NBER WORKING PAPER SERIES

HOW WIDE IS THE BORDER?

Charles Engel
John H. Rogers

Working Paper No. 4829

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1994

We would like to thank the participants in the workshops at the Federal Reserve Board, the NBER Summer Institute, Penn State, Santa Cruz and Washington for very helpful comments. The National Science Foundation provided support for this project under grant number SBR-932078. This paper is part of NBER's research program in International Finance and Macroeconomics. Any opinions expressed are those of the authors and not those of the Board of Governors of the Federal Reserve System or the National Bureau of Economic Research.

NBER Working Paper #4829
August 1994

HOW WIDE IS THE BORDER?

ABSTRACT

Failures of the law of one price explain much of the variation in real C.P.I. exchange rates. We use C.P.I. data for U.S. cities and Canadian cities for 14 categories of consumer prices to examine the nature of the deviations from the law of one price. The distance between cities explains a significant amount of the variation in the prices of similar goods in different cities. But, the variation of the price is much higher for two cities located in different countries than for two equidistant cities in the same country. By our most conservative measure, crossing the border adds as much to the volatility of prices as adding 2500 miles between cities.

Charles Engel
Department of Economics
University of Washington
Seattle, WA 98195
and NBER

John H. Rogers
Board of Governors of the
Federal Reserve System
20th Street & Constitution Avenue, NW
Washington, DC 20551

The failure of the law of one price in international trade has been widely documented. It should be no surprise that similar goods sold in different locations have different prices. Indeed, Debreu (1959) in Theory of Value defines goods to be different if they are not sold in the same location: "Finally wheat available in Minneapolis and wheat available in Chicago play also entirely different economic roles for a flour mill which is to use them. Again, a good at a certain location and the same good at a different location are *different* economic objects, and the specification of the location at which it will be available is essential." Only when costs are borne to transport wheat from Chicago to Minneapolis will the miller in Minneapolis consider the Chicago wheat equivalent to the Minneapolis wheat. But, can the international failure of the law of one price be attributed entirely to this segmentation of markets by physical distance, or are there other factors, such as nominal price-stickiness, that help to explain the failure?

Recent evidence suggests that not only are failures of the law of one price significant, but that they play an important, if not dominant role in the behavior of real exchange rates. For example, Engel (1993) examines the conditional variance of the prices of similar goods across borders and compares that to the variance of the relative price of similar goods within borders. As that paper explains, much of the neoclassical theory of real exchange rate movements relies on variations of relative prices of different goods within a country's borders. In fact, Engel finds that with a few exceptions, the price of similar goods across borders is much more variable than within-country relative prices. For example, the price of men's clothing in the U.S. relative to the price of men's clothing in Canada has a greater conditional variance than the price of men's clothing in the U.S. relative to the price of restaurant meals in the U.S. and a greater variance than the price of men's clothing in Canada relative to the price of restaurant meals in Canada. Most of the variation in real exchange rates appears to be attributable to failures of the law of one price.

Furthermore, Rogers and Jenkins (1994) document that the persistence of the real exchange rate can mostly be explained by the persistence of failures of the law of one price. In the post-Bretton Woods era, one cannot reject at standard confidence levels that the real U.S. dollar/ Canadian dollar exchange rate follows a random walk. But, Rogers and Jenkins find that there is generally more persistence in the price of similar goods across borders than there is in the relative price of different goods within a country's borders. While it may be implausible that the prices of similar goods in Canada and the U.S. have a unit root (whether they are traded or not), it nonetheless appears that the high persistence of the real exchange rate is most directly linked to persistent deviations from the law of one price.

One reason that the price of similar goods might vary in different locations is that the markets for the goods are separated geographically. Recent work in international trade, spearheaded by Krugman (1991) and including empirical work by Frankel, Stein and Wei (1993) and McCallum (1993),[1] suggests that much of the pattern of international trade can be explained by geographical considerations. Countries are more likely to trade

[1] The McCallum paper is in a sense complementary to this one, in that it uses data from the states in the U.S. and provinces in Canada to measure the effects on the volume of trade of crossing international border as opposed to intranational borders.

with neighbors because transportation costs are lower. We note that transportation costs may be a reason why markets are segmented.

We examine the importance of transportation costs in this paper by relating measures of the spread or variance of the price of similar goods to the distance between markets. We employ consumer price data disaggregated into fourteen categories of goods. We make use of data available for nine Canadian cities and fourteen cities in the United States. The basic hypothesis is that the volatility of the price of similar goods between cities should be positively related to the distance between those cities.

However, we also entertain the possibility that the variance of the price of similar goods in two cities in different countries could be different than the volatility of the price in two cities equally far apart but in the same country. That is, the border matters. The recent literature on pricing to market (for example, Dixit (1989), Dornbusch (1987), Feenstra (1989), Froot and Klemperer (1989), Giovannini (1988), Kasa (1992), Knetter (1989, 1993), Krugman (1987) and Marston (1990)) has examined markets that are segmented by borders, although it has not paid particular attention to why the border as opposed to distance leads to market segmentation.

There are a few reasons why the border might matter. One reason is that the price of a consumer good might be sticky in terms of the currency of the country in which the good is sold. Goods sold in the U.S. might have sticky prices in U.S. dollar terms, and goods sold in Canada might have sticky prices in Canadian dollar terms. The nominal exchange rate is, in fact, highly variable. In this case, the cross-border prices would fluctuate along with the exchange rate, but the within-country prices would be fairly stable. There actually may be a connection between geographical market segmentation and price stickiness. It would be easier for a producer in one location to maintain a fixed nominal price if it were difficult for consumers to obtain goods from other locations, or if it were difficult for potential competitors to import supplies from nearby locations and enter the market undercutting the current sellers. So, market segmentation may reinforce price stickiness.

