Cross-Border Prices, Costs and Mark-ups[*]

Gita Gopinath
Harvard University

Pierre-Olivier Gourinchas
University of California at Berkeley

Chang-Tai Hsieh
University of Chicago

This Version: July 4, 2008
Very Preliminary: Do not quote/circulate

Abstract

We use data on retail prices and wholesale costs for detailed products at the barcode level from 325 stores of a large grocery chain to measure the effect of the US-Canadian border on market segmentation. Theoretically, we use a model of pricing and location on the circle to document possible patterns of cross-border prices. Empirically, we find clear evidence of international market segmentation. Cross-border price gaps are significantly higher than within country price gaps. Using the precise geographic location of each store, we find that UPC level prices and wholesale costs are discontinuous at the border. Our findings indicate that most differences in cross border prices arise from differences in an apparently tradable component of costs and not from systematic mark-up differences.

1 Introduction

To what extent do national borders and national currencies impose costs that segment markets across countries? Some of the central questions in international economics, ranging from the transmission of shocks across borders, to the effectiveness of stabilization policies, to the impact of changes in nominal exchange rates, hinge on the answer to this question.

A large literature has attempted to examine the effect of the border by looking at deviations from the law of one price (LOP), generally finding evidence of of deviations from the LOP that are large, volatile, and remarkably correlated with the nominal exchange rate.[1] In particular, a seminal contribution by Engel and Rogers (1996) starkly highlights the extent to which markets are segmented across countries by showing that, after controlling for the distance between cities, the volatility of changes of price indices for disaggregated product categories between U.S. and Canadian cities are much larger than that observed across cities in the same country.

The literature has followed up on Engel and Rogers' influential paper in two directions: by studying prices of products at a much more disaggregated level, and by decomposing the border effect of prices of a small number of specific products into marginal costs and markups.[2] A related literature also looks at the pass-through from exchange rates to import or consumption prices.[3] In a recent and interesting contribution Broda and Weinstein (2007) estimates border effects using a large amount of barcode-level price data collected at the consumer level. Unlike Engel and Rogers (1996), that paper concludes that the degree of price segmentation is similar across and within borders.

However, a recent paper by Gorodnichenko and Tesar (2005) points out that estimates of the border effect in regressions like those used by Engel and Rogers (1996) and Broda and Weinstein (2007) are unidentified unless the degree of within-country price dispersion is the same across countries. More generally, since both within-country and cross-country relative prices are endogenously

---

[1]See Rogoff (1996) and Goldberg and Knetter (1997) for comprehensive reviews of this literature.

[2]Crucini and Shintani (2006) and Crucini, Telmer, and Zachariadis (2005) for instance examine the retail price of narrowly defined product categories such as "Washing Powder", across countries within the European Union. Others focused on specific goods, such as *The Economist* magazine (Ghosh and Wolf (1994)) or *Ikea*'s furniture products (Haskel and Wolf (2001); Hassink and Schettkat (2001)). Parsley and Wei (2007) decompose the price of a *Big Mac* across countries into variation in marginal costs and variation in markups. Goldberg and Verboven (2005) study the automobile car market in Europe. See Goldberg and Verboven (2001) for a survey.

[3]See Goldberg and Knetter (1997), as well as the more recent contributions of Campa and Goldberg (2005, 2006); Gopinath and Rigobon (2007) and Gopinath, Itskhoki, and Rigobon (2007).

influenced by the presence of the border, this prevents any inference from estimated border coefficients in the absence of a model that derives the distribution of prices as an endogenous outcome of a potential barrier at the border.

This paper takes up the challenge laid down by Gorodnichenko and Tesar (2005). We derive a model along the lines of Salop (1979)'s circular city model that endogenously determines the distribution of prices within and across countries, in the presence of a border cost and heterogeneity in marginal costs. The model makes two predictions. First, the impact of border costs is observed only through changes in prices "close" to the border, and has little effect on the pricing decision "far" away from the border. Second, it predicts that when border costs are large enough, markets are effectively segmented across countries. In that case, it becomes impossible to estimate the size of border costs since border costs essentially become irrelevant to a firm's pricing decision. Instead, the analysis of price gaps across borders can recover a lower bound on border costs, and this lower bound could well be zero.

We then use a new dataset to estimate the cost of the U.S.-Canadian border within the framework of our model. The dataset has detailed product-level information at the bar-code level from 250 US stores (in 19 states) and 75 Canadian stores (in 5 provinces) from a large food and drug retailer, for every week between January 2004 and June 2007.[4] There are three important features of this data.

First, we have data on the price of detailed identical products sold by the same retailer in both countries. While the level of detail is similar to that used by Broda and Weinstein (2007), one key difference is that our data captures prices charged by the *same* retailer in all locations, while the Broda and Weinstein (2007) data contains prices at which consumers purchase a particular good without any control for retailer heterogeneity. Our findings will then be complementary to theirs, since they will help address the question of whether the large variation in retail prices within a country results from heterogeneity across retailers in the price they set for the exact same good, or whether a particular retailer's prices vary across locations. This distinction is important because models of deviations from the LOP due to pricing to market refer to pricing decisions by a single firm.

---

[4]The dataset contains 98,517 UPCs in the US, 33,179 UPCs in Canada, and 6,347 uniquely matched UPCs in both countries. The total number of price observations across stores and time is close to 40 million.

Second, we have information on the wholesale cost and the gross profit of every product for every store and for every week. This allows us to decompose the variation in retail prices into their markups and marginal costs components without the need to estimate demand and supply systems.

Third, we know the precise geographic location of the stores in our data. This information is crucial because a central prediction of our model is that the impact of the border on prices depends upon the distance to the border. We use it to estimate the deviation from the LOP between stores that are located right across the border from each other.[5]

In short, our data set allows us to answer the following questions: Do we see deviations from the LOP between stores that are located right across the border? If so, do these deviations from the LOP reflect differences in in relative marginal costs or in relative markups? The traditional interpretation has been that deviations from the LOP reflect either variations in the 'transportation, distribution and retail' component of marginal costs, or variations in markups across locations. Our data allows us to answer precisely this question.

We report three key findings: First, we find large and heterogenous deviations from the LOP across stores located close to the border. Second, the median of these deviations moves almost one-to-one with changes in the US-Canada nominal exchange rate. Lastly, most of the time-series movements in cross-country relative prices arise from movements in relative costs, and not from changes in relative mark-ups. This last finding – that price discontinuities move one to one with cost discontinuities – is consistent with full segmentation of markets, in which case price gaps only provide a lower bound of the border cost.

Our work leaves several questions unanswered. First, the finding that wholesale costs appear to be differ significantly across the border suggests that wholesale markets are segmented across countries, even for wholesalers servicing the same retailer. An obvious question that remains to be answered are the sources of this segmentation of the wholesale market. Second, we are only able to conduct our analysis for a single retailer and it remains to be seen whether the effect of the border is similar at the producer level. Third, we only look at the discontinuity of prices across the border. To compare border costs to within-country arbitrage costs we plan to apply the same procedure to discontinuities within country borders.[6] Finally, we exploit movements in relative costs that are

---

[5]Holmes (1998) uses a similar approach to estimate the effect of right-to-work laws on employment across US states.

[6]We present evidence (in Section 3) that cross border price gaps exceed within country price gaps.

3

driven by variation in the nominal exchange rate. However, to estimate border costs as opposed to just the lower bound on these costs, one may need much larger variation in relative costs than what we observe in our data.

2 Data Source

We have access to weekly store-level data for a sample of 325 grocery stores from a retail chain that operates in the United States and Canada. This chain is one of the leading food and drug retailers in North America. It operates directly or through its subsidiaries nearly 1800 stores in a broad range of geographic locations and socio-economic neighborhoods (1400 stores in the US and 400 in Canada).[7]

The data set contains weekly total sales, quantities sold, wholesale unit cost as well as a measure of per-unit gross profit for 125,349 unique Universal Product Codes (UPC) in 52 distinct product groups, for 250 stores in 19 US states, and 75 stores in 5 Canadian provinces, between Januray 2004 and June 2007 (178 weeks).[8,9] The total number of observations across stores and time is close to 40 million. Most of these observations are concentrated in the processed and unprocessed food and beverages categories, housekeeping supplies, books and magazines, and personal care products. This level of disaggregation allows for a very precise identification of products. For instance, in our data, a 25 fl.oz Perrier Mineral Water with a Lemon Twist and a 25 fl.oz Perrier Mineral Water with a Lime Twist are two separate members of the Soft Beverages product group.

Figure 1 plots the location of the stores in our data. Most US stores are located in the Western and Eastern corridors, in the Chicago area, Colorado and Texas, while most Canadian stores are located along a relatively narrow horizontal band running close to the border with the US.

---

[7]The data sharing agreement between this retailer and the research community allowed this data to be used by, among others, Chetty, Looney, and Kroft (2007); Eichenbaum, Jaimovich, and Rebelo (2007); Gicheva, Hastings, and Villas-Boas (2007) and Einav, Leibtag, and Nevo (2007). Unlike these papers we focus on the cross-country dimension of the data.

[8]All UPCs fall within the following structural hierarchy: (a) Business Group (e.g. DRF, Dairy, Refrigerated, Frozen Foods); (b) Business Unit (Dairy and Refrigerated Foods); (c) Product Group or 2-digit SMIC (36 Refrigerated Dairy); (d) Category or 4 digit SMIC (3601 - Milk/Milk Substitutes); (e) Class or 6 digit SMIC (3601 01 - Mainstream Milk); (f) Subclass or 8 digit SMIC (3601 01 01 - Whole Milk); (g) Subsubclass or 10 digit SMIC (3601 01 01 05 - 1/2 Gallon Whole Milk).

[9]The US stores are located in Alaska, Arizona, California, Colorado, Delaware, Hawaii, Idaho, Illinois, Maryland, Montana, Nebraska, New Jersey, New Mexico, Oregon, Pennsylvania, South Dakota, Texas, Virginia and Washington, as well as the District of Columbia. The Canadian stores are located in Alberta, British Columbia, Manitoba, Ontario and Saskatchewan.

Of the 125,349 unique UPCs in our dataset, our first task is to find the set of 'matched UPCs', i.e. the set of products that appear in identical form in at least one Canadian and one US store, in at least one week. It represents the set of goods for which we can evaluate deviations from the law of one price. This dataset of matched UPCs contains 6,347 unique products, or about 5% of the original dataset. This decline in matched products across the border is an important effect emphasized in Broda and Weinstein (2007) that carries across to our dataset. It underlies the importance of working with unique products, as identified by their UPCs. When comparing price indices across countries at higher levels of aggregation, one is likely to suffer from a serious composition bias.[10]

Table 1 reports some information on the number of distinct products (among matched goods) per product group, per store-week and per store-pair-week in our data. The distribution of products across product groups is very skewed, with a median around 16 products and a mean (unreported) of 117. The average store in our data carries 682 distinct matched products for which we have data in a typical week. We can match about 380 of these for within-country store-pairs in a given week, and 242 for cross-border store-pairs.

The data set contains information on "Gross" and "Net" price. Both prices exclude (US) sales and (Canadian) value-added taxes. The net price can differ from the gross price when there are sales (coupons, promotions). It is always smaller than or equal to the gross price and exhibits significantly more variability. We also have information on the "Whole-Sale Cost" which refers to the list price at which the retailer purchases a given product from the wholesaler.[11] These costs need not represent the true cost to the retailer given that there are typically freight and transport costs as well as retail allowances, i.e. rebates provided by the wholesaler to the retailer or vice versa. To correct for this, we can use data on "Adjusted Gross Profits" per unit to back out the "Net Cost", or imputed cost of goods. The precise link between the whole-sale cost and the net

---

[10]One possibility is that similar products do not have the same UPC in the US and Canada due to different labeling requirements (e.g. language, nutrition) in the two countries. Since an important aspect of our study is to identify *identical* goods in both countries, we treat such products as different goods.

[11]The whole-sale cost data is of superior quality to the data that is available from the Dominick's data base, since the latter is an average cost measure, while the data base we use has weekly cost information.

cost is as follows:[12]

$$\text{Net cost} \quad = \quad \text{Whole-sale cost} - \text{Allowances} \tag{1}$$
$$= \quad \text{Net Price} - \text{Adjusted Gross Profits}$$

At short horizons, with rent, capital and labor taken as given, it is natural to interpret the net cost as a measure of the marginal cost of goods.[13] Equivalently, "Adjusted Gross Profits" measure the mark-up at the product and store level. At longer horizons, adjusted gross profits represent an upper bound for the product mark-up.

