1 = \beta_1/\sigma$, $\theta_2 = \beta_2/\sigma$, and $\theta_3 = -\gamma_1/\sigma$. Plugging in these estimates, canceling out the σ, we obtain $(de/e)/(dD/D)$ as a function of the estimated parameters and the level of the real exchange rate. This calculation tells us the percent change in the real exchange rate required to compensate someone for a percentage increase in the distance or duration of the cross-border trip.

---

[36]The CD with the highest fraction commuters is Essex (35%), just across the border from Detroit. The next highest CD has just 13% commuters.

To obtain the change in expenditure, X, that would be required as compensation for the trip we note that expenditure in CAD is given by e times expenditure in USD. Holding USD-denominated expenditure constant, we have $dX/X = de/e$. We thereby arrive at the following formula for the travel cost:

$$\frac{dX}{dD} = \frac{-\hat{\theta}_3}{\hat{\theta}_1 + 2\hat{\theta}_2 \ln e} \left[\frac{X}{D}\right].$$

At the 2010 average real exchange rate of $e = 0.8846$, the first factor is given by -0.611 for distance (using $\hat{\theta}$ from column (3) of Table 6) and -1.02 for time (based on column (3) of Table 7). The second factor shown in brackets, X/D, is less straightforward to determine. We use the car-weighted median distance (or duration) of a round trip for daytrippers for D. This works out to 36 miles or 1.8 hours (including a 26 minute border wait in each direction). For X we use \$51, the 2010 median expenditure (in USD) of daytrippers who spent a positive amount, as calculated from the ITS.

Plugging in these values we obtain a travel cost of US \$0.87 per mile or \$29.69 per hour. These figures are in line with the \$0.89 per mile reimbursement rate for government travel within Ontario,[37] and 2010 Canadian median hourly wages of US \$23.34 per hour.[38] Using means instead of medians for D (56 miles) and X (\$152) leads to travel cost estimates of \$1.66/mile and \$68.34/hour. These travel cost estimates are on the high end of the range reported in the literature on shopping within national markets.[39]

Table 8: Travel cost estimates

Distribution	$d \ln e/d \ln D$	US \$/mile		$d \ln e/d \ln T$	US \$/hour	
		median	average		median	average
$\zeta(i) \sim$ Normal	-0.611	0.87	1.66	-1.023	29.69	68.34
$\zeta(i) \sim$ Logistic	-0.618	0.88	1.68	-1.124	32.63	75.10
$\zeta(i) \sim$ Gumbel	-0.597	0.85	1.63	-0.946	27.47	63.23

The normality assumption for individual heterogeneity can be replaced with assumptions of logistic or Gumbel distributions. While each distributional assumption leads to different estimated coefficients, their relative values change very little. Thus,

---

[37]See `http://www.njc-cnm.gc.ca/directive/travel-voyage/s-td-dv-a2-eng.php`. All Canadian dollar figures in this section were converted to USD using the 2010 average exchange rate of 1.03 CAD/USD.

[38]See CANSIM Table 2820070.

[39]Chiou and Muehlegger (2008) estimate that consumers would be willing to travel to a location 2.7 miles further away to save \$1 on cigarettes. This equates to a travel cost of 18.5 cents per mile. Manuszak and Moul (2009) estimate a marginal cost of around 50 cents per mile for consumers of gasoline in the Chicago area. Thomadsen (2005) estimates a travel cost of around \$1.50 per mile for consumers choosing fast food restaurants in Palo Alto.

we see in Table 8 that $-\hat{\theta}_3/(\hat{\theta}_1 + 2\hat{\theta}_2 \ln e)$ evaluated in 2010 ranges from -0.597 (Gumbel) to -0.618 (logistic), with the normal distribution in the middle. The monetary travel costs differ by only a few cents per mile. We are reassured that the results are not fragile to specific distributional assumptions. This is not to say that all distributions would yield the same results. However, it seems reasonable to infer that the travel cost results are unlikely to vary much so long as a bell-shaped distribution for heterogeneity is assumed.