The sticky-price explanation is a natural one that has been addressed in previous literature. Our test is in part inspired by Mussa (1986), who noted that the variance of the real exchange rate based on all goods in the consumer price index is larger for Toronto versus Chicago, Vancouver versus Chicago, Toronto versus Los Angeles and Vancouver versus Los Angeles than it is for Toronto versus Vancouver and Chicago versus Los Angeles when there are floating exchange rates between the U.S. and Canada. He attributes this pattern to sticky prices, or, in his terms, nominal exchange regime non-neutrality. The result does not carry through to periods when exchange rates are fixed between Canada and the U.S.[2] Within the recent literature on pricing-to-market, Marston (1990) and Giovannini (1988) specifically consider the role of nominal price-stickiness.

An alternative explanation for why the border might matter springs from the observation that it is unlikely that there is much arbitrage that occurs in the final goods market. The price of the final good reflects not only the price of the actual product sold, but also the service that brings the good to the market (advertising, retailing, etc.) It is possible that much of the interregional variation in prices of a good reflect variation in the costs of this marketing service. These marketing services are likely to be highly

[2] Our study considers only a floating rate period, because data is not available for our array of cities and sub-categories of the CPI during fixed-exchange rate periods.

labor intensive. To the extent that national labor markets are more separated than local labor markets within a country, there would be more variation in cross-border prices than within-country prices.

Finally, there might be direct costs to crossing borders because of tariffs and other trade restrictions.

In section 1 we set out an extremely simple model that captures these elements. Section 2 discusses the data used in our study. In section 3 we present empirical results. The concluding section answers the question posed in the title of the paper and considers the implications of our findings.

1. A Model

This section lays out a simple model that is meant to serve as a point of reference for our empirical findings in section 3. We should emphasize that we view the results in section 3 as being exploratory in nature, and not meant to support one specific model over another. A full general equilibrium examination of a model of trade with transportation costs is contained in Dumas (1992).

Assume there are two locations. There is a monopoly producer of final output in each location. Final output is non-traded. Final output is produced using an intermediate input, which is traded. We take this approach because it captures in a rough way the notion that there is a non-traded element in the sale of final consumer goods. That element is the marketing service that accompanies the sale of the good. Few goods are of the homogeneous type sold in auction markets. There is a great deal of perceived and actual product variety. Marketers spend resources advertising and retailing goods. These marketing services typically are labor intensive. To the extent that labor services are immobile across regions, the marketing service will not be tradeable. This notion of trade in intermediate inputs but not in final products has been captured in a perfectly competitive environment by Sanyal and Jones (1982).

In the "home" location, $p_i = \beta_i c_i$, where p_i is the price of the product i in that location. There is a constant cost per unit, given by c_i. (That is, the unit cost is independent of the scale of production.) The markup over cost, β_i (> 1), is related inversely to the elasticity of demand (so, with a constant elasticity of demand ε_i, $\beta_i = \varepsilon_i/(\varepsilon_i-1)$).

We will assume that marketing services are provided by a factor of production whose return is w, and that this factor payment is the same for all goods at a specific location.

The cost per unit, c_i, is equal to $\alpha_i w^{\gamma_i} q_i^{1-\gamma_i}$. q_i is the price of the intermediate input used in the production of good i in the home location. The parameters α_i and γ_i are product specific. The share of marketing costs in total costs is given by γ_i. So, $p_i = \beta_i \alpha_i w^{\gamma_i} q_i^{1-\gamma_i}$.

In the "away" location, $\tilde{p}_i = \beta_i \alpha_i \tilde{w}^{\gamma_i} \tilde{q}_i^{1-\gamma_i}$. We let the parameters of the production function and demand curve be the same at the two locations for simplicity, although it would be trivial to generalize this. The price of the factor that provides marketing services, and the price of the intermediate good can be different across the locations.

The intermediate good can be traded, but at a cost which is related to distance. The intermediate good sold in the away location at price $\tilde{q}_i$ can be purchased at the domestic location for a price of $\tilde{q}_i d_i$, where d_i is the proportional transport cost, $d_i > 1$. (This notion of transport costs is referred to by Samuelson (1954) as "iceberg" transport costs. Only a fraction

$1/d_i$ of the intermediate good survives the move from the away location to the home location.) So, q_i must be $\le \tilde{q}_i d_i$. Likewise, if the intermediate good sells domestically at a price q_i, then the foreign producer can buy the good for a price of $q_i d_i$, so that $\tilde{q}_i \le q_i d_i$. This implies $\tilde{q}_i/d_i \le q_i \le \tilde{q}_i d_i$.

It follows that

(1) $$k_i d_i^{\gamma_i - 1} < p_i/\tilde{p}_i < k_i d_i^{1-\gamma_i},$$

where $k_i = (w/\tilde{w})^{\gamma_i}$.

Since our empirical work is in natural logs, it is useful to write

(2) $$\ln(k_i) - (1-\gamma_i)\ln(d_i) < \ln(p_i/\tilde{p}_i) < \ln(k_i) + (1-\gamma_i)\ln(d_i).$$

Transportation costs are probably relatively constant over time (so d_i is constant). If k_i is also constant, then the potential range of $\ln(p_i/\tilde{p}_i)$ is $2(1-\gamma_i)\ln(d_i)$. So, the potential range of variation is related only to the transportation costs, and the share of marketing services in the final product.