3 Preliminary analysis of LOP deviations at the border

3.1 Median deviations over time

As a first pass at the data, the top left part of figure 2 reports the median average cross-border price gap over time. That is, for each week and each UPC, we compute the log-deviation between the average Canadian and US net prices. The figure reports the median of that distribution over time. When positive, this number indicates that more Canadian goods tend to be have a higher price than the corresponding US good. The figure indicates that the average price gap has increased over time from roughly -5% in June 2004 to 15% in June 2007. The figure also reports (the dashed line on the right-axis) the (log) US/CAN nominal exchange rate. The overall correlation between the two series is striking: the evolution over time in the median price gap mirrors almost perfectly the evolution of the nominal exchange rate.

Using the identity:

$$\ln p_t^i = \ln\left(\text{cost}_t^i\right) + \ln\left(\text{markup}_t^i\right),$$

the top-right and bottom left panels perform the same exercise for the imputed (net) cost and

---

[12]According to our retailer's definition, allowances consist of the sum of shipping allowances, scan allowances, direct-store-delivery case billback allowances, header flat allowances, late flat allowances, new item allowances, *minus* the sum of buying allowances, freight allowances, overseas freight and distress and other allowances.

[13]This requires that freight, transport and retail allowances are measured correctly at the product level. If these items are instead established at a more aggregated level, this would affect our measure of marginal costs. E.g., one may imagine that a soft drink manufacturer negotiates global allowances on a broad range of drinks sold to the retailer; similarly, it may be difficult to assess the transport cost & freight of a bottle of milk. Unfortunately, we don't have a breakdown of allowances between their different components. We proceed under the assumption that the net cost corresponds to an economic measure of marginal costs.

the resulting markup. Looking at the two figures side by side reveals a striking fact: most of the movements in the median price gap result from corresponding movements in relative costs, while relative mark-ups show barely any response to the fluctuations in the exchange rate. This result is robust to the definition of the price (gross versus net) or of the costs (wholesale versus imputed). Figure 3 reports the same analysis using gross prices and wholesale costs.

Prices in our sample change very frequently. The median frequency across UPC's using net prices (gross prices) is 0.41 (0.22) implying a median duration of 2.4 (4.5) weeks.[14] Despite the high median frequency, there are a significant fraction of goods that do not change price during their entire time in the sample. To ensure that these goods do not drive the results we divide the sample into high and low frequency adjusters depending on whether their frequency is above or below the median. Figure 4 presents the results for the high and low frequency groups respectively.

Both figures reveal that the movement of the median price gap is similar for high and low frequency adjusters. In both cases, we find that the median price gap increases over time. The contribution of imputed costs is smaller for the frequency adjusters. Finally, the median markup gap movements are small relative to prices and costs.

Overall, the evidence indicates that the median price gap moves with the nominal exchange rate even for UPC's that adjust prices more frequently and that costs differences play an important role.

3.2 Dispersion Across UPC's

The previous figures do not tell us anything about the dispersion of price gaps across UPCs at a point in time. Figure 5 reports the distribution of the cross-border price gap across UPCs for the first week of 2004 and the first week of 2007. The figure shows that there is a large dispersion of price gaps across UPCs at any given point in time. Hence, while the median moves closely with the exchange rate, the price gap for any individual UPC is likely to be dominated by idiosyncratic factors. This feature is also documented in Crucini and Shintani (2007).

Figure 6 reports the same distribution for the cross-border average cost gap and markup gap.

---

[14]The frequency number was arrived at as follows: we estimated the frequency of price adjustment for each UPC-store combination; Then we estimated the average frequency across these store combinations for each UPC. We then estimated the median within each category and the median across these categories.

The figures indicate significant dispersion in relative costs across the border, and a much tighter distribution of markup differences across the border.

3.3 Dispersion Across Stores

Finally, Tables 2, 3 and 4 below report some raw statistics for the extent of price dispersion within and across US and Canadian stores. Table 2 reports statistics for the gross price and net price charged across stores in the US and Canada in the first week of 2005. USA-USA (resp. CAN-CAN) report prices for store-pairs located within the US (resp. Canada), while CAN-US examines prices for cross-border store pairs. With 250 US stores and 75 Canadian stores, there are 31125 US-US store-pairs, 2775 CAN-CAN ones and 18750 cross border pairs. Define p_{jt}^i as the gross US dollar price of product i in store j at time t. We construct the (log) price gap between two stores j and k for good i as $q_{jkt}^i = \ln p_{jt}^i / p_{kt}^i$. If the law of one price holds, $q_{jkt}^i = 0$.

The median number of common UPCs for store pairs is 373 (405) within the US (Canada) and 248 for cross-border pairs.[15] Columns (1)-(3) report the mean, median and standard deviation of price differences for store-pairs. Statistics of this distribution are reported in the rows. The median across stores of the median price gap is 0 both within US and within Canada pairs. This result corroborates the evidence in Broda and Weinstein (2007) and Crucini, Telmer, and Zachariadis (2005) that price differentials are centered around zero within countries. Cross-border store pairs however have a large median gap of 6.5 percentage points. Since the U.S. store is always treated as store of reference, this implies that Canadian prices were 6.5 percent higher than US prices in the first week of 2005.[16]

Similarly, the median of the absolute gross price gap is larger for cross-border stores (13%) as compared to either the within-US (6.5%) or the within-Canada (0%) pairs. This difference is even larger when using net prices. The fact that there is less price dispersion within-Canada as compared to within-US is consistent with the evidence in Gorodnichenko and Tesar (2005) and Engel and Rogers (1996), and unlike Broda and Weinstein (2007). Similarly, the median absolute gross whole-sale price gap is 7 times larger for cross-border store pairs as compared to within U.S. store pairs. This difference is small for mark-ups.

---

[15]The median number of UPCs differs from Table 1 because we are only looking at a single week of data.

[16]Since these are pre-tax prices, the 7% Canadian value-added tax (or GST) cannot account for the result.

In Section A in the appendix we present results from estimating regressions along the lines of Engel and Rogers (1996) and find that the border coefficient is sizeable and robust to various specifications including excluding within Canada store pairs. However, in the rest of the paper, we depart from this regression framework and analyze the effect of the border using a regression discontinuity approach.

4 Circular World

To motivate the regression discontinuity design we employ in the next section, we consider a simple model along the line of Salop (1979) circular city model of horizontal differentiation. We introduce a border that splits the circle of unit length into two countries: country A and country B. Points on the circle are indicated by ω. This circular world is represented in figure 7.

4.1 Stores

There are N_A stores in region A and N_B stores in region B. The total number of stores is $N_{AB} \equiv N_A + N_B$. We fix the location of stores exogenously. This will be consistent with the empirical analysis which will focus on short horizons during which the location of firms will be taken as predetermined. The borders will be at $\omega = 0$ and $\omega = \frac{N_A}{N_{AB}}$. We refer to stores by their location: we parameterize store locations on the circle by the variable $\omega_i = (2i - 1)/2N_{AB}$ where $i = 1, 2, ...N_A, N_A + 1, N_A + 2, ...N_{AB}$. The stores closest to the border are $i = 1, N_A, N_A + 1, N_{AB}$. Stores are assumed to be equidistant from each other and located at

$$\omega_i = \frac{1}{2N_{AB}}, \frac{3}{2N_{AB}}, ..., \frac{2N_A - 1}{2N_{AB}}$$

for country A, and at

$$\omega_i = \frac{2N_A + 1}{2N_{AB}}, \frac{2N_A + 3}{2N_{AB}}, ..., \frac{2N_{AB} - 1}{2N_{AB}}$$

for country B.

If $N_A = N_B$ the two countries will be evenly sized in terms of consumers and firms. We allow for $N_A \neq N_B$. Stores in country i produce with a constant marginal cost c_i that may be different across countries.

4.2 Consumers

Consumers are uniformly distributed on the unit circle. All stores sell the same homogenous good and each consumer buys one unit of the good. Consumers have preferences over stores because, all else equal, they strictly prefer to shop in stores that are closer to their location. Consumers incur a cost t per unit distance. This cost could be transportation costs or could reflect the consumer's value of time. In addition there is a cost b incurred if the consumer crosses the border. The utility of a consumer located at ω is given by

$$u(\omega) = \nu - \theta p - t|\omega_i - \omega| + bI(\omega_i, \omega)$$

where $I(\omega_i, \omega)$ is an indicator function that takes the value 1 whenever the consumer and store are in different countries and zero otherwise. θ captures the own price elasticity of demand. t can be viewed as capturing the degree of substitutability across locations. The higher is t lower is the degree of substitution. We assume that ν is large enough so that all consumers purchase one unit of the good.

4.3 Demand Functions

We will assume that the parameters of the model are such that all firms earn positive profits in equilibrium. This implies that consumers will not shop at stores that are further than $1/N_{AB}$ from their own location. In particular, between stores i and $i - 1$, there will be a consumer who is indifferent between shopping at either store.

4.3.1 Interior Stores

Let us first consider the demand faced by an interior store, that is, $i \neq \{1, N_A, N_A + 1, N_{AB}\}$. The location $\underline{\omega}$ of the consumer who is indifferent between stores i and $i - 1$ satisfies

$$p_{i-1} + t(\underline{\omega} - \omega_{i-1}) = p_i + t(\omega_i - \underline{\omega})$$

Similarly the location $\bar{\omega}$ of the consumer indifferent between stores i and $i+1$ is given by

$$p_i + t(\bar{\omega} - \omega_i) = p_{i+1} + t(\omega_{i+1} - \bar{\omega})$$

As a result, the total demand for products at store i can be expressed as

$$D_i(p_{i-1}, p_i, p_{i+1}) = \bar{\omega} - \underline{\omega} = \frac{1}{N_{AB}} + \frac{p_{i+1} - 2p_i + p_{i-1}}{2t}.$$

The interior store in country j chooses p_i to maximize static profits,

$$\pi_i = (p_i - c_j)D_i(p_{i-1}, p_i, p_{i+1})$$

Due to the symmetry of the problem the solution takes the following simple form for stores in the interior of country A :

$$p_i = (\hat{p}_A - c_A - \frac{t}{N_{AB}})\frac{\cosh \kappa(i - \frac{N_A+1}{2})}{\cosh \kappa(\frac{N_A-1}{2})} + c_A + \frac{t}{N_{AB}}, \tag{2}$$

where $\kappa \equiv \cosh^{-1} 2 \approx 1.317$, $\hat{p}_A = p_1 = p_{N_A}$ represents the price at the store closest to the border in country A and cosh is the hyperbolic cosine function.[17]

By analogy, the interior solution for country B is

$$p_i = (\hat{p}_B - c_B - \frac{t}{N_{AB}})\frac{\cosh \kappa(i - \frac{N_B+1}{2} - N_A)}{\cosh \kappa(\frac{N_B-1}{2})} + c_B + \frac{t}{N_{AB}}. \tag{3}$$

where $\hat{p}_B = p_{N_{AB}} = p_{N_A+1}$ represents the price at the store closest to the border in country B.

As equations 2 and 3 indicate, all else equal, prices are positively related to marginal costs c_i and to the inverse of the elasticity of substitution across locations t. They are also negatively related to the total number of competitors, N_{AB}. Interior prices are affected by the border cost b indirectly through its effect on border prices $\hat{p}_A$ and $\hat{p}_B$. The strength of this effect is decreasing in the distance from the border, as one would expect.

---

[17]The hyperbolic cosine function is given by $cosh(x) = \frac{e^x + e^{-x}}{2}$.

4.3.2 Border Stores

For stores adjacent to the border, namely $i \in \{1, N_A, N_A + 1, N_{AB}\}$, there will be several distinct possibilities, depending on parameter values, that we specify below. In what follows, we consider the case where the number of stores N_i is large enough, in the sense that $\cosh \kappa(\frac{N_i - 3}{2}) / \cosh \kappa(\frac{N_i - 1}{2}) \approx \exp(-\kappa)$.[18]

Proposition 1 *[Full Segmentation] When the following condition is satisfied:*

$$|c_A - c_B| < b$$

the marginal consumer is at the border (i.e. national markets are fully segmented). Then:

(i) the prices of stores at the border are given by

$$\hat{p}_A = c_A + \frac{t}{N_{AB}} \frac{3e^{\kappa} - 1}{2e^{\kappa} - 1} \; ; \; \hat{p}_B = c_B + \frac{t}{N_{AB}} \frac{3e^{\kappa} - 1}{2e^{\kappa} - 1} \tag{4}$$

(ii) The difference in prices of border stores moves one to one with the difference in costs, i.e. $\partial(\hat{p}_A - \hat{p}_B)/\partial(c_A - c_B) = 1$.