4.6 Consumer welfare effects of policy changes

The net benefits accruing to crossers, shown in equation 10, can be expressed as the product of two factors:

$$G_c = \underbrace{(B - \ln \tau_c + \mathbb{E}[\zeta \mid \zeta > -B + \ln \tau_c])}_{\text{Average crosser's gain}} \underbrace{F[B - \ln \tau_c]N_c}_{\text{Number of crossers}}. \qquad (21)$$

To a first approximation, the percentage change in crosser welfare brought about by a change in the determinants of $B - \ln \tau_c$ will be given by the sum of the percentage changes in the number of crossers, n_c, and the average gain each crosser expects to obtain, G_c/n_c. We therefore quantify these components separately.[40]

With ζ distributed $\mathcal{N}(\mu, \sigma^2)$, we can compute the average crosser's gain as

$$G_c/n_c = (B - \ln \tau_c) + \mu + \sigma \frac{\phi[(\mu + B - \ln \tau_c)/\sigma]}{\Phi[(\mu + B - \ln \tau_c)/\sigma]} = \sigma \left(\mathbf{Z}\hat{\boldsymbol{\theta}}_c + \frac{\phi[\mathbf{Z}\hat{\boldsymbol{\theta}}_c]}{\Phi[\mathbf{Z}\hat{\boldsymbol{\theta}}_c]} \right),$$

where $\mathbf{Z}_c$ is the vector of explanatory variables and $\boldsymbol{\theta}$ is the coefficient vector. The second equality comes from $(B - \ln \tau_c + \mu)/\sigma = \mathbf{Z}\hat{\boldsymbol{\theta}}_c$ (the prediction index obtained from the fractional probit regressions). Without being able to identify σ, levels of G_c/n_c cannot be determined but we can determine the percentage change resulting from any contemplated change in the $\mathbf{Z}_c$ vector. To quantify the aggregate effect of policy changes, it is necessary to aggregate over the effects at each census division, multiplying by N_c to give greater weight to larger divisions.

In Tables 9 and 10 we show the effect of two possible changes. Table 9 shows the effect, in two different years, on the number of cross-border trips from a decrease in the real exchange rate of 10%. This is equivalent to a strengthening of the Canadian Dollar. These estimates were derived by calculating, for each month in the corresponding year, the number of car trips from each Census Division had the RER in that month been 10% lower than its actual value. These counterfactual values were then aggregated across all census-divisions in the province and compared to the predicted values using the specification of Column 2 in Table 6. The years that we

---

[40]The difference between their sum and the total welfare effect is negligible in the experiments we conduct.

analyze are 2000—when the Canadian Dollar was near its all-time low—and 2007, when it reached its highest value in decades.

Table 9 reveals differences in exchange rate elasticities across locations and time. As implied by equation 14 and the normal distribution for shocks, the elasticities are larger for census divisions further from the border. At a given point in time, an appreciation of the RER shifts up the benefits of crossing for all census-divisions and therefore for all provinces, leading to proportional rises in the elasticities from 2000 (when the CAD was weak) to 2007 (when it reached a high). The elasticities rise due to the convex relationship between the crossing benefits and the log RER.

The implied crossing elasticities can be compared to those obtained in the trade literature to gain perspective on the extent of consumer arbitrage. When the Canadian dollar is weak, the Canada-wide elasticity of 0.99 is similar to the estimated elasticities of Blonigen and Wilson (1999) for Canada-US trade in goods.[41] In years with strong CAD, our results suggest elasticities for travel that are twice as large as those typically observed for goods. Stronger travel effects make some sense in light of the fact that travelers can respond immediately to price differences whereas traders need to make various up-front investments in marketing, distribution, and logistics.

The model indicates that the home appreciation gives rise to aggregate gains of 26.84% in 2007. Most of this, 24.4%, comes from increased propensity to cross. Welfare changes for the average crosser contribute 2.09%.[42] The gains to the average crosser are approximately twice as high when the appreciation starts from an already strong Canadian dollar. The biggest percentage gains to the average crosser are obtained in census divisions close to the border.