In fact, k_i is probably not constant, in which case the potential range of $\ln(p_i/\tilde{p}_i)$ is

$$\mathrm{Max}(\ln(k_i)) - \mathrm{Min}(\ln(k_i)) + 2(1-\gamma_i)\ln(d_i).$$

Before we introduce sticky prices and discuss how the currency the good is priced in may matter, it is helpful to summarize why the border might be important in determining the range of variation in $\ln(p_i/\tilde{p}_i)$ when prices are flexible. First, d_i might be related to crossing a border. That is, the transportation costs might be higher if the border is crossed because of barriers to trade. We have modeled transportation costs as being symmetric. Even if trade barriers on the intermediate goods were present in only one direction, or if the trade barriers were not symmetric, there would still be a range in which the relative prices of final goods could fluctuate.

Moreover, $Max(ln(k_i)) - Min(ln(k_i))$ may be related to the border. For example, while there is labor market segmentation between any two locations, it is apt to be greater between locations in different countries. There are cultural and political reasons why workers would not be willing to cross international borders in order to take advantage of a higher wage, even when they are willing to relocate within a country. So, variation in the price of non-traded leisure could lead to a larger value for $Max(ln(k_i)) - Min(ln(k_i))$, which reflects the range of variation in the relative marketing costs for the final consumer goods.

Now, suppose prices have to be set one period ahead in terms of the domestic currency. Let p_i be the dollar price in the home location. Let p_i^* be the Canadian dollar price, so sp_i^* is the U.S. dollar price, where s is the U.S. dollar per Canadian dollar exchange rate. From above, if prices are set a period ahead of time, assuming costs are incurred a period ahead of time:

(3) $$p_i(t) = \beta_i \alpha_i w(t-1)^{\gamma_i} q_i(t-1)^{1-\gamma_i}$$

Here, q_i refers to the U.S. dollar price of the intermediate good, and w is the dollar cost of marketing. The price in period t depends on the costs in period t-1. Similarly,

(4) $$p_i^*(t) = \beta_i \alpha_i w^*(t-1)^{\gamma_i} q_i^*(t-1)^{1-\gamma_i}$$

These prices are expressed in Canadian dollars.

We shall assume that the production costs are incurred a period ahead of time, when the nominal price is set. So,

(5) $$k_i(t-1))-(1-\gamma_i)ln(d_i) < ln(p_i(t)/s(t-1)p_i^*(t) < ln(k_i(t-1))+(1-\gamma_i)ln(d_i),$$

where $k_i(t-1) = (w(t-1)/s(t-1)w^*(t-1))^{\gamma_i}$

Of course our data on relative prices across borders compares prices using the current exchange rate. That is, we have data on $p_i(t)/s(t)p_i^*(t)$. We can write

$$(6) \quad \ln(k_i(t-1))-(1-\gamma_i)\ln(d_i)-\Delta\ln(s(t)) < \ln(p_i(t)/s(t)p_i^*(t)) < \ln(k_i(t-1))+(1-\gamma_i)\ln(d_i)-\Delta\ln(s(t)),$$

where $\Delta\ln(s(t)) \equiv \ln(s(t)) - \ln(s(t-1))$.

Now the range of potential fluctuation in $\ln(p_i(t)/s(t)p_i^*(t))$ is

$$\mathrm{Max}(\ln(k_i(t-1))-\Delta\ln(s(t))) - \mathrm{Min}(\ln(k_i(t-1))-\ln(d_i)-\Delta\ln(s(t))) + 2(1-\gamma_i)\ln(d_i).$$

The border will matter because of the exchange rate fluctuation, as well as for the other reasons discussed above.

In our empirical work, our measure of the dispersion of $\ln(p_i(t)/s(t)p_i^*(t))$ between locations is not exactly this range of potential fluctuation. First, equation (6) gives expressions for the minimum and maximium possible value of the price variable, but the actual price variable may not hit these bounds. Second, we do not use the spread between the minimum and maximum values of $\ln(p_i(t)/s(t)p_i^*(t))$ as our measure of dispersion. Instead we use measures of volatility such as the standard deviation or the spread between the 10th and 90th percentile of $\ln(p_i(t)/s(t)p_i^*(t))$. The minimum and maximum values might reflect unusual events that we want to exclude for reasons of robustness. Still, under reasonable assumptions it is likely that our measure of the variation in $\ln(p_i(t)/s(t)p_i^*(t))$ is related to distance and the border.

For example, one measure of the variation in prices that we employ in section 3 is the spread between the 90th percentile and 10th percentile of $\ln(p_i(t)/s(t)p_i^*(t))$. We can write

$$p_i(t)/s(t)p_i^*(t)= (w(t-1)^{\gamma_i}q_i(t-1)^{1-\gamma_i})/(s(t)w^*(t-1)^{\gamma_i}q_i^*(t-1)^{1-\gamma_i}).$$