Proposition 1 corresponds to the case where the difference in prices between border stores, $|\hat{p}_A - \hat{p}_B|$, is smaller than the border cost b. In this case the demand functions are independent of prices on the other side of the border, and markets are completed segmented. The observed difference in prices at the border is independent from the border cost b and only provides a lower bound on its true value. The solution to the interior prices are then determined by substituting the solution for the border prices into equations (2) and (3).

Proposition 2 *[Partial Segmentation]*

1. The marginal consumer for the border stores is located in country A if

$$c_A - c_B > b \tag{5}$$

In that case, the prices of stores at the border are given by

$$\hat{p}_A = \frac{(4 - e^{-\kappa})(j_A + b) + (j_B - b)}{(4 - e^{-\kappa})(4 - e^{-\kappa}) - 1} \; ; \; \hat{p}_B = \frac{(4 - e^{-\kappa})(j_B - b) + (j_A + b)}{(4 - e^{-\kappa})(4 - e^{-\kappa}) - 1} \tag{6}$$

---

[18]In practice, this requires $N_i > 10$. Note also that this restriction is not required in the symmetric case where $N_A = N_B$.

where

$$j_A = \left(3 - e^{-\kappa}\right)\left(c_A + \frac{t}{N_{AB}}\right) \ ; \ j_B = \left(3 - e^{-\kappa}\right)\left(c_B + \frac{t}{N_{AB}}\right)$$

2. *The marginal consumer for the border stores is located in country B if*

$$c_B - c_A > b$$

In that case, the prices of stores at the border are given by

$$\hat{p}_A = \frac{(4 - e^{-\kappa})(j_A - b) + (j_B + b)}{(4 - e^{-\kappa})(4 - e^{-\kappa}) - 1} \ ; \ \hat{p}_B = \frac{(4 - e^{-\kappa})(j_B + b) + (j_A - b)}{(4 - e^{-\kappa})(4 - e^{-\kappa}) - 1} \tag{7}$$

This is the case when $|\hat{p}_A - \hat{p}_B| > b$. In this case, the demand functions depend on prices on the other side of the border and the border parameter b enters the pricing equations. As long as condition (5) is satisfied the price difference increases in b. When b gets high enough, however, markets become segmented and we move to the full segmentation of Proposition 1, where the cross-border price difference is independent from b.

In Figure 8 we plot prices as a function of the distance to the border, where the border is represented by the strong vertical line at 0. Prices for country A (country B) are to the right (left) of the border. The border parameter b is assumed to be high enough that the marginal customer is at the border, that is prices are determined as in Proposition 1. For the left figure we assume that $c_A = c_B$ and $N_A = N_B$. That is, in the absence of the border, all firms would charge the same price. However, because of a border the stores close to the border are shielded from competition from stores across the border and they charge a higher price. However, there is no difference in border prices. As stated earlier, this does not imply that the border costs is 0, but just that it cannot be estimated from price differences across borders. For the figure on the right we allow costs to differ on each side of the border, with costs in region A being greater than costs in region B: $c_A > c_B$. This generates a discontinuity at the border. In the simulation, since $N_A = N_B$, the discontinuity is exactly equal to the difference in costs.[19] As we will see in the empirical section, this seems to be the relevant case.

In Figure 9 the border parameter b is set to 0. All else is the same as in the previous figure. As mentioned earlier, in the figure to the left all prices are identical regardless of how close to the border they are. In the figure to the right, there is a price discontinuity that arises from the

---

[19] This is also approximately true if N_A and N_B are sufficiently large.

differences in costs. The magnitude of this discontinuity is smaller than the difference in costs. When $N_A = N_B = 20$, $\hat{p}_A - \hat{p}_B = [(3 - e^{-\kappa})/(5 - e^{-\kappa})][(c_A - c_B) + 2b]$. The derivative of $(\hat{p}_A - \hat{p}_B)$ relative to $(c_A - c_B)$ is strictly less than 1.

As compared to Figure 8, where markets in country A are completely segmented from markets in region B, in Figure 9, the stores on either side of the border compete for customers. This explains why, in the case when $(c_A - c_B) > 0$, the border store in country A charges a lower price compared to the interior stores in country A, while the border store in country B charges a higher price than its interior stores.

In Figure 10 we assume that $c_A - c_B \equiv \Delta c$ follows an AR(1) process $\Delta c_t = \mu + 0.99 \Delta c_{t-1} + \varepsilon_t$, where $\sigma_\varepsilon = 0.03$. For a particular realization of shocks we plot the path of price differences at the border together with the path of the cost differences. In the figure on the left the two paths are identical as shown in Proposition 1, while in the figure on the right, which is the case when border costs equal zero and refers to Proposition 2, the effect on prices is dampened. The ratio of $(\hat{p}_A - \hat{p}_B)$ to Δc is a constant that is strictly less than 1.

To summarize, in this section we presented a simple model of horizontal differentiation where a homogenous good is sold by multiple stores and consumers have preferences over store location. Competition across stores leads to prices being interdependent, with the extent of influence of competitors prices on a stores prices declining in the distance from the store. Some of the simple intuitive insights that arise from this model that we use in our empirical analysis are as follows.

Firstly, if countries are completely symmetric, the endogenous distribution of prices within countries are identical and the border cost cannot be estimated by comparing price differences across borders. Consequently, regressions along the lines of Engel and Rogers (1996) and Broda and Weinstein (2007) reveal no information about the size of the border cost. An estimate of zero does not imply the absence of border costs. This point is distinct from the one made in Gorodnichenko and Tesar (2005) who emphasize the problems that arise with heterogeneity across countries.

Secondly, prices of stores that are far from the border are minimally affected by the size of the border cost. This was evident from equations 2 and 3. As seen in the right panel of Figures 8 and 9 prices of stores far from the border barely changes even when the border cost as a ratio of t declines from 20% to 0. The effect of the border is observed mostly for those stores close to the

14

border. In most of the existing literature, owing to lack of data, no distinction is made between stores that are close to the border versus those that are far from it. Since we have data on the precise geographical location of stores we can pay attention to prices that carry information about border costs. The next section presents results for price gaps as a function of distance from the border similar to Figures 8, 9 and 10.

Third, if border costs are sufficiently high, markets are perfectly segmented and the magnitude of border costs does not affect pricing decisions. Consequently, even if one had all the relevant data available for costs, level of competition, elasticities etc... the cost of the border could not be estimated. In that case, price differences at the border provide a lower bound on the true size of border costs and price differences should move one-to-one with cost differences.

Lastly, in general, prices depend on many factors such as the cost of traveling across distances t, the number of competitors N_i, the own price elasticity of demand θ, besides the cost of production in each location c_i and the border cost b. Parameters such as t and θ can typically vary with the demographic structure of households in different locations (income, age of households etc...). In our empirical section we will attempt to control for heterogeneity in factors that affect prices, so that the cross border variation in prices can be attributed to differences in costs interacting with the border.

5   Regression Discontinuity Design

This section examines the effect of the border using a regression discontinuity design.[20] Prices across stores can differ because of differences in market conditions or in costs. Our goal is to control for differences in market conditions for stores close to the border, examine the discontinuities in prices and compare it to costs.

To isolate the effect of the border from the effect of market conditions we will use one of the main advantages of our data, which is the precise geographic location of each store. Stores located far apart can face heterogenous market conditions. Consequently the border effect can be compounded

---

[20]See Imbens and Lemieux (2007) for a practical guide to the RD framework. A non-exhaustive list of recent studies applying RD designs includes Card, Chetty, and Weber (2006); Chay and Greenstone (2005); Chay, McEwan, and Urquiola (2005); DiNardo and Lee (2004) and Card, Mas, and Rothstein (2006). See also the February 2008 special issue of the Journal of Econometrics.

with other effects. We limit our attention to stores close to the border that face less heterogeneity in market conditions. Holmes (1998) uses a similar approach to estimate the effect of right-to-work laws on employment across US states.

Specifically, consider the following model of the relationship between the price p_j^i of product i in store j and various covariates:

$$\ln p_j^i = \alpha^i + \gamma^i C_j + \beta' X_j + \epsilon_j^i \tag{8}$$

where C_j is a dummy variable which is equal to 1 if the store is located in Canada and 0 if the store is located in the United States, X_j measures other characteristics of market j, and ϵ_j captures unobserved store characteristics. The parameter of interest is γ^i. The problem for inference is that the unobserved characteristics may not be independent from the location of the store. However, if the unobserved characteristics are a continuous function of the distance between the stores, we can control for these unobserved characteristics by introducing the distance between the stores as an additional regressor. In short, the equation we estimate is:

$$\ln p_j^i = \alpha^i + \gamma^i C_j + \beta' X_j + \theta^i D_j + \epsilon_j^{i\prime} \tag{9}$$

D_j is the distance (in kilometers) of store j to the border, where we define distance as positive for US stores and negative for Canadian stores. Following Imbens and Lemieux (2007), we estimate γ_i using a local linear regression approach. This local linear regression is estimated for US and Canadian stores within a certain distance to the border.[21] The optimal distance is selected using a standard bandwidth selection, based on the cross validation procedure advocated by Imbens and Lemieux (2007).[22] γ^i answers the question: how do prices change when one crosses from $D_j = \varepsilon$ to $D_j = -\varepsilon$, where ε is some small number.

The key assumption of the regression discontinuity approach is that the unobserved characteristics do not change discontinuously at the border. Although we can not test this assumption directly, we do three things to assess its plausibility. First, we examine whether the observable characteristics change abruptly at the border. The argument here is that if the observable charac-

---

[21] All store level observations beyond this cut-off are effectively discarded

[22] The procedure looks for the minimum value of the cross-validation criterion in 100km increments. The optimal bandwidth ranges from 100km to 700km. For most product-groups week pairs, the optimal bandwidth is either 100km, 350km or 500km.

teristics do not change discontinuously at the border, then this is also likely to be the case for the unobservable characteristics. In the same spirit, we compare the estimates of γ^i with controls for observable characteristics, and without controls. Third, we provide estimates of the border effect over time, exploiting the 20% nominal devaluation of the U.S. dollar from 2004 through 2007. Even if market features are different for the U.S. and Canada cities that are very close to the border, these differences are likely to be fairly stable over time.

We begin by plotting the distribution of the distance of the stores in our sample from the U.S.-Canadian border (in kilometers). Figure 11 plots the density of all the stores in our sample as a function of distance from the border, where distance is defined as negative for Canadian stores and positive for the US stores.

As can be seen, all Canadian stores are less than a 1000 kms from the border, while many stores in the U.S. are more than 1000 kms from the border. Obviously, the geographical concentration of economic activity in the U.S. is very different from that in Canada, highlighting Gorodnichenko and Tesar (2005)'s caution about estimates that do not take cross-country differences in within-country heterogeneity into account. Although this is less of a concern with our approach, since we are only looking at U.S. and Canadian stores that are physically close to each other, we will also work with a sample of stores located in Oregon, Washington, and British Columbia (21 Canadian and 41 U.S. stores) where market conditions are likely to be more homogenous. Figure 12 plots the distribution of distance to the border for this subsample of west coast stores.

In figure 13 we depict the regression discontinuity for some relevant covariates. As suggested by our circular world model, the covariates that we should consider are ones that potentially affect demand characteristics. We look at the number of Supermarkets in each local market,[23] the population density, the proportion of people aged 0-19 and aged 65 and up, the proportion of black people, the year the store was opened and household income.[24]

Graphically, for several of these variables no stark discontinuity is apparent. We formally test for this and find that when all stores are included there is some discontinuity at the border for the

---

[23]These are establishments in NAICS 445110 (Supermarkets and Other Grocery Stores, but not Convenience Stores)

[24]We obtained the data for the U.S. from the US population census and economic census data base and from Statistics Canada for Canada. There is a difference in the level of disaggregation at which the data is collected because Canadian data is collected at the county level while U.S. data is collected by zip code.

age variables as well as the proportion of African Americans. When we restrict attention to our west coast subsample of stores, these discontinuities disappear, but we find some discontinuities for the fraction of seniors as well as median household income. We include controls for these variables when exploring the price gaps.

5.1 Regression Discontinuity Estimates

We begin our investigation of the price gaps across the border by plotting the average gross price for three detailed products against the distance of the store from the border.[25] The figures illustrate that the effect of the border may vary substantially across products. For the product "Perrier Water Regular" and for "Kleenex Lotion Upright", plotted in Figures 15 and 16 there is a clear discontinuity at the border and importantly this discontinuity is not evident within country borders. On the other hand, for "Cascade Regular Powder" there is no apparent discontinuity at the border (Figure 17).