Table 9: Impact of a 10% Canadian dollar appreciation on same-day travel

Year:	2000		2007	
	% Δ Trips (n_c)	% Δ Gains (G_c/n_c)	% Δ Trips	% Δ Gains
Canada	9.90	0.91	24.44	2.09
New Brunswick	7.75	1.06	18.87	2.44
Quebec	12.31	0.89	30.77	2.05
Ontario	9.79	0.92	24.21	2.11
Toronto (140 km)	13.28	0.89	33.12	2.06
Hamilton (75 km)	12.01	0.95	29.93	2.18
Niagara (24 km)	9.84	1.07	24.03	2.47
Manitoba	12.02	0.90	30.07	2.07
Saskatchewan	12.96	0.81	32.67	1.87
Alberta	14.13	0.77	36.26	1.76
British Columbia	10.16	1.00	24.76	2.33

---

[41]Blonigen and Wilson's average elasticity is 0.81. Two thirds of 146 estimates less than one.

[42]The remainder, 0.31%, is attributable to the weighted product of the changes.

In Table 10 we show the effect of increasing wait times at the border. We use the specification from Column 3 of Table 7. This specification had assumed a wait time of 26 minutes at the border. In our counterfactual experiment we double this to 52 minutes.[43] This naturally decreases the likelihood of cross-border trips by Canadians. However, now there are significant differences across provinces, and almost no variation over time. The smallest effects of the increased wait times are in the provinces of Alberta and New Brunswick. These provinces do not have large cities close to the border. Since the wait time is incurred by all travelers, those driving longer distances pay a proportionately lower cost. By contrast, a province such as Ontario has a large population located very close to the border and therefore our model predicts a very large decrease in trips for a given increase in wait times.[44] The welfare losses to the average crosser from increased wait times range from under 2% for Alberta to almost 10% for Niagara county. The predicted changes in crossings and welfare do not change much over time since the effect of travel costs is independent of the value of the RER.

Table 10: Impact of a doubling of border wait times on same day trips

Year:	2000		2007	
	% Δ Trips	% Δ Gains	% Δ Trips	% Δ Gains
Canada	-55.20	-4.64	-54.68	-4.63
New Brunswick	-49.83	-5.22	-49.23	-5.18
Quebec	-54.20	-4.55	-53.95	-4.58
Ontario	-58.39	-5.08	-57.52	-5.03
Toronto (140 km)	-43.51	-4.11	-42.95	-4.15
Hamilton (75 km)	-52.23	-6.18	-52.15	-6.18
Niagara (24 km)	-62.47	-9.75	-62.36	-9.75
Manitoba	-51.56	-4.00	-51.59	-4.02
Saskatchewan	-51.49	-2.27	-51.32	-2.24
Alberta	-48.97	-1.68	-49.06	-1.66
British Columbia	-53.67	-6.10	-53.42	-6.21

One final counterfactual we consider is to "turn off" the estimated 9/11 effect. As we reported earlier, the post-9/11 period had a 32% reduction in same-day crossings relative to what the model would have predicted based on the evolution of the real exchange rate, gas prices and incomes. The average crosser incurs a 3.4% reduction in welfare.

---

[43]Note that this increase in wait times needs to occur for exogenous reasons such as reduced staffing at the border or an increase in the time taken to process vehicles. Increases in wait times due to an increased number of cars arriving at the border will confound our predictions.

[44]See Figure B.2 in the Appendix to understand the differences in the geographical distribution of population across Canadian provinces.