So,

$$\ln(p_i(t)/s(t)p_i^*(t)) = \ln(k_i(t-1))-\Delta\ln(s(t))+(1-\gamma_i)\ln(x_i(t)),$$

where $x_i(t) = q_i(t-1)/s(t-1)q_i^*(t-1)$. We have that $\ln(x_i(t))$ is bounded above by $\ln(d_i)$ and below by $-\ln(d_i)$. For example, suppose $(1-\gamma_i)\ln(x_i(t))$ takes on the value $(1-\gamma_i)\ln(d_i)$ with probability 1/3, $-(1-\gamma_i)\ln(d_i)$ with probability 1/3, and 0 with probability 1/3. Suppose that $\ln(k_i(t-1)) - \Delta\ln(s(t))$ takes on three values: $-z$ with probability 1/3, 0 with probability 1/3, and z with probability 1/3. Assume the distributions are independent, and $z > 2(1-\gamma_i)\ln(d_i)$. The probability table is given by:

Value	Probability
$(1-\gamma_i)\ln(d_i)+z$	1/9
z	1/9
$z-(1-\gamma_i)\ln(d_i)$	1/9
$(1-\gamma_i)\ln(d_i)$	1/9
0	1/9
$-(1-\gamma_i)\ln(d_i)$	1/9
$-z+(1-\gamma_i)\ln(d_i)$	1/9
$-z$	1/9
$-z-(1-\gamma_i)\ln(d_i)$	1/9

The upper and lower 10th percentiles depend on $(1-\gamma_i)\ln(d_i)$ and on the volatility of $\ln(k_i(t-1)) - \Delta\ln(s(t))$.

The model of this section demonstrates that the volatility of the price of a good in two different locations might be related to the distance between those two locations. In our model, this comes about not because the final good might be transported and goods arbitrage occurs with final products. Instead, it occurs because of arbitrage in the intermediate goods market. We also argue that the volatility of the final goods price might be different between two equidistant cities if those cities are located in the same country and if they are not. That is because, first, there is more labor market segmentation between countries than within countries, so that there can be more volatility in relative marketing costs. Also, transportation costs may be incurred at border crossings. And, if there is nominal price stickiness, there will be fluctuations in cross-border prices as the nominal exchange rate varies.

2. Data

We use consumer price data from 23 North American cities for 14 disaggregated consumer price indexes. The data cover the period June 1978 to June 1993.

For our purposes, it is natural to choose the U.S. and Canada as the countries to study. First of all, the countries are adjacent -- they are not separated by large bodies of water. Were it not for the country borders, one would expect more trade to occur between Toronto and New York than between New York and Los Angeles. Indeed, there are no other examples of market economies that are as large in area (so that there can be significant distances between major cities within a country) and adjacent. Also, trade has been relatively free between the two countries. If the border matters, it is unlikely that it matters because of trade restrictions. The fact that both countries are mostly English-speaking and have similar cultural and political traditions suggests that there is likely to be more cross-border labor migration than between most countries.

Our data from the U.S. was obtained from the Bureau of Labor Statistics. The fourteen goods from the U.S. are listed on the left hand side of Table 1. All of the price data (for both countries) are seasonally unadjusted.

We use comparable price data for Canada that was obtained from Statistics Canada. There is not always an exact match between the price indexes available in Canada and those available in the U.S. However, we were able to construct indexes for the 14 categories of goods in Canada, in some cases by using even more disaggregated Canadian indexes. For example, the U.S. data contains a series on Men's and Boy's Apparel. There is no comparable series in Canada. However, we can obtain from Canada individual series on Men's Wear and Boy's Wear. We then construct a Men's and Boy's Apparel series for Canada by taking a weighted average of the Men's Wear series and the Boy's Wear series.[3] This type of construction was needed to arrive at five of the fourteen Canadian price series.[4] Table 1 indicates how these series were derived.

These categories of goods are mutually exclusive. Together they comprise 94.6% of purchases (using the weights in the U.S. consumer price index).

Monthly price data was used for nine Canadian cities: Calgary, Edmonton, Montreal, Ottawa, Quebec, Regina, Toronto, Vancouver and Winnipeg.

[3] The weights come from the current weights used in the U.S. consumer price index, which we obtained from the BLS.

[4] We did not use what may seem like the more obvious approach of comparing prices of Men's Apparel across countries and the price of Boy's Apparel across countries, because on a city by city basis data is not available from the BLS at that disaggregated a level for the U.S.

Monthly price data for the U.S. is available for four "core" cities: New York, Philadelphia, Chicago and Los Angeles. In addition, for five cities there is data that is released in even-numbered months: Dallas, Detroit, Houston, Pittsburgh and San Francisco. For five other cities, there is data available in odd-numbered months: Baltimore, Boston, Miami, St. Louis, and Washington.[5]

Consumer price data is closer to being monthly average data than point-in-time data. Typically to get the price of a single product, several outlets are sampled during the month. The outlets are not all sampled on the same day. The change in the price of the product from the previous month is calculated as the average change across the various outlets. For the cities that report data every second month, the prices are for the second month of the interval (rather than an average across both months.)

In order to nullify a potential bias, we use a monthly average U.S. dollar / Canadian dollar exchange rate from the Citibase tape. Averaging tends to reduce the volatility of the series. So, if we were to use an exchange rate at a specific point in time, but use price data which is essentially averaged, we would introduce volatility into our measure of cross-border prices. That is compensated for by taking the monthly average exchange rate.

For each good, we calculated the inter-city relative prices. Thus, when we are using only the Canadian cities and the core U.S. cities, for each good there are 78 inter-city prices (13 cities x 12/2). Adding the five even month

[5] Data for Cleveland is available every other month. However, the data switched from being odd month to even month in the middle of our sample. Also, at the beginning of the sample, Detroit data was monthly, but switched to even month, while the reverse is true for San Francisco. We make use only of the even month data for these two cities.

U.S. cities adds another 75 prices, and adding the five odd month U.S. cities adds another 75 prices.