The left panel of figure 18 plots the kernel density of the border effect for roughly 1,200 products in the first week of 2004 and 2007, with controls only for distance. This figure confirms the visual impression from figures 15-17 of considerable heterogeneity in the border effects. Some effects are negative while others are positive. The median absolute discontinuity is 14%. In addition, we see clearly that the distribution of the border effect shifts to the right after 2004. In words, the price in Canadian stores close to the border exceed that in U.S. stores close to the border for more products in 2007 relative to 2004. This figure is to be compared to figure 5. The regression discontinuity confirms our preliminary results on median cross border price gaps.

The left panel of figure 19 provides additional information on the relationship between the exchange rate and the border effect. The figure reports the median border effect across product groups in every week, against the US/Canadian dollar nominal exchange rate. As can be seen, there is virtually a one-to-one correspondence between movements in the median price border effect and the exchange rate. In January 2004, the median price gap was 5 percent lower in Canada relative to the US. By the middle of 2007, the median price gap was now 15 percent higher in Canada. Over this time period, the U.S. dollar depreciated by roughly 20 percent relative to the Canadian dollar.

---

[25]This data is for the first week of 2004.

The right panels of Figure 18 and 19 report the same information when we introduce the covariates in the border effect regression. The main results are unchanged. The distribution of border treatments in the right panel of figure 18 seems to be shifted to the right compared to the left panel. This indicates that after controlling for covariates, the border effect may in fact increase. Over time, figure 19 reports that the median border effect across borders co-moves significantly with the nominal exchange rate.

To investigate the role of costs and mark-ups in the movements in price gaps, Figure 20 plots the distribution of regression discontinuity estimates of the border effect on the wholesale costs and on markups over wholesale costs in 2004 and 2007 (from here on, the estimates do not control for covariates). There is some evidence that border effect on markups shifted to the left from 2004 through 2007, suggesting that the depreciation of the U.S. dollar lowered markups in Canadian stores relative to that in U.S. stores. However, the overwhelming movement is in the costs. Figure 21 presents the related median regression discontinuity estimates for the wholesale costs and mark-up over time. This figure confirms the impression that wholesale costs moved very closely with the exchange rate and the effect on markups while negative was of a much smaller magnitude. Since wholesale costs can be viewed as the most 'traded' component of the retailers costs, the discontinuity in this component of costs is particularly striking. The evidence from the regression discontinuity design about the importance of segmentation in wholesale costs is then consistent with the evidence reported in Section 3.

As a robustness check, we estimate discontinuities using only the west coast stores close to the border and the results are reported in Figure 22. The results are consistent with the findings that use all stores.

In Figure 23 we overlay the plots for movements in the cost gap and the price gap (along with the log of the exchange rate). The left figure corresponds to all stores and the right figure to west coast stores. As is evident the two series move closely together which is consistent with the two markets being fully segmented, given the differences in costs that be observe in our sample. This is similar to the left panel of Figure 10 in the theoretical section.

Taken together, the empirical evidence in this section delivers four main results. First, there is a great deal of heterogeneity in the "effect" of the border on prices, with both negative and

positive price gaps. Second, the feature that price gaps move almost one to one with costs gaps suggests that the two markets are fully segmented. In that case, our model indicates that price gaps provide a lower bound on the border costs. Since we find significant price gaps, we conclude that the effect of the border is sizeable. Third, the fact that the results in this section, where we restrict attention to border stores, is so similar to the ones obtained when all stores were used is also evidence of the fact that the markets are fully segmented, for otherwise there should be significant differences between border stores price gaps and price gaps of interiors store. Fourth, it appears that whole-sale markets are highly segmented, even when servicing the same store. This is especially striking since the wholesale component is the most tradable component of overall costs.

6   Discussion and Conclusion

The paper delivers three important results. First, this paper takes up the challenge laid down by Gorodnichenko and Tesar (2005). We provide a simple model along the lines of Salop (1979) to interpret and measure the effect of the border. The model shows conditions under which one can estimate a lower bound on the effect of the border from discontinuities in prices close to the border. We implement this strategy empirically using a new dataset that contains bar-code level prices for 6,347 goods from a large food and drug retailer, for every week between January 2004 and June 2007. Using a regression discontinuity design, we find significant price discontinuities at the border. To the question: is there a significant border effect, we find that the answer is an overwhelming yes.

Moreover, we establish that the media price are strongly correlated with movements in the US-Canada nominal exchange rates. Finally, we also find that movements in cross-border relative prices reflect mostly movements in marginal costs, while markups remain more stable.

Our finding raise important questions for future research. For instance, we find it quite striking that the most tradable component of overall retail costs (the wholesale cost) appear so segmented even when servicing the same retailer in different locations. In future work, we plan to investigate this question more fully. We also propose to compare within-country price discontinuities to cross-border price discontinuities more systematically.

References

BRODA, C., AND D. WEINSTEIN (2007): "Understanding International Price Differences Using Barcode Data," Discussion paper, Chicago GSB.

CAMPA, J., AND L. GOLDBERG (2005): "Exchange Rate Pass-Through into Import Prices," *Review of Economics and Statistics*, 87(4), 679–690.

———— (2006): "Distribution Margins, Imported Inputs, and the Sensitivity of the CPI to Exchange Rates," Federal Reserve Bank of New York.

CARD, D., R. CHETTY, AND A. WEBER (2006): "Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," Working Paper 12639, National Bureau of Economic Research.

CARD, D., A. MAS, AND J. ROTHSTEIN (2006): "Tipping and the Dynamics of Segregation in Neighborhoods and Schools," Working Papers 894, Princeton University, Department of Economics, Industrial Relations Section.

CHAY, K. Y., AND M. GREENSTONE (2005): "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, 113(2).

CHAY, K. Y., P. J. MCEWAN, AND M. URQUIOLA (2005): "The Central Role of Noise in Evaluating Interventions that Use Test Scores to Rank Schools," *American Economic Review*, 95(4), 1237–1258.

CHETTY, R., A. LOONEY, AND K. KROFT (2007): "Salience and Taxation: Theory and Evidence," Working Paper 13330, National Bureau of Economic Research.

CRUCINI, M. J., AND M. SHINTANI (2006): "Persistence in Law-Of-One-Price Deviations: Evidence from Micro-Data," Working Paper 06-W17, Department of Economics, Vanderbilt University.

———— (2007): "Microeconomic Sources of Real Exchange Rate Variability," Working paper, Vanderbilt University.

CRUCINI, M. J., C. I. TELMER, AND M. ZACHARIADIS (2005): "Understanding European Real Exchange Rates," *American Economic Review*, 95(3), 724–738.

DiNardo, J., and D. S. Lee (2004): "Economic Impacts of New Unionization On Private Sector Employers: 1984-2001," *The Quarterly Journal of Economics*, 119(4), 1383–1441.

Eichenbaum, M., N. Jaimovich, and S. Rebelo (2007): "Reference Prices and Nominal Rigidities," mimeo, Stanford.

Einav, L., E. Leibtag, and A. Nevo (2007): "Errors in Self-Reported Data: A Cross-Validation of Homescan Data," mimeo, Stanford.

Engel, C., and J. Rogers (1996): "How Wide is the Border?," *American Economic Review*, 86, 1112–1125.

Ghosh, A. R., and H. C. Wolf (1994): "Pricing in International Markets: Lessons From The Economist," Working Paper 4806, National Bureau of Economic Research, Inc.

Gicheva, D., J. Hastings, and S. Villas-Boas (2007): "Revisiting the Income Effect: Gasoline Prices and Grocery Purchases," Working Paper 13614, National Bureau of Economic Research.

Goldberg, P., and M. Knetter (1997): "Goods Prices and Exchange Rates: What Have We Learned?," *Journal of Economic Literature*, 35(3), 1243–1272.

Goldberg, P., and F. Verboven (2001): "The Evolution of Price Dispersion in the European Car Market," *Review of Economic Studies*, pp. 811–48.

———— (2005): "Market Integration and Convergence to the Law of One Price: Evidence from the European Car Market," *Journal of International Economics*, pp. 49–73.

Gopinath, G., O. Itskhoki, and R. Rigobon (2007): "Currency Choice and Exchange Rate Pass-through," Harvard University.

Gopinath, G., and R. Rigobon (2007): "Sticky Borders," *Quarterly Journal of Economics, forthcoming*.

Gorodnichenko, Y., and L. Tesar (2005): "A Re-Examination of the Border Effect," Working Paper 11706, National Bureau of Economic Research.

Haskel, J., and H. Wolf (2001): "The Law of One Price–A Case Study," *Scandinavian Journal of Economics*, 103(4), 545–58.

HASSINK, W. H., AND R. SCHETTKAT (2001): "On Price-Setting for Identical Products in Markets without Formal Trade Barriers," Discussion Paper 315, Institute for the Study of Labor (IZA).

HOLMES, T. (1998): "The Effect of State Policies on the location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, 106(4), 667–705.

IMBENS, G., AND T. LEMIEUX (2007): "Regression Discontinuity Designs: A Guide to Practice," Working Paper 13039, National Bureau of Economic Research, Inc.

PARSLEY, D., AND S. WEI (2007): "A Prism into the PPP Puzzles: The Micro-Foundations of Big Mac Real Exchange Rates," *Economic Journal*, 117(523), 1336–1356.

ROGOFF, K. (1996): "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, 34, 647–668.

SALOP, S. (1979): "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 10, 141–156.

7 Tables and Figures

	Number of Unique Products								
	per product group			per store-week			per store-pair-week		
	median	10%	90%	mean	10%	90%	mean	10%	90%
US	15.5	2	268	692	509	889	378	263	504
Canada	16	2	262	649	417	825	384	234	530
Both countries	16.5	2	269	682	487	874			
Cross-border pairs							242	148	328

Table 1: Descriptive Statistics. The table reports the median number of unique products per product group, per store per week, and per store-pair per week. The table also reports the first and last deciles.

	Mean	Median	St. Dev.	Mean Absolute	Med. Absolute
	(1)	(2)	(3)	(4)	(5)
			Panel A: Gross Prices		
			USA-USA store-pairs (31125)		
Median	0.012	0.000	0.133	0.083	0.065
Average	0.018	0.008	0.131	0.084	0.055
St. Dev.	0.041	0.028	0.001	0.001	0.001
			CAN-CAN store-pairs (2775)		
Median	0.007	0.000	0.090	0.033	0.000
Average	0.009	0.000	0.094	0.037	0.004
St. Dev.	0.001	0.001	0.030	0.021	0.010
			CAN-USA store-pairs (18450)		
Median	0.068	0.065	0.220	0.166	0.130
Average	0.067	0.064	0.223	0.170	0.131
St. Dev.	0.048	0.044	0.028	0.027	0.027
			Panel B: Net Prices		
			USA-USA store-pairs (31125)		
Median	0.004	0.000	0.155	0.091	0.049
Average	0.010	0.006	0.152	0.091	0.050
St. Dev.	0.038	0.027	0.034	0.030	0.031
			CAN-CAN store-pairs (2775)		
Median	0.009	0.000	0.079	0.034	0.000
Average	0.012	0.000	0.084	0.035	0.002
St. Dev.	0.022	0.001	0.028	0.018	0.007
			CAN-USA store-pairs (18450)		
Median	0.367	0.311	0.3318	0.387	0.314
Average	0.365	0.309	0.318	0.388	0.315
St. Dev.	0.047	0.046	0.034	0.042	0.044

Table 2: Deviations from the Law of One Price for Retail Prices: Panel A refers to Gross prices and Panel B refers to net prices. The table reports within and between-country statistics (the rows) for the mean, median, standard deviation, mean absolute and median absolute (log) price gap within store-pairs (the columns) for the first week of 2005.