5   Conclusion

In this paper we have addressed the extent of market segmentation by studying the behavior of US-Canada border crossings and their relationship to arbitrage gains. Our findings support the hypothesis of a porous border with partial market segmentation and reject the idea of isolated markets. Crossings are heavily influenced by arbitrage opportunities and the exchange rate elasticity of crossings increases with the appreciation of the domestic currency. These results are not in conflict with previous evidence of pricing-to-market across borders so long as consumer response to arbitrage opportunities is finite. Two forces in our model prevent travelers from fully arbitraging the price differences between the two countries. First, consumers face large marginal travel costs. Our estimates range between \$30 and \$68 per hour (or \$0.87 and \$1.66 per mile). Second, individuals are heterogeneous. While the majority of border crossers live less than 18 miles from the border, the majority of Canadians reside more than 81 miles away.

References

Asplund, M., R. Friberg, and F. Wilander (2007, February). Demand and distance: Evidence on cross-border shopping. *Journal of Public Economics 91*(1-2), 141–157.

Baggs, J., E. Beaulieu, L. Fung, and B. Lapham (2010). Exchange rate movements and firm dynamics in canadian retail industries. Working Paper.

Blonigen, B. and W. Wilson (1999). Explaining armington: what determines substitutability between home and foreign goods? *Canadian Journal of Economics*, 1–21.

Boivin, J., R. Clark, and N. Vincent (2011). Virtual borders: Online nominal rigidities and international market segmentation. Forthcoming, *Journal of International Economics*.

Brockmann, D., L. Hufnagel, and T. Geisel (2006). The scaling laws of human travel. *Nature 439*(7075), 462–465.

Burstein, A. and N. Jaimovich (2009). Understanding movements in aggregate and product-level real-exchange rates. Working Paper.

Campbell, J. R. and B. Lapham (2004, September). Real exchange rate fluctuations and the dynamics of retail trade industries on the u.s.-canada border. *American Economic Review 94*(4), 1194–1206.

Chiou, L. and E. Muehlegger (2008). Crossing the Line: Direct Estimation of Cross-Border Cigarette Sales and the Effect on Tax Revenues. *The B.E. Journal of Economic Analysis & Policy (Contributions) 8*(48).

Di Matteo, L. and R. Di Matteo (1993). The determinants of expenditures by canadian visitors to the united states. *Journal of Travel Research 31*(4), 34–42.

Di Matteo, L. and R. Di Matteo (1996). An analysis of canadian cross-border travel. *Annals of Tourism Research 23*(1), 103–122.

Dornbusch, R., S. Fischer, and P. A. Samuelson (1977, December). Comparative advantage, trade, and payments in a ricardian model with a continuum of goods. *American Economic Review 67*(5), 823–39.

Engel, C. and J. H. Rogers (1996, December). How wide is the border? *American Economic Review 8*(5), 1112–25.

Ferris, J. (2010). Quantifying non-tariff trade barriers: What difference did 9/11 make to canadian cross-border shopping? *Canadian Public Policy 36*(4), 487–501.

Ferris, J. S. (2000). The Determinants of Cross Border Shopping: Implications for Tax Revenues and Institutional Change. *National Tax Journal 53*(4), 801–824.

Ford, T. (1992). International outshopping along the canadaunited states border: The case of western new york. Working Paper.

Goldberg, P. K. and M. M. Knetter (1997, September). Goods prices and exchange rates: What have we learned? *Journal of Economic Literature 35*(3), 1243–1272.

González, M., C. Hidalgo, and A. Barabási (2008). Understanding individual human mobility patterns. *Nature 453*(7196), 779–782.

Gopinath, G., P.-O. Gourinchas, C.-T. Hsieh, and N. Li (2011). International prices, costs and mark-up differences. *American Economic Review 101*(6), 2450–86.

Gorodnichenko, Y. and L. L. Tesar (2009, January). Border effect or country effect? seattle may not be so far from vancouver after all. *American Economic Journal: Macroeconomics 1*(1), 219–41.

Knight, B. G. and N. Schiff (2010, January). Spatial competition and cross-border shopping: Evidence from state lotteries. NBER Working Papers 15713, National Bureau of Economic Research, Inc.