We also use data on the distance between cities. We use two separate measures of distance, both obtained from the Automap (version 2) software. One measure is the great circle distance, and the other is the quickest driving time distance. Our results were not affected by the choice of distance measure, so all results reported use the great circle distance.

Table 2 presents selected summary statistics. There seem to be two general characteristics of the data in this table. First, by any of our measures, the volatility of prices between two Canadian cities is lower on average than the volatility between a U.S. city and another city on average (whether the city pairs are both in the U.S. or one is in the U.S. and one in Canada.) Second, the average distance between cities within Canada and within the U.S. are comparable, and slightly less than the average cross-border distances. It is not at all obvious from these summary statistics that the border would have much explanatory power for price dispersion, but that is what we shall find in the next section.

3. Empirical Results

In this section we try to explain the volatility of the prices of similar goods sold in different locations by the distance between the locations and other explanatory variables, including a dummy variable for whether the cities are in different countries.

Let P^i_{jk} be the log of the price of good i in location j relative to the price of good i in location k. We calculate the volatility of P^i_{jk} in two different ways. One is simply the standard deviation of P^i_{jk}, and the other is the difference between the 90th percentile value of P^i_{jk} and the 10th percentile value of P^i_{jk}.

We take three different measures of P^i_{jk}. One is the actual log of the relative price. The second is the first difference in the log of the relative price. The third is a filtered measure of the relative price.

For the third measure of P^i_{jk}, we regress the log of the relative price on twelve seasonal dummies and six monthly lags (or in the case of the data that is bi-monthly, on three bi-monthly lags and six seasonal dummies). We then take the two-month ahead in-sample forecast error from this regression as our measure of P^i_{jk}. The in-sample forecast errors cover the period February 1979 to June 1993. The notion is that we are capturing a measure of the unanticipated component of short-run price shocks. This is consistent with the explanation for volatility in P^i_{jk} presented in section 1. The volatility arises from shocks to intermediate goods prices. If there are differences in prices of intermediate goods that are anticipated in advance, less costly ways of moving goods between locations can be worked out. There will be less volatility caused by anticipated changes in intermediate goods prices than by unanticipated changes in these prices.

The dependent variable in our regressions, then, is a measure of the volatility of P^i_{jk}. We have six such measures -- the standard deviation and the spread between the 90th and 10th percentile for each of the three measures of P^i_{jk}.

As described in section 2, for each good i, when we confine ourselves to analysis of those cities for which we have monthly data (the Canadian cities and the four core U.S. cities) there are 78 values of P^i_{jk} at each date. Since we have 14 goods, we have a total of 1092 cross-sectional data points for P^i_{jk} at each date. Hence, for each volatility measure we have 1092 observations on the dependent variable.

When we include the data for the U.S. cities for which there is bi-monthly data, we add 150 more observations of the dependent variable for each good. That gives us a total of 3192 observations in these regressions.

As explanatory variables for the volatility of P^i_{jk} we first include a measure of the distance and the square of the distance between location j and k. We hypothesize that the distance has a positive effect on the volatility of P^i_{jk}, but that the coefficient on the square of distance is negative. The reason we think that there is a concave relation between the volatility of P^i_{jk} and distance is that at longer distances shippers might employ air or rail that would not be economical over short distances. So, shipping costs do not rise linearly with distance.

We also include a dummy variable for whether locations j and k are in different countries. For reasons we have explained, we expect the coefficient on this variable to be positive.

Because the different goods may have different degrees of price volatility, we include a dummy variable for each good. Since there is a constant term, we do not include a dummy for the first good.

In the regressions that use the data from the U.S. cities that only report data bi-monthly, we include dummy variables for whether either location j or k (or both) is from one of the cities that reports data in even months, and another for cities that report data in odd months. We do this so that we can account for additional volatility that may be introduced by measurement error from the less frequent observation of prices.

An issue we must consider is that U.S. prices may be more volatile than Canadian prices even for equidistant cities. For example, the U.S. might be a more heterogeneous country. Labor markets could be less well integrated, or goods markets may be less integrated, so there can be greater discrepancies in prices between locations. Alternatively, there may be differences in methodologies for recording prices that leads to greater discrepancies in prices between locations in one country compared to the other. To deal with this problem, we include a dummy variable if the price involved a U.S. city -- that is, if both cities were in the U.S., or if one was in the U.S. and one was in Canada.[6]

Table 3 reports the results from the regressions when P^i_{jk} is measured as the output from the filter. Regressions 1 and 2 are for the core cities, using the standard deviation of prices and the spread between the 90th and 10th percentile of prices, respectively. The analogous regressions for all the cities are labelled Regressions 3 and 4.

First, consider the results from the analysis of only the cities that report monthly data. The results are very similar whether volatility is measured as the standard deviation or the spread. The distance and distance-squared variables are significant and of the hypothesized sign.[7] So, as the recent trade literature suggests, geography is important. To some extent the failure of the law of one price can be explained by the market segmentation caused by distance between cities.

---

[6] Distance and the border turn out to be highly significant in the regressions whether or not this dummy variable is included.

[7] All t-statistics reported use White's (1980) heteroskedasticity-consistent measure of the standard errors.

But, the border variable is also highly significant and large in magnitude. In the Conclusions section we discuss a method of comparing the economic significance of the distance variables to the border dummy.

Table 3 also reports results using the entire sample of cities. The results are essentially unchanged. The distance variables are significant and of the expected sign. The border dummy is highly significant. Indeed, both distance measures and the border dummy become more significant with the entire sample of cities as compared to the sample of core cities.