	Mean	Median	St. Dev.	Mean Absolute	Med. Absolute
	(1)	(2)	(3)	(4)	(5)
		Panel A: Gross Wholesale Prices			
		USA-USA store-pairs (31125)			
Median	0.004	0.000	0.097	0.039	0.000
Average	0.015	0.000	0.189	0.057	0.002
St. Dev.	0.030	0.003	0.027	0.043	0.006
		CAN-CAN store-pairs (2775)			
Median	-0.000	0.000	0.085	0.018	0.000
Average	-0.001	0.000	0.085	0.019	0.000
St. Dev.	0.005	0.000	0.024	0.006	0.000
		CAN-USA store-pairs (18450)			
Median	0.167	0.120	0.329	0.236	0.142
Average	0.181	0.121	0.412	0.257	0.145
St. Dev.	0.048	0.029	0.182	0.057	0.028
		Panel B: Imputed Costs			
		USA-USA store-pairs (31125)			
Median	0.008	0.000	0.213	0.083	0.020
Average	0.021	0.002	0.276	0.099	0.027
St. Dev.	0.041	0.011	0.151	0.048	0.024
		CAN-CAN store-pairs (2775)			
Median	-0.003	0.000	0.135	0.047	0.002
Average	-0.004	0.000	0.138	0.047	0.003
St. Dev.	0.011	0.000	0.033	0.011	0.003
		CAN-USA store-pairs (18450)			
Median	0.427	0.338	0.368	0.447	0.343
Average	0.441	0.345	0.438	0.464	0.350
St. Dev.	0.058	0.048	0.169	0.058	0.048

Table 3: Deviations from the Law of One Price for Wholesale Prices: Panel A refers to Gross wholesale prices and Panel B refers to imputed costs. The table reports within and between-country statistics (the rows) for the mean, median, standard deviation, mean absolute and median absolute (log) wholesale price gap within store-pairs (the columns) for the first week of 2005.

	Mean	Median	St. Dev.	Mean Absolute	Med. Absolute
	(1)	(2)	(3)	(4)	(5)
USA-USA store-pairs (31125)					
Median	-0.006	0.000	0.246	0.133	0.072
Average	-0.010	-0.001	0.308	0.144	0.072
St. Dev.	0.048	0.020	0.141	0.050	0.031
CAN-CAN store-pairs (2775)					
Median	0.011	0.000	0.140	0.062	0.001
Average	0.015	0.003	0.141	0.062	0.019
St. Dev.	0.026	0.007	0.032	0.021	0.016
CAN-USA store-pairs (18450)					
Median	-0.067	-0.023	0.283	0.183	0.122
Average	-0.075	-0.034	0.344	0.202	0.124
St. Dev.	0.056	0.039	0.142	0.052	0.026

Table 4: Differences in Log Mark-ups: The table reports within and between-country statistics (the rows) for the mean, median, standard deviation, mean absolute and median absolute (log) mark-up gap within store-pairs (the columns) for the first week of 2005. The mark-up is the ratio of net price to imputed cost.
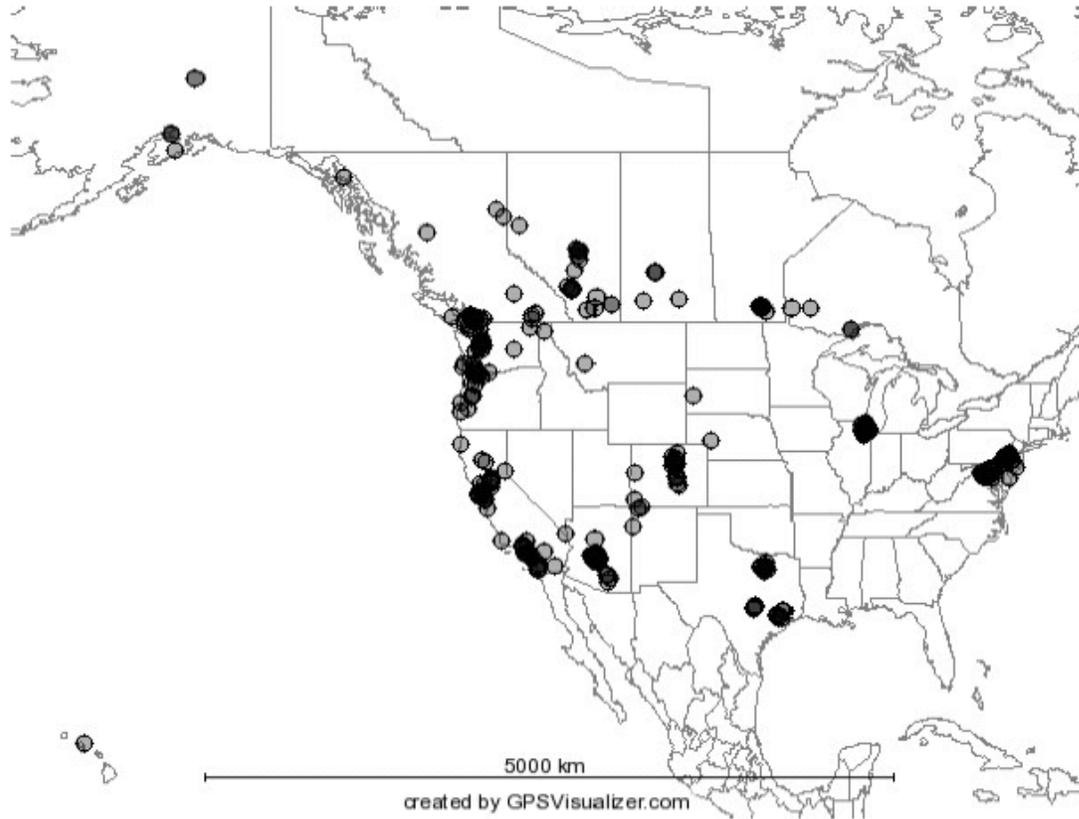
Figure 1: Map of the 325 retail North American stores in our data (250 U.S. and 75 Canada)
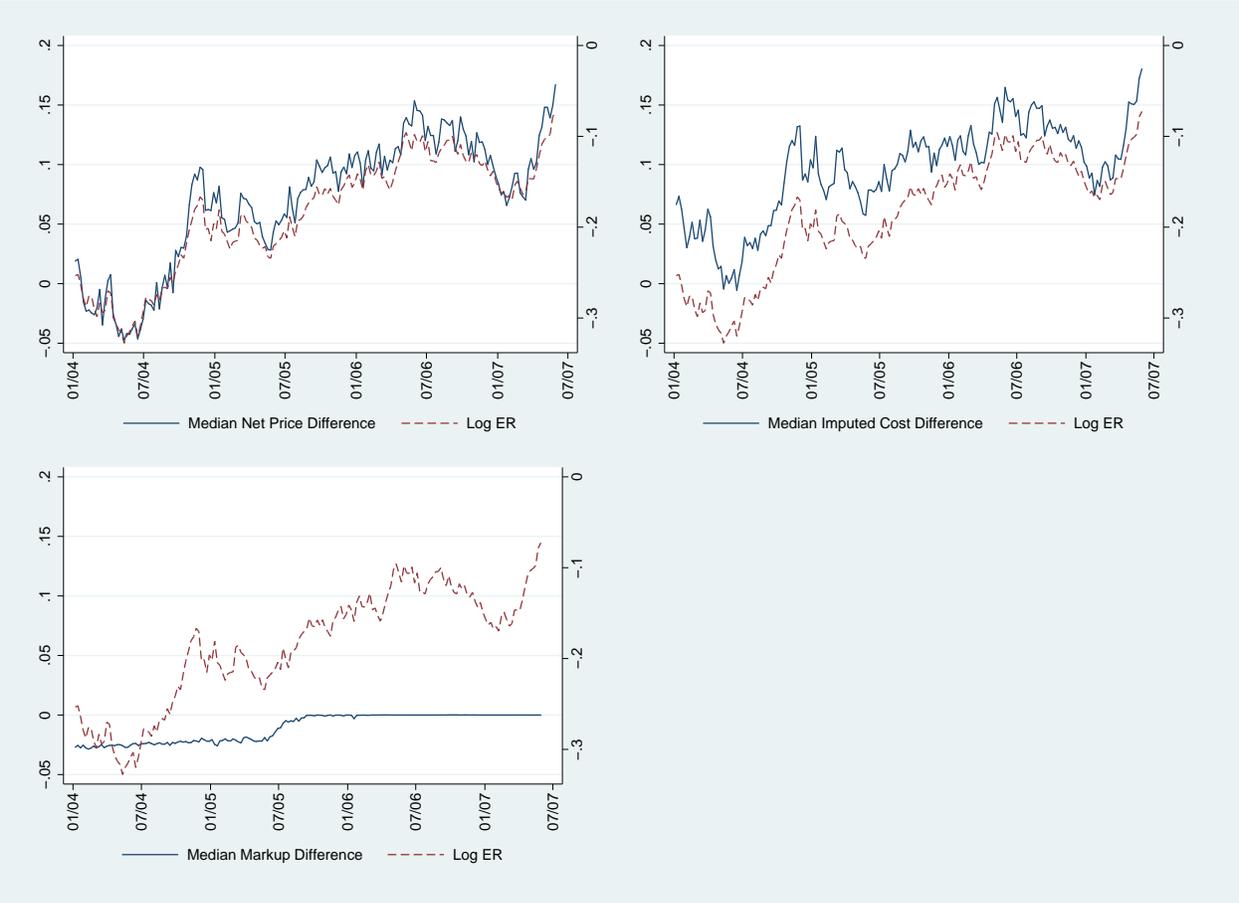
Figure 2: Median net price, imputed cost and markup cross-border gap and exchange rate
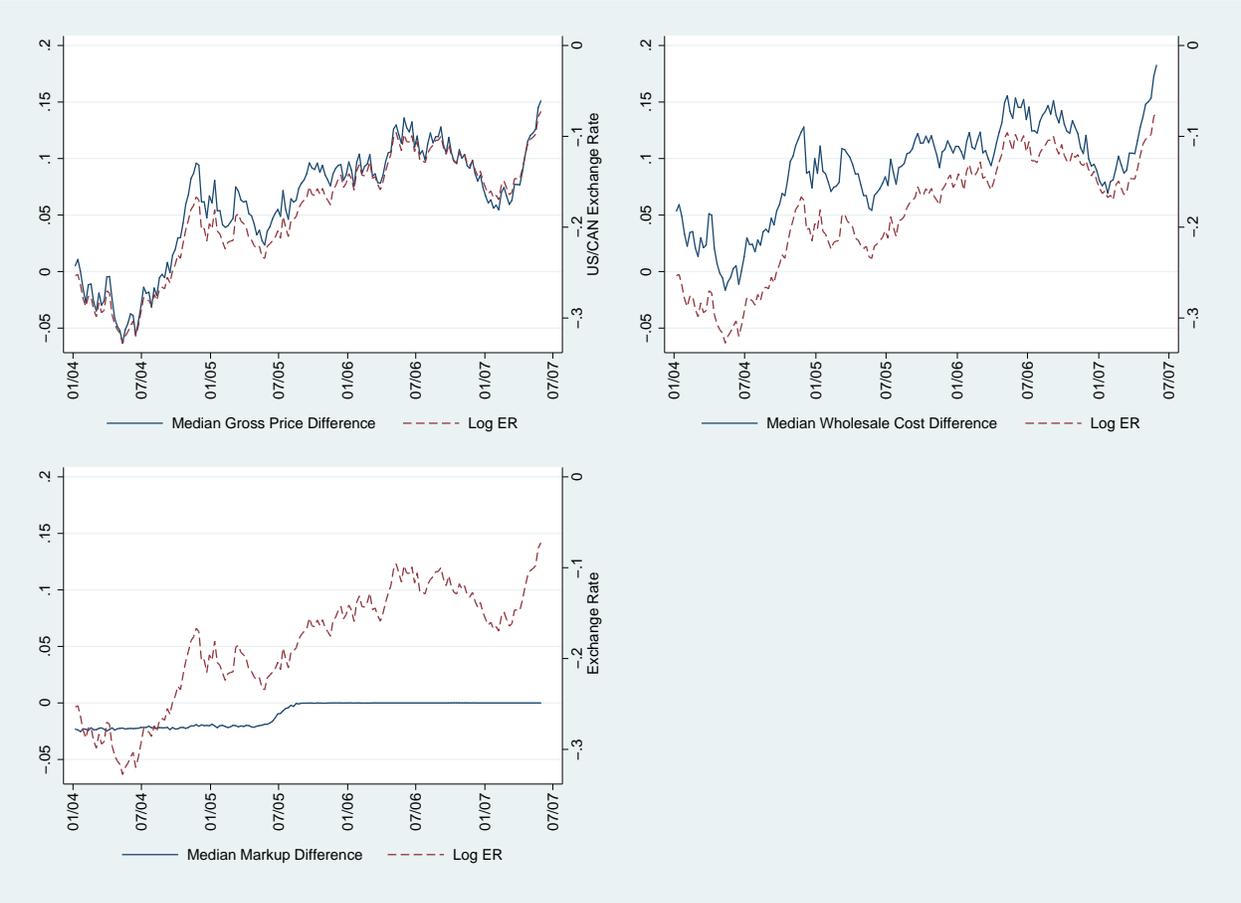
Figure 3: Median gross price, wholesale cost and markup cross-border gap and exchange rate

Figure 4: Median net price, imputed cost and markup cross-border gap and exchange rate for high (top) and low (bottom) frequency adjusters.