Manuszak, M. D. and C. C. Moul (2009). How far for a buck? tax differences and the location of retail gasoline activity in southeast chicagoland. *The Review of Economics and Statistics 91*(4), 744–765.

Papke, L. and J. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics 11*(6), 619–632.

Porter, D. (2009, July). Looney's leap: Mind the (price) gap. Special report, BMO Capital Markets Economics.

Thomadsen, R. (2005). The effect of ownership structure on prices in geographically differentiated industries. *RAND Journal of Economics 36*(4), 908–929.

Appendices

A Data construction

A.1 Crossing fractions

Each observation in the ITS data is a questionnaire filled out by a Canadian resident returning to Canada from a trip to the US. This includes people who enter by car, bus, train, air, foot, boat etc. A maximum of one questionnaire is given to each traveling party. We keep only those observations where the traveling party exited and re-entered Canada by car. We also restrict the sample to people who reside in one of the 7 provinces that share a land border with the United States: New Brunswick, Quebec, Ontario, Manitoba, Saskatchewan, Alberta and British Columbia. This leaves us with 646,223 questionnaires over 20 years (1990–2009).

These questionnaires are handed out at the various border crossing ports, but not in a representative manner (either across ports, or across months of the year for a given port). Therefore, Statistics Canada has assigned weights to each questionnaire in order to address non-representative sampling and non-response. Applying these weights makes the data representative at the annual level for each port-factor-group (PFG).[45] However, we also want to exploit within-year variation in the exchange rate, and therefore require representative data on monthly travel. More importantly, we also require representative data at the level of each Census Division (CD) in order to examine the effect of the geographic distribution of residents on their propensity to travel. In order to construct data that are representative for each CD in each month, we construct our own weights.

Each questionnaire is associated with a particular CD and a port of entry into Canada. It also provides the month of travel and the length of the trip.[46] Therefore, each observation is CD-port-month-trip length combination. For notational clarity, we suppress subscripts for month and trip length. Define r_{cp} as the number of respondents from census division c passing through port of entry p. Define r_c as total

---

[45]A PFG is a combination of a port of entry, length of stay, and mode of travel. For example, the PFG defined as Blaine–1 night–automobile is the set of traveling parties that entered Canada at the Blaine, BC port, having claimed to have spent one night in the US.

[46]We construct the length of trip from the reported dates of exit and entry. We assign the month of travel as the calendar month in which the vehicle entered Canada.

respondents (across all CDs) at port p: $r_p = \sum_c r_{cp}$. Let n_p be the true number of crossers at port p which we obtain on a monthly basis from Cansim Table 427-0002. To estimate crossings by census division, $\hat{n}_c$, we first allocate n_p across census divisions using shares of response counts: $\hat{n}_{cp} = (r_{cp}/r_p)n_p$. Alternatively, one can think of this as the weighted sum of questionnaire respondents, r_{cp}, where weights are given by n_p/r_p, the number of actual crossers per respondent at a given port-month. Summing over all p for a given c we obtain $\hat{n}_c = \sum_p r_{cp} n_p / r_p$.

The estimated crossing fraction is given by dividing $\hat{n}_c$ by our estimate of cars at risk, $N_c = \text{Pop}_{ct} \times \text{CPC}_c \times 30$. Census division populations, Pop_{ct}, are available annually from Cansim Table 051-0034, provided by Statistics Canada. Car registration data used for generating CPC_c come from Statistics Canada publication 53-219-XIB ("Road Motor Vehicle Registrations 1998").

A.2   Driving distances and times to the border

We calculate the distance from each Canadian Census Division (similar to a US county) to the nearest ports D_c using two methods. The primary method takes advantage of geographically detailed information at the level of Census Subdivisions (similar to US Census Tracts). The 250 CDs have an average of 20 subdivisions. We obtained Subdivision centroid information from the Standard Geographical Classification of 2001 and used Google's driving distance application to measure the road distance and time from each centroid to the nearest crossing port. We obtained two measures: the median and the average distances for each CD. These two metrics are very similar for the majority of CDs except for two CDs in Ontario where the average distance is heavily influenced by outlier (low population and high distance to the border) subdivisions. We therefore used medians in our estimations. The results using averages do not differ much in terms of exchange rate or distance elasticities but the province and income effects are influenced by the two outliers.