In all of the regressions, the coefficient on the dummy variable for whether the price includes a U.S. city is positive, large and significant. U.S. prices seem to be more variable than Canadian prices. We also find that the dummies for the cities in which prices are measured in even months and in odd months are positive and highly significant, suggesting that the less-frequent measurement induces added variability.

Table 4 reports the results for the regressions in which P^i_{jk} is measured simply as the log of the relative prices, and Table 5 reports the regressions when P^i_{jk} is measured as the first difference in the log of the relative prices. The qualitative results are exactly as reported for the regressions using the filter measure of P^i_{jk}.

We tried several extensions to test the robustness of our results. One variation we tried was to alter the period covered by the data. We eliminated the early 1980s from our sample. This was the period in which the U.S. dollar experienced large swings in its value. We estimated the model over the period September 1985 to June 1993. Again, there was virtually no change in the results in these regressions.

We also split the sample at January 1990, when the Canadian - U.S. Free Trade Agreement went into effect. If trade barriers are an important reason why the border variable is economically significant in explaining price dispersion, one would expect that the magnitude of this variable would decline after 1989. In fact, we found a slight tendency in the opposite direction -- the estimated border coefficients were larger in the post-1989 period.

We also estimated the models in Tables 3-5 with a separate dummy variable for each city (but no dummy variable for U.S. cities as a group, since that would introduce perfect collinearity).[8] We still find that the distance variables are significant at the 95% level and of the hypothesized sign. However, the standard errors of the coefficient estimates are larger, and the t-statistics smaller on the distance variable. This may reflect that there is some collinearity between distance and some of the city dummies. For example, the dummy variables for Miami and Vancouver are positively correlated with distance. We also find in these regressions that the border dummy is significant. In fact, the coefficients on the border dummy are larger than in Tables 3-5, and the t-statistics are now around 30.

There is a problem with the quadratic specification of the distance measure in our regressions: the derivative of volatility with respect to distance turns negative at a distance of about 1500 miles. This might imply that distance is in fact not a useful indicator of how segmented the markets are. However, given that both the linear term and the squared distance term are strongly statistically significant, we believe that there is evidence for a concave relation between the volatility of prices and distance.

[8] Also, in this specification, there was no constant term, but a dummy variable for each of the 14 goods. Again, the point here is to avoid perfect collinearity.

In Table 6 we report the results of another specification of the functional form relating volatility to distance. In this table, we relate the volatility to the log of distance. Regression 1 in table 6 measures relative prices as the filtered price, and measures volatility using the standard deviation. Regression 2 measures prices in the same way and takes the spread between the 10th and 90th percentiles as the measure of volatility. Both regressions use the data from all of the cities. We report the results only from these two regressions, but we find the same results qualitatively for both of our measures of volatility, all of our measures of the price and using either all of the cities or just the core cities.

We find that the distance variable is positive and significant. The coefficients on the dummy variables for the border remain highly significant (as do the other coefficients we discussed in relation to Tables 3-5.) We did not perform a formal specification test to compare this regression to the analogous regressions (Regressions 3 and 4 from Table 3) using the quadratic specification of the volatility-distance relationship. However, we note that the adjusted R-squared statistics reported in the two tables are identical at .74 for the standard deviation measure of volatility and slightly larger (.76 as compared to .75) for the spread measure of volatility. So, the log distance specification performs at least as well as the quadratic specification.

The fact that the derivative of volatility with respect to distance turns negative at 1500 miles using the quadratic specification suggests, of course, that there are some distant cities that have fairly stable prices relative to each other. This observation leads us to the hypothesis that firms in some cities are price leaders. Specifically, we hypothesize that New York and Los

Angeles are price leaders. Firms in other cities in part adjust their prices to match price changes in these two large cities. To test this hypothesis, we included a dummy variable for whether the city pair included either Los Angeles or New York. The results of these regressions are reported in Table 6 as Regression 3 and Regression 4.

The specification of Regressions 3 and 4 in Table 6 are the same as those of Regressions 1 and 2 in that table, except that the L.A.-N.Y. dummy variable is included. We find that the coefficient on the L.A.-N.Y. dummy is negative, as expected, and highly significant. Furthermore, compared to regressions 1 and 2, the coefficients on the distance variable are larger, and the coefficients on the border dummy are smaller (although the t-statistics are still large). The implications for interpreting the relative strengths of the distance and border effects are discussed in the next section.

One other convex specification of the distance variable we tried is one in which we hypothesize that after a certain critical distance (arbitrarily chosen to be 1700 miles) additional distance does not contribute at all to volatility. So, in our alternative model, there is a linear relation between volatility and distance for distances up to 1700 miles, and then after 1700 miles the derivative of volatility with respect to distance is zero. This functional form is implemented by coding all distances that are over 1700 miles as being exactly 1700 miles, then regressing volatility on distance and the other variables included in the regression. The results are similar qualitatively to all of the other results we have reported, although, as we shall discuss in the next section, with this specification the border variable has the smallest quantitative effect on price dispersion relative to the distance variable among all of the specifications we tried.

The results in Tables 3-6 impose the restriction that the effect of crossing the border and the effect of adding distance be the same for all goods. The only variation allowed between goods was in the intercept term for the volatility. Table 7 considers the effects of distance and crossing the border on volatility for each of the goods separately. For each good, Table 7a reports the regressions of the standard deviation of the filtered price data on a constant, a dummy for whether the cities are in different countries, distance and distance squared, and a dummy variable for whether one of the cities was a U.S. city. There are also dummy variables for whether the price involves a city whose data is reported only in even-numbered or odd-numbered months. Table 7b reports similar regressions, but uses the log-distance specification.