30
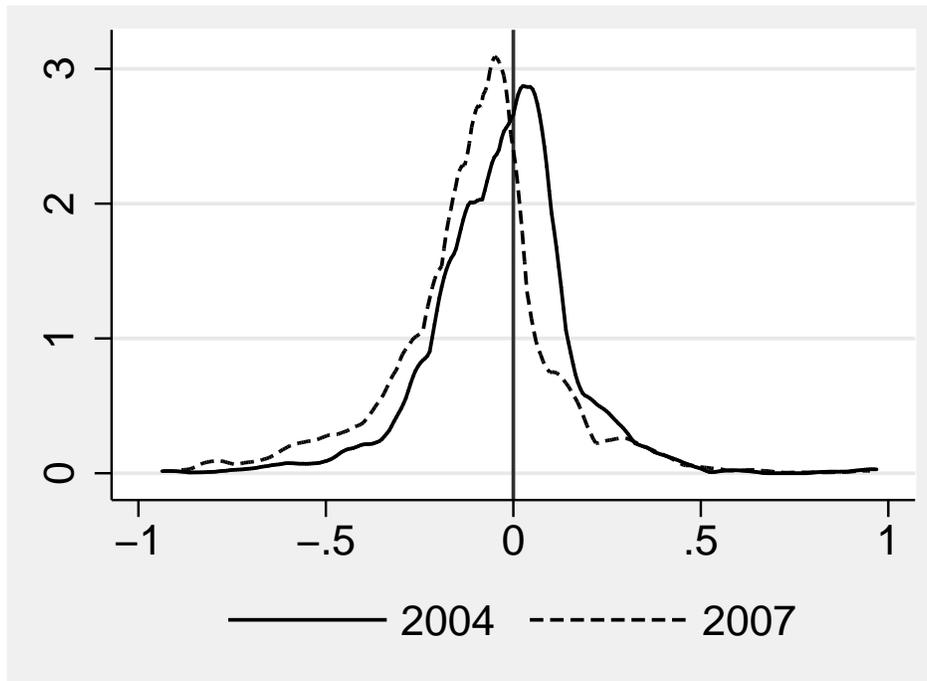
Figure 5: The dispersion of cross-border average UPC price gap in the first week of 2004 and 2007
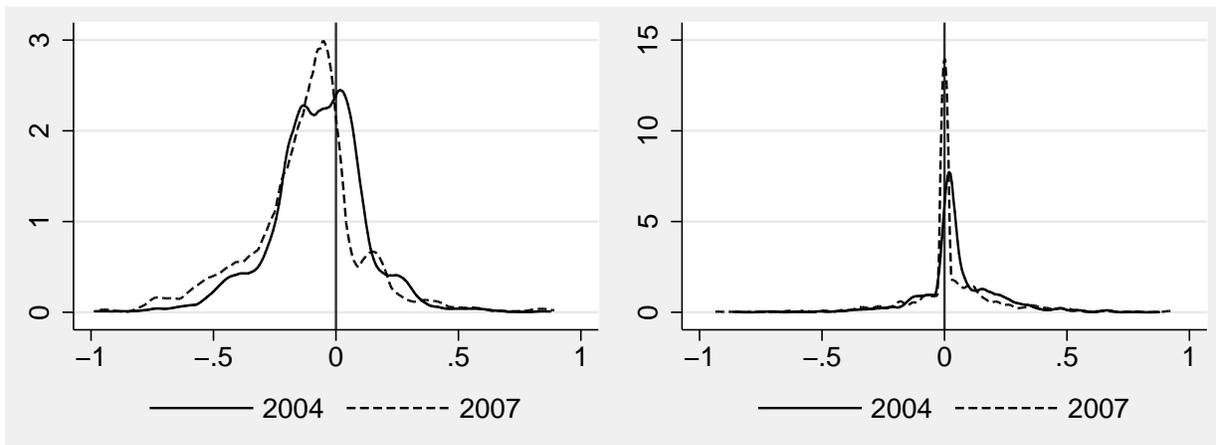
Figure 6: The distribution of cross-border average cost (left) and markup (right) gaps in the first week of 2004 and 2007

Figure 7: Circular World
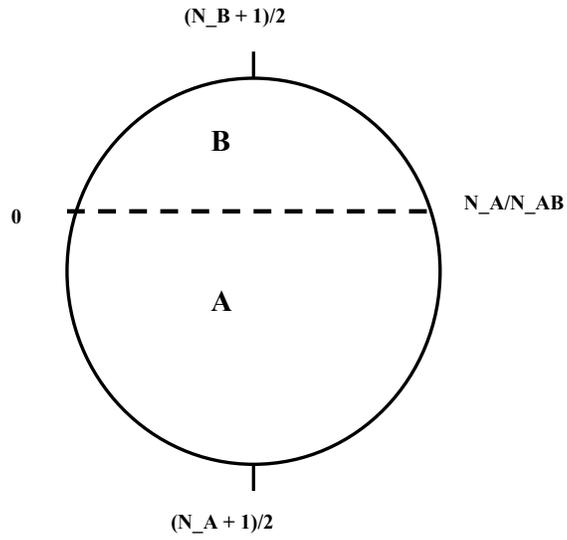
Figure 8: Price Discontinuity at the Border
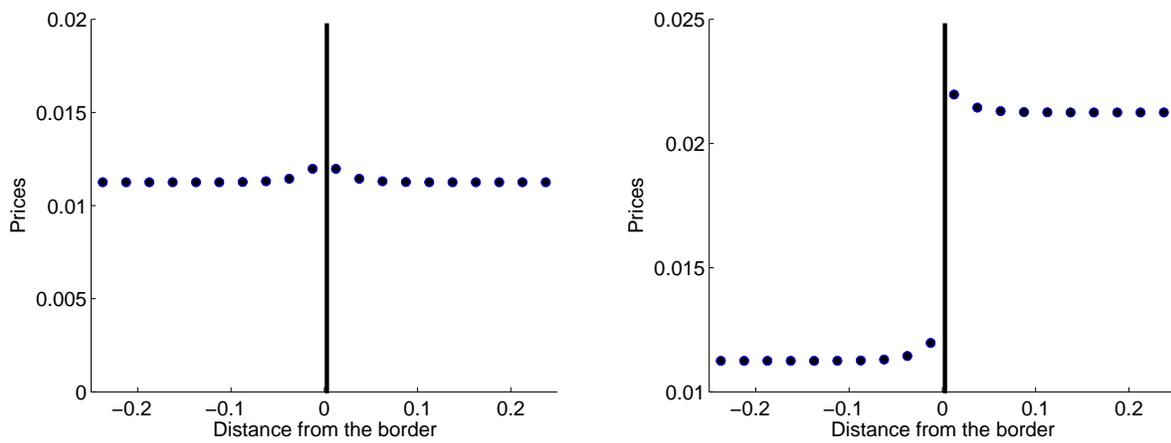
Note: For the left figure the parameters are $N_a = N_b = 20$, $t = 0.05$, $c_a = c_b = 0.01$ and for the figure on the right the parameters are $N_a = N_b = 20$, $t = 0.05$, $c_a = 0.02 > c_b = 0.01$

Figure 9: Price Discontinuity at the Border
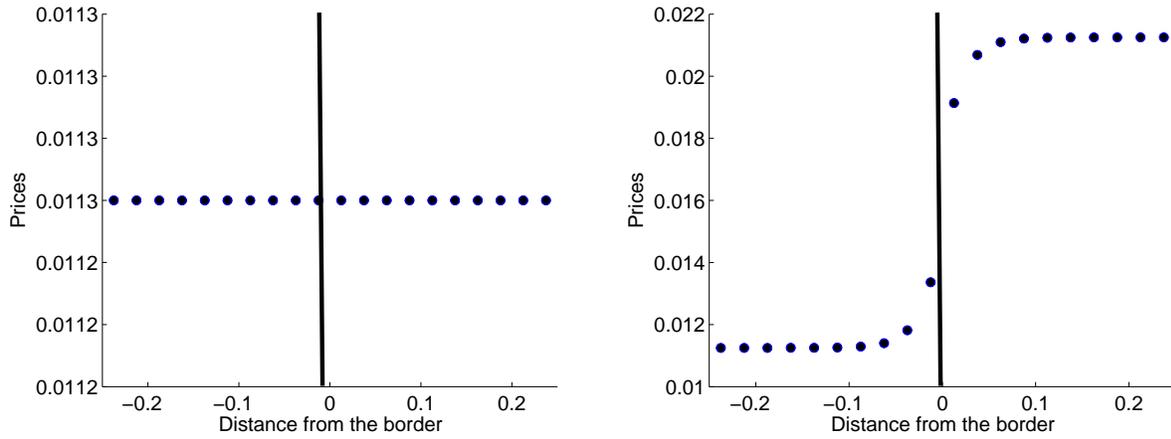
Note: For the left figure the parameters are $N_a = N_b = 20$, $t = 0.05$, $c_a = c_b = 0.01$, $b = 0$ and for the figure on the right the parameters are $N_a = N_b = 20$, $t = 0.05$, $c_a = 0.02 > c_b = 0.01$, $b = 0$.

Figure 10: Time Path of Price Discontinuity at the Border
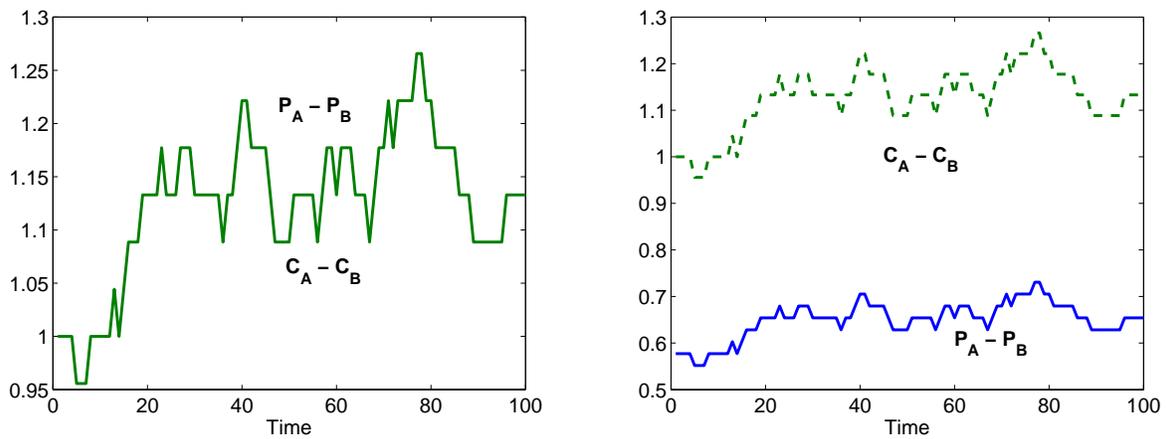
Note: The parameters are $N_a = N_b = 20$, $t = 0.05$. For the left figure $|p_{border\ A} - p_{border\ B}| < b$, which is the case of Proposition 1 and the right figure assumes $b = 0$, which is consistent with the case in Proposition 2.