The secondary method of calculating distances (employed in columns (5) and (6) of Table 7) takes into account the fact that crossers from a given census division do not always use the same port. At the CD level, we know shares of crossers from each CD that cross at 102 different ports. We use the average shares of the top 5 ports over the 1990 to 2010 period to construct weighted average distance and time from the CD's geographic centroid. This measure generates several outliers in large CDs that have centroids that are far from the border but populations that are concentrated close to the border.

A.3   Prices, exchange rates, and incomes

Exchange rates obtained from Pacific Exchange Rate Service (`fx.sauder.ubc.ca`). The US Consumer Price Index is the US city average for all items and all urban consumers, not seasonally adjusted (Series ID CUUR0000SA0 from `bls.gov/cpi#`

`data`). Canadian prices are from CANSIM Table 3260020, 2009 basket, all items. We choose July 1993 as the base period because in that month the nominal exchange rate was equal to the annual purchasing power parity rate provided by the OECD and thus the RER was approximately 1. Prices for regular unleaded gasoline at self service filling stations are obtained from CANSIM Table 3260009 for a major urban centre for each of the border provinces. We obtained median household income from the CHASS Canadian Census Analyser for the years 1991, 1996, 2001, and 2006. We linearly interpolated and extrapolated around July of each census year to obtain the monthly data from 1990 to 2010.
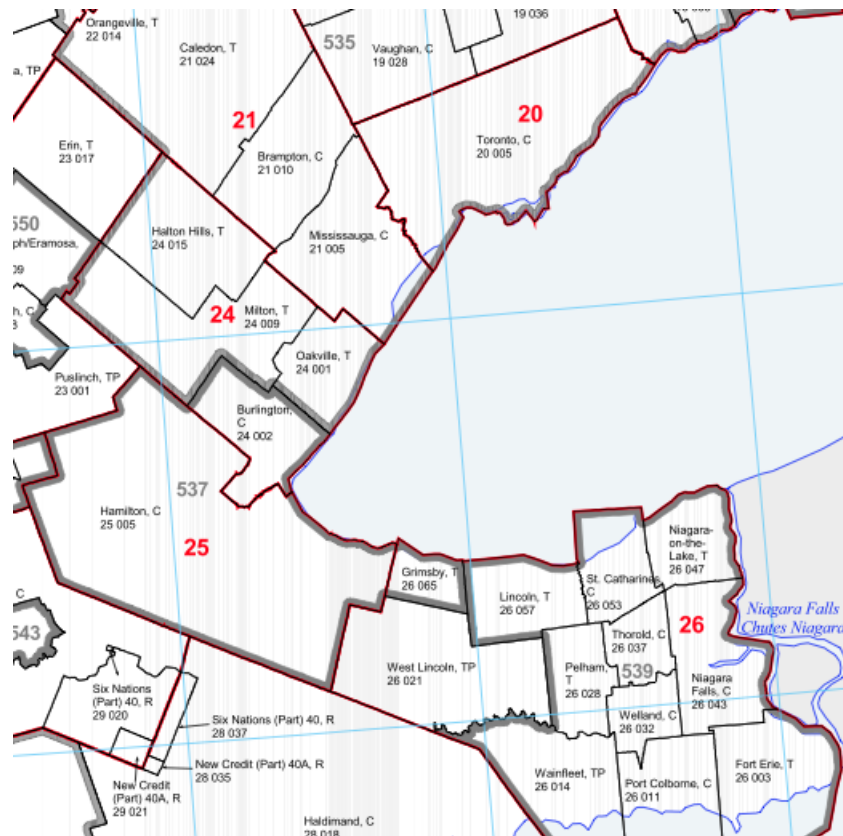
B Additional Figures