The results of Table 7a are strongly supportive of the view that distance matters for volatility, but that the border is also extremely important. The coefficient on the border dummy is positive and highly significant for twelve of the fourteen goods. Only for two clothing variables -- women's and girls' clothing and footwear -- does the dummy variable have the wrong sign. It is not significantly negative for those two goods, however. The distance variables all are of the right sign except for women's and girls' clothing and entertainment, and they are almost always significant. For the two goods where the coefficient on distance is negative, the coefficient on distance squared is positive. So, while the hypothesis that distance and the border increase market segmentation does not hold for every good, it holds well across most of the goods.

The results in Table 7b are similar, except that the distance variable does not perform as well. The distance variable is of the wrong sign for three of the fourteen goods, but it is not significantly negative for any of these goods. It is significantly positive for six of the fourteen goods, and insignificantly positive for the other five goods. The border dummy has the hypothesized sign for twelve of the fourteen goods, and is significant in all of these cases. In the two cases in which it is negative, it is insignificant. So, in no case is there a significant distance or border coefficient that is of the incorrect sign.

The next section examines the relative importance for volatility of distance and of crossing the border.

4. Conclusions

We have seen that physical distance plays a significant role in explaining the failure in the law of one price between two locations. But, physical distance alone does not explain the variability in prices of similar goods if the two locations are in different countries -- the border matters.

It is useful to get a handle on how important the border is relative to physical distance. Let $\hat{b}_1$ be the estimated coefficient on the distance variable, and $\hat{b}_2$ be the estimated coefficient on the squared distance variable in the regressions reported in Tables 3-5. The effect of adding one more mile of distance between cities on the volatility of relative prices is equal to $(\hat{b}_1+2\hat{b}_2\cdot\text{distance})$. We must evaluate this derivative at a particular distance. A natural value to evaluate this derivative is the average distance between cities. The average distance between all of the cities in our sample is 1230 miles. So, the effect of adding a mile of distance on volatility is $\hat{b}_1+2460\hat{b}_2$.

We want to compare the effect of adding a mile of distance to the effect of crossing the border. So, consider the regression in Table 2 that includes the data for all of the cities, where the measure of volatility is the standard deviation of P^{i}_{jk} (Regression 3 in Table 2). The effect on the standard deviation of adding one more mile of distance is $4.56\times10^{-6} - 2460\times(1.62\times10^{-9}) = 5.75\times10^{-7}$. The effect of crossing the border is to increase the standard deviation by 4.21×10^{-3}. Hence, crossing the border is equivalent to adding $(4.21\times10^{-3})/(5.75\times10^{-7}) = 7182$ miles between the cities!

This measure of the relative importance of the border is typical in our regressions. For each regression in Tables 3-5, we report an "Implied Width of Border". In Tables 3-5, these distances range from 6481 miles to the largest of 23,383 miles.[9]

Table 6 uses the log of distance specification. Again, we evaluate the effect of distance at the average distance and find that the border ranges from a width of 4310 miles up to 9663 miles. As we explain in Section 3, one other specification we used was one in which the distance relation is linear and positive up to 1700 miles, and flat after that. The coefficient on distance for distances less than 1700 miles is taken as the derivative of volatility with respect to distance. When the Los Angeles - New York dummy variable is included, we find that crossing the border adds as much to volatility as traveling 2500 miles within a country.[10] Of all of our regressions, this is the one in which the relative effect of the border is the smallest.

---

[9] In fact, for the regressions using the logs of the levels of the prices, the implied width of the border is much larger than for the other regressions.
[10] Under the assumption that distance adds to volatility as it would if cities were less than 1700 miles apart.

The conclusions are striking. Apparently two cities that lay directly across the border will find larger failures of the law of one price than two cities on the opposite ends of the same country.

The major message of our empirical results is not just that the border matters for relative price variability -- it is that both distance and the border matter. The literature on pricing-to-market has emphasized that when markets are segmented that price discrimination can occur. The finding that distance is important in explaining price differences between locations lends support to this literature and the associated work on geography and trade. But our findings seem to suggest that there is more than standard price discrimination behavior involved in cross-border price movements. If tariffs and trade barriers were significant between the U.S. and Canada, our finding of a strongly significant border effect would be consistent with typical models of market segmentation. However, among pairs of countries in the world, the trade barriers between the U.S. and Canada are quite low, and the border variable has a large influence on volatility even after the Free Trade Agreement went into effect. It does not seem possible that trade restrictions can account for the degree of cross-border price variability that we find.

Our study still cannot determine whether segmented labor markets which are reflected in variation in marketing costs, or sticky nominal prices account for the importance of the border variable. If our results do in fact indicate sticky nominal prices, then they also shed some light on the price-setting process. We have found that the distance between markets influences prices, suggesting that price setters take into account prices of nearby competitors. It is probably not too far-fetched to infer that firms would respond more to changes in prices of near substitutes, whether the nearness is in geographical space or product space. For example, when the price of Crest toothpaste changes, producers of Colgate toothpaste respond to a greater extent than they would when the price of Dial soap changes. So, a reasonable model of price stickiness must take into account how isolated the market is for the product of the price setter. There appears to be grounds for a marriage between the New Keynesian literature on menu costs and the new trade literature emphasizing the role of geography.