33
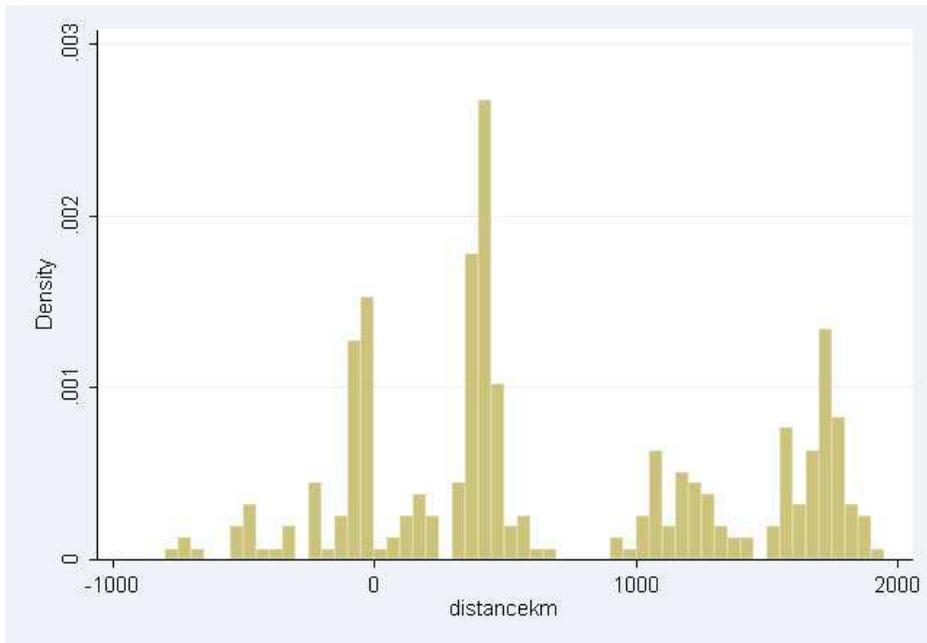
Figure 11: Distance to the Border (All Stores)
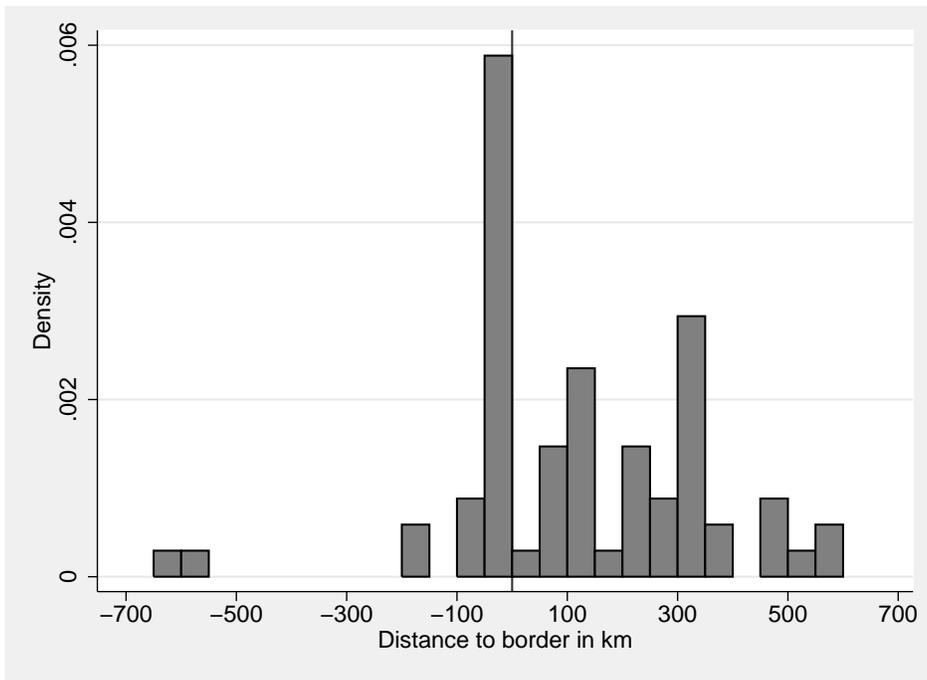
Figure 12: Distance to the Border (West Coast Stores)
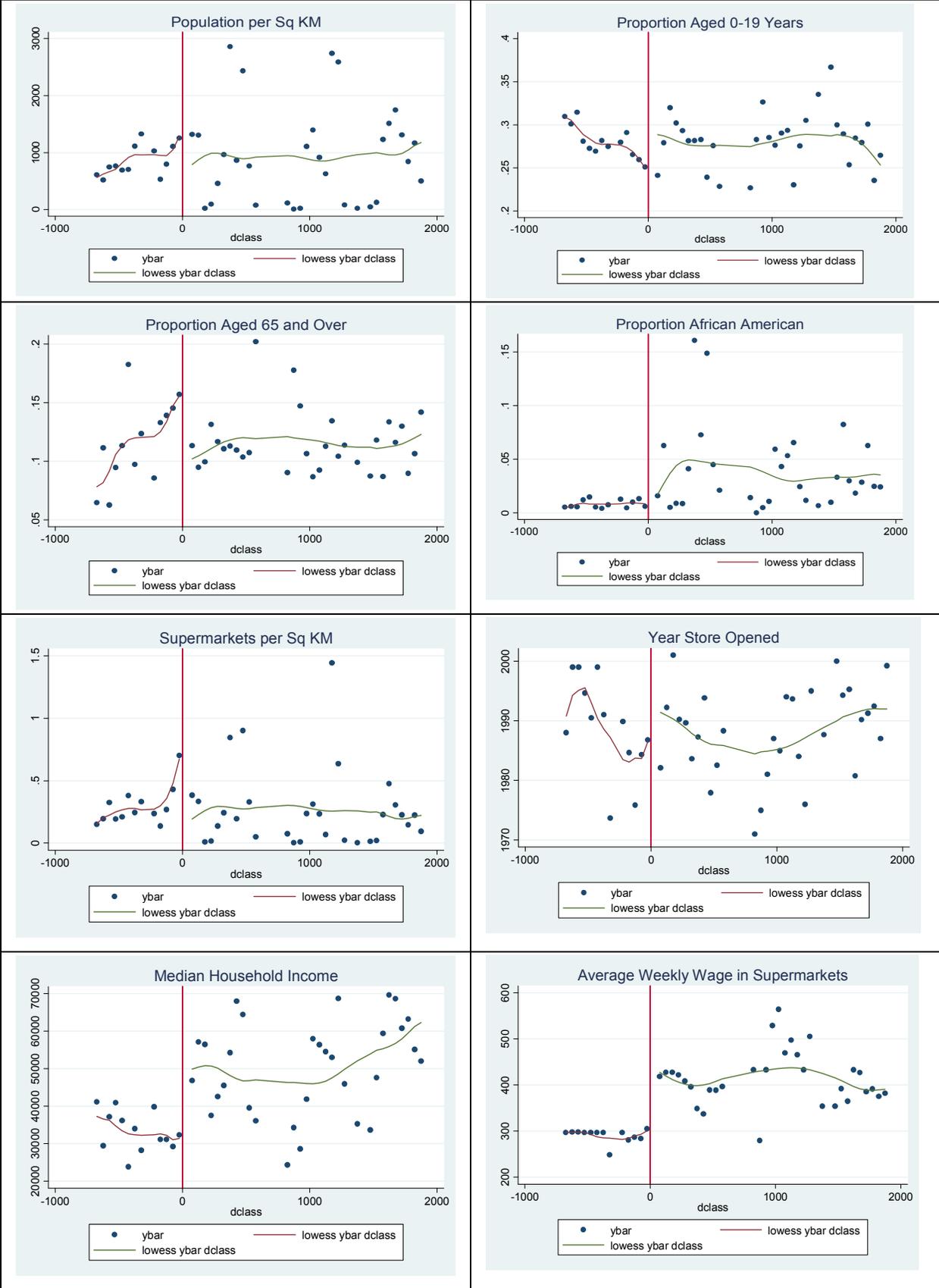
Figure 13: Graphical Depiction of Regression Discontinuity for Covariates
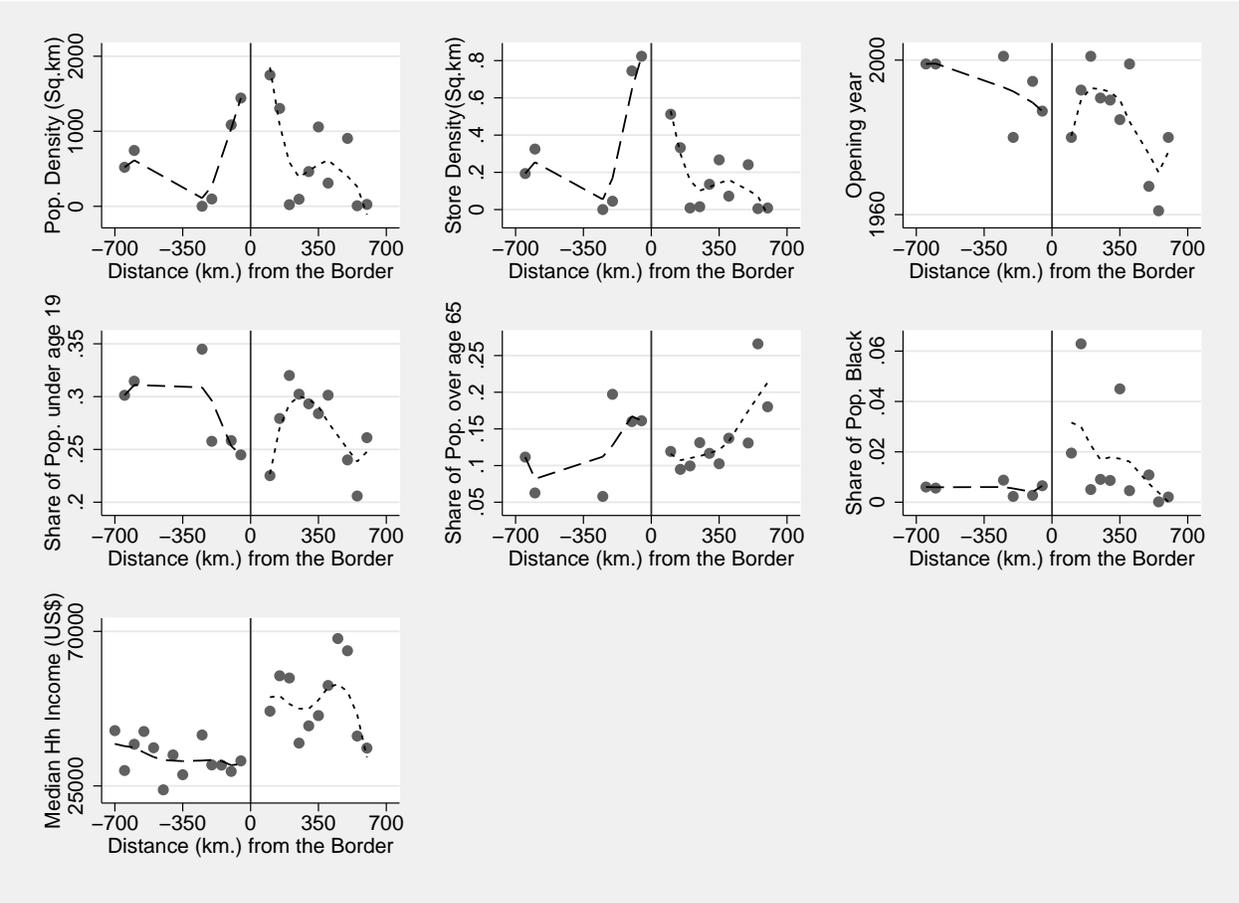
Figure 14: Graphical Depiction of Regression Discontinuity for Covariates (West Coast Stores)
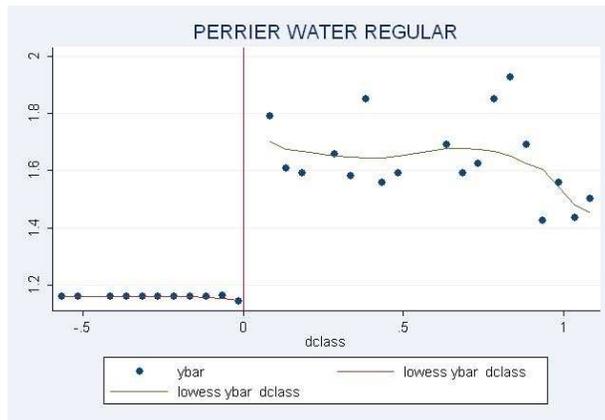
Figure 15: Graphical Depiction of Regression Discontinuity for Perrier Water Regular
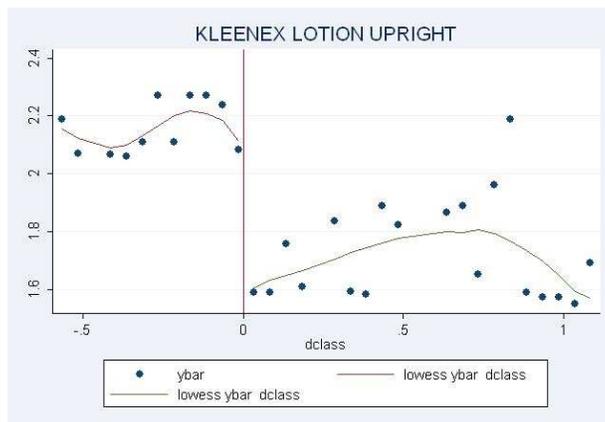