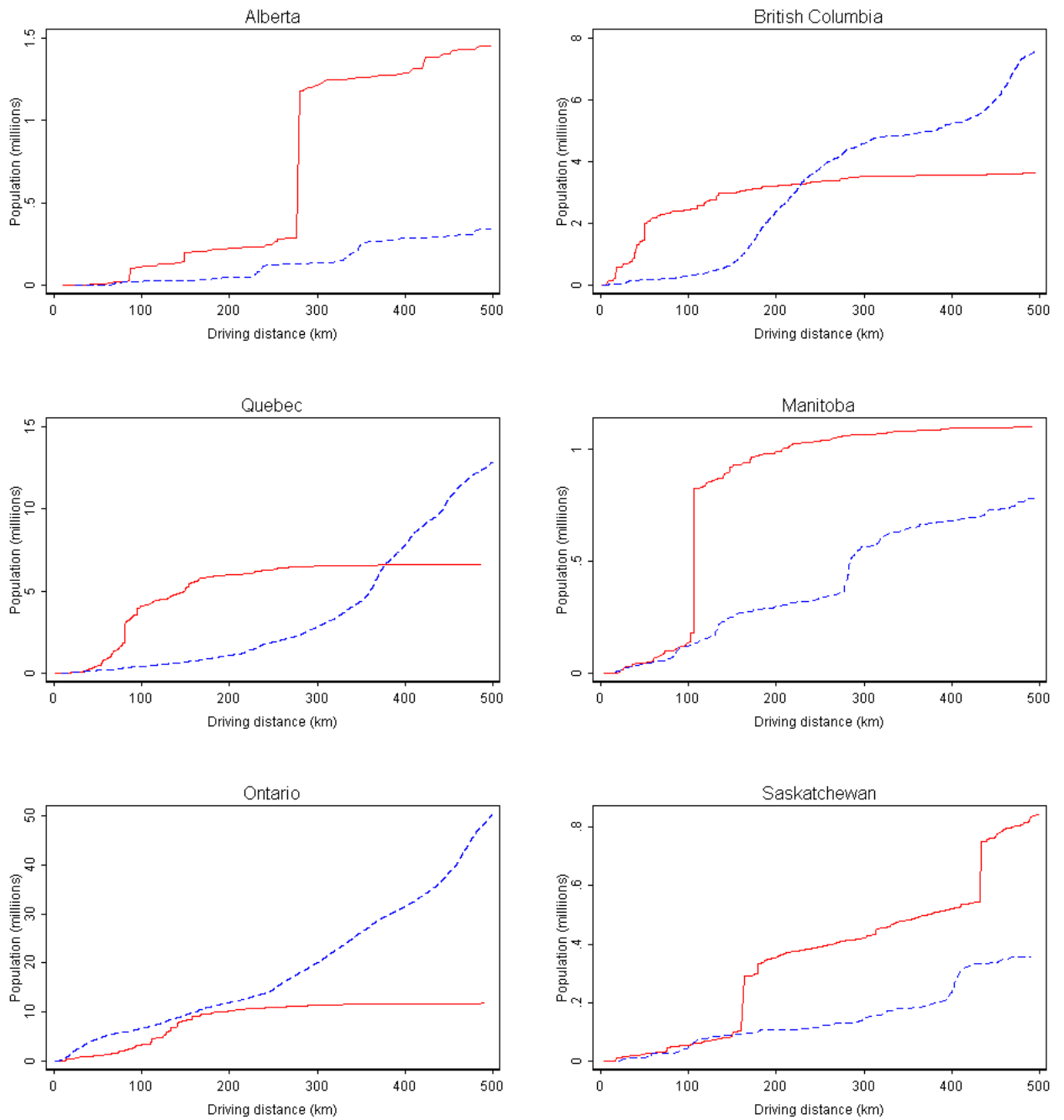

Figure B.1: Census Divisions in Southeastern Ontario

Figure B.2: Accumulated Population and Distance to the Border