Figure 16: Graphical Depiction of Regression Discontinuity for Kleenex Lotion Upright
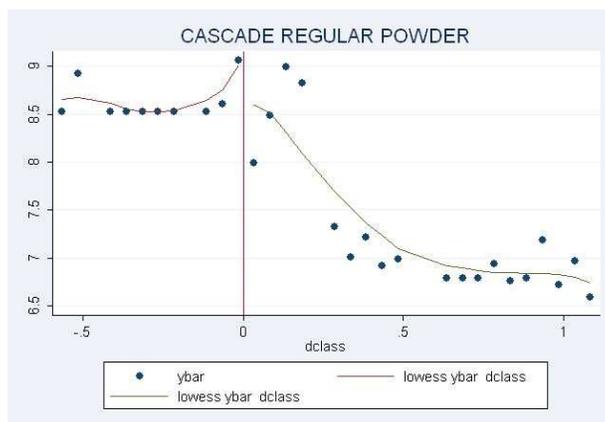

Figure 17: Graphical Depiction of Regression Discontinuity for Cascade Powder Regular

Figure 18: Distribution of border discontinuity in prices without (left) and with (right) covariates

Figure 19: Median border discontinuity of prices over time without (left) and with (right) covariates
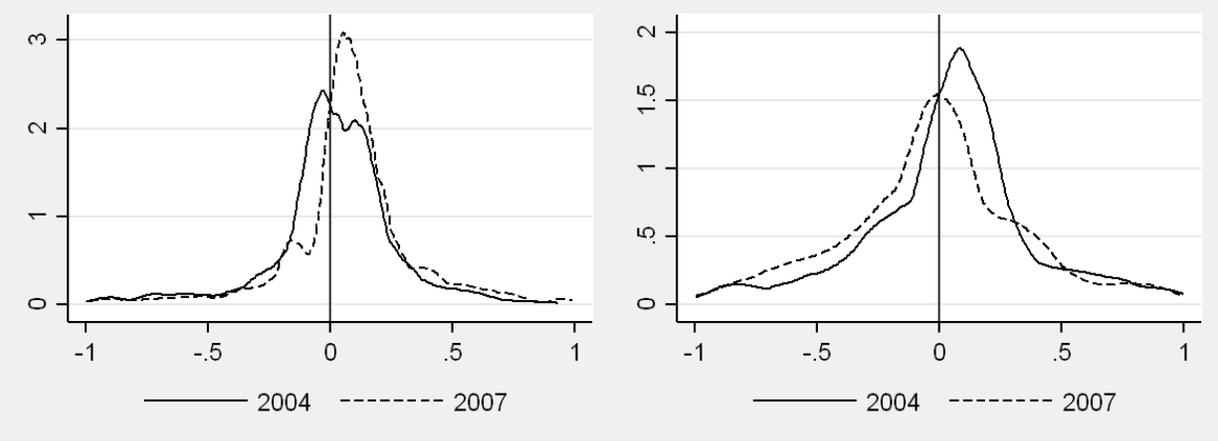
Figure 20: Distribution of Border Discontinuity in Imputed Costs (left) and Markups (right)
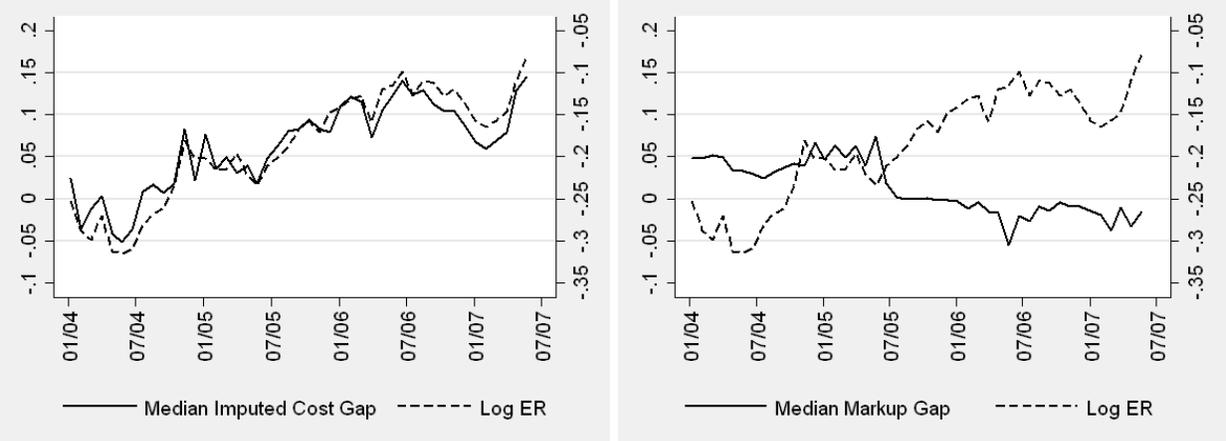

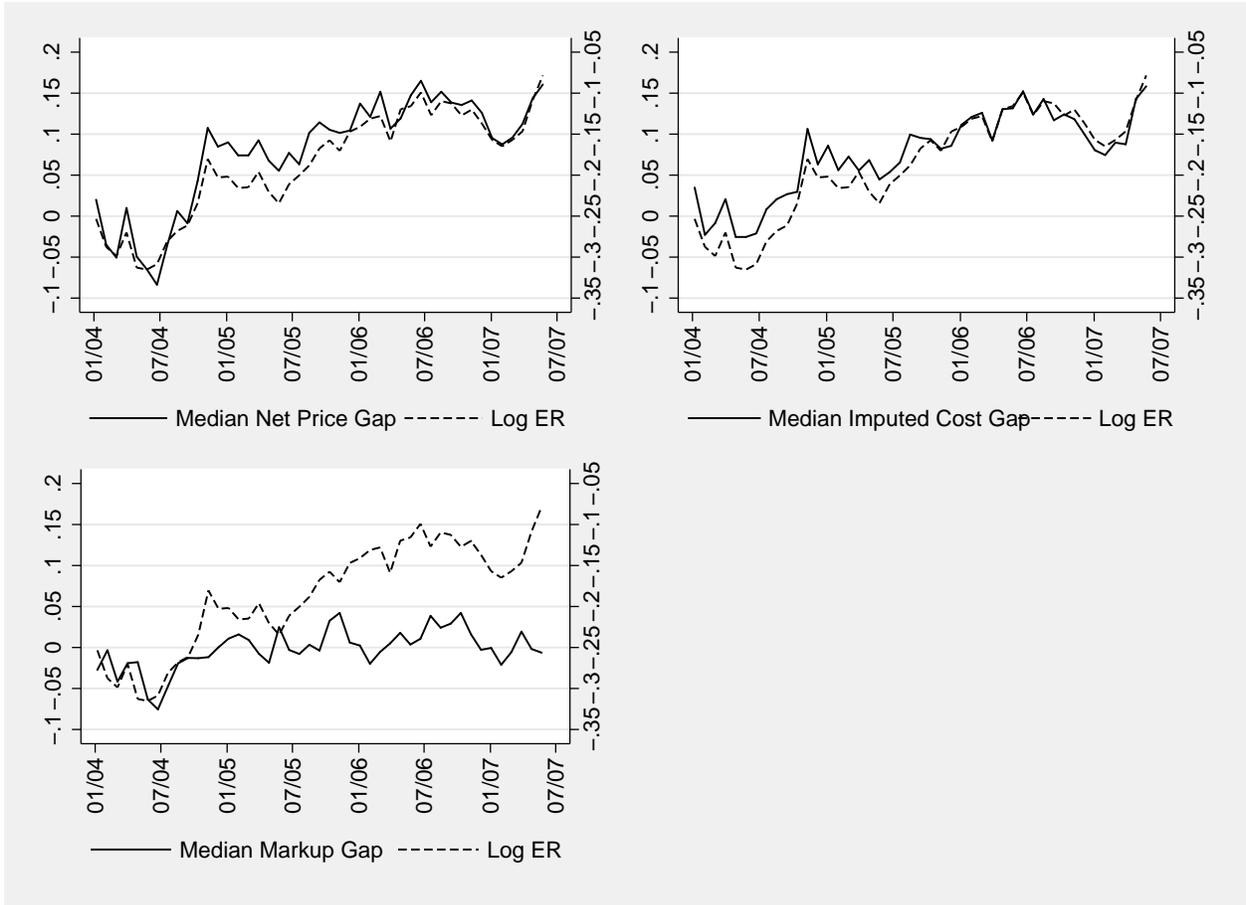Figure 21: Median border discontinuity of Imputed Costs (left) and Markups (right) over time

Figure 22: Median border discontinuity of Prices, Costs and Mark-Ups over time for West Coast Stores
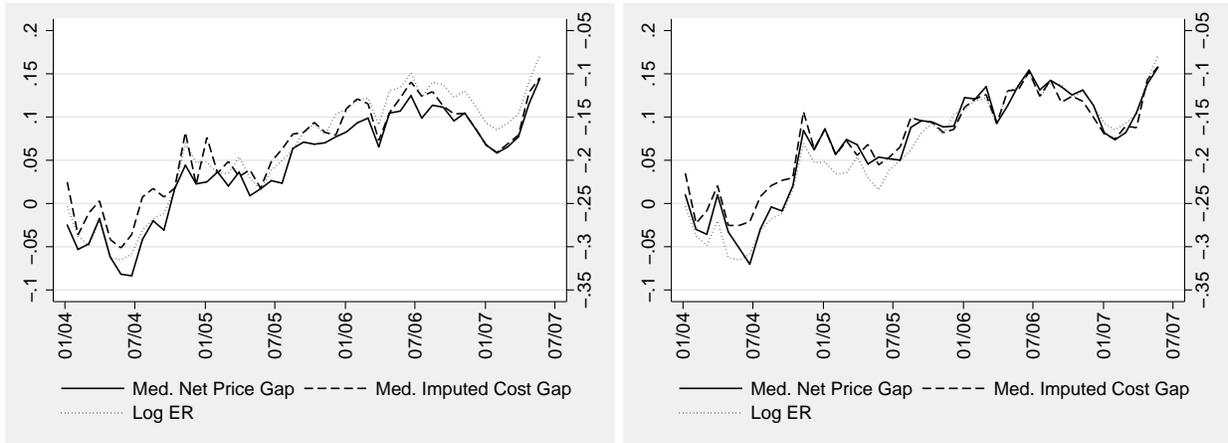

Figure 23: Co-movement in Price and Cost Gaps over time

A   Engel and Roger's (1996) Border Regressions

This appendix presents evidence on regressions ala Engel and Rogers (1996). While this is not our preferred specification, we present these estimates mostly for comparability with the earlier literature.

In Table 5 we regress the square of the (log) price difference or the absolute (log) price difference on log distance and a dummy for the border. We include store fixed effects and robust standard errors clustered by store-pair are reported in parenthesis. We find a very significant and positive effect of distance on price gaps. The border dummy in this regression is also sizeable and positive.[26] This is the case even if we compare only within-US pairs to cross-border pairs and exclude within-Canada pairs. (Gorodnichenko and Tesar (2005) found that when this regression was estimated using the data in Engel and Rogers (1996), the cofficient on the border became negligible.)

Broda and Weinstein (2007) also use barcode data for the US and Canada and reports statistics as in Tables 2-5. A quick comparison of our estimates to those in that paper suggests important differences. We find significant differences in the behavior of prices across borders, while Broda and Weinstein (2007) find that the price gaps across borders are similar to those within borders. Secondly, these authors find little heterogeneity across countries in the within-country price dispersion. The differences could arise because the nature of goods covered in the two studies can be significantly different, or, probably more importantly, because we observe the pricing behavior of a single retailer, while the Broda and Weinstein (2007) study potentially compares prices across retailers, over potentially large geographical units (a whole region for their Canadian data).

---

[26]The coefficient on the 'width of the border' is commonly defined as $\exp\left(\beta_{\text{border}}/\beta_{\text{log distance}}\right).$

Dependent Variable	Square of Price Difference	Absolute Price Difference
All Pairs		
Log Distance	0.336	1.366
	(0.004)**	(0.010)**
Border Dummy	7.926	13.473
	(0.016)**	(0.024)**
Observations	16320298	16320298
"Width" of the Border	1.79E+10	19141
Excluding CAN-CAN pairs		
Log Distance	0.36	1.45
	(0.003)**	(0.009)**
Border Dummy	7.06	13.22
	(0.518)**	(0.434)**
Observations	15334220	15334220
"Width" of the Border	3.29E+08	9111
Excluding US-US pairs		
Log Distance	0.06	0.44
	(0.014)**	(0.018)**
Border Dummy	10.78	22.01
	(0.497)**	(0.697)**
Observations	5230079	5230079
"Width" of the Border	1.07E+78	5.3E+21

Table 5: Engel-Rogers Border Regression. Coefficients and standard errors are multiplied by 10^2. ** Significant at the 1% level. Store dummies are included. Robust standard errors, clustered by store-pair, are in parentheses.