References

Debreu, Gerard, 1959, Theory of Value, (New Haven: Yale University Press).

Dixit, Avinash, 1989, "Hysteresis, Import Penetration, and Exchange Rate Pass-Through," Quarterly Journal of Economics 104, 205-228.

Dornbusch, Rudiger, 1987, "Exchange Rates and Prices," American Economic Review 77, 93-106.

Dumas, Bernard, 1992, "Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World," Review of Financial Studies 5, 153-180.

Engel, Charles, 1993, "Real Exchange Rates and Relative Prices: An Empirical Investigation," Journal of Monetary Economics 32, 35-50.

Feenstra, Robert C., 1989, "Symmetric Pass-Through of Tariffs and Exchange Rates under Imperfect Competition: An Empirical Test," Journal of International Economics 27, 25-45.

Frankel, Jeffrey; Ernesto Stein and Shang-Jin Wei, 1993, "Continental Trading Blocs: Are They Natural or Super-Natural?", National Bureau of Economic Research working paper no. 4588.

Froot, Kenneth A. and Paul D. Klemperer, 1989, "Exchange Rate Pass-Through when Market Share Matters," American Economic Review 79, 637-654.

Giovannini, Alberto, 1988, "Exchange Rates and Traded Goods Prices," Journal of International Economics 24, 45-68.

Kasa, Kenneth, 1992, "Adjustment Costs and Pricing-to-Market: Theory and Evidence," Journal of International Economics 32, 1-30.

Knetter, Michael N., 1989, "Price Discrimination by U.S. and German Exporters," American Economic Review 79, 198-210.

Knetter, Michael N., 1993, "International Comparisons of Pricing-to-Market Behavior," American Economic Review 83, 473-486.

Krugman, Paul, 1987, "Pricing to Market When the Exchange Rate Changes," in Sven W. Arndt and J. David Richardson, eds., Real-Financial Linkages among Open Economies (Cambridge: M.I.T. Press).

Krugman, Paul, 1991, "Increasing Returns and Economic Geography," Journal of Political Economy 99, 483-499.

Marston, Richard C., 1990, "Pricing to Market in Japanese Manufacturing," Journal of International Economics 29, 217-236.

McCallum, John, 1993, "National Borders Matter: Regional Trade Patterns in North America," Department of Economics, McGill University.

Rogers, John H., and Michael Jenkins, 1994, "Haircuts or Hysteresis? Sources of Movements in Real Exchange Rates," Journal of International Economics, forthcoming.

Samuelson, Paul, 1954, "The Transfer Problem and the Transport Costs, II: Analysis of Effects of Trade Impediments," Economic Journal 64, 264-289.

Sanyal, Kalyan and Ronald Jones, 1982, "The Theory of Trade in Middle Products," American Economic Review 72, 16-31.

White, Halbert, 1980, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," Econometrica 48, 817-838.

TABLE 1

CATEGORIES OF GOODS IN DISAGGREGATED CONSUMER PRICE INDEXES

# of good	United States	Canada
1	Food at home	Food purchased from stores
2	Food away from home	Food purchased -- restaurants
3	Alcoholic beverages	Alcoholic beverages
4	Shelter	Shelter - .2135*(water, fuel and electricity)
5	Fuel and other utilities	Water, fuel and electricity
6	Household furnishings and operations	Housing excluding shelter
7	Men's and boys' apparel	.8058*(Men's wear)+.1942*(boys' wear)
8	Women's and girls' apparel	.8355*(Women's wear) + .1645*(girls' wear)
9	Footwear	Footwear
10	Private transportation	Private transportation
11	Public transportation	Public transportation
12	Medical care	Health care
13	Personal care	Personal care
14	Entertainment	.8567*(Recreation) + .1433*(reading material)

Table 2: Selected Summary Statistics

Vol. Measure / Country Pairs:	U.S.-U.S.	CAN.-CAN.	U.S.-CAN.
Std. Dev. (FIL)	.027	.014	.031
Spread (FIL)	.063	.031	.073
Std. Dev. (LEV)	.082	.039	.100
Spread (LEV)	.213	.099	.261
Std. Dev. (DIF)	.034	.017	.037
Spread (DIF)	.078	.036	.087
Avg. Distance Between Cities	1,024 mi. (66 pairs)	1,124 mi. (36 pairs)	1,346 mi. (126 pairs)

Notes: Column entries give the mean value of price volatility across all inter-city combinations within the U.S., within Canada, and across the U.S.-Canadian border respectively. Std. Dev. indicates that the measure of volatility is the standard deviation of the relative price series. Spread indicates that volatility is measured as the difference between the 90th and 10th percentile values of the relative price series. FIL indicates that prices are measured as the forecast error from the filter, while LEV (DIF) indicates that prices are measured as the log of relative prices (first-difference of the log of relative prices). The average "great circle" distance between cities is given in the final row.

Table 3: Regressions Relating Price Volatility to Distance and the Border
(Prices measured as forecast error from filter)

Indep. Variable	Regression 1	Regression 2	Regression 3	Regression 4
Distance	5.75×10^{-6} (4.51)	1.31×10^{-5} (4.55)	4.56×10^{-6} (5.56)	1.01×10^{-5} (5.25)
Distance^2	-2.10×10^{-9} (-4.47)	-4.66×10^{-9} (-4.38)	-1.62×10^{-9} (-5.60)	-3.60×10^{-9} (-5.26)
Border Dummy	4.38×10^{-3} (3.44)	1.17×10^{-2} (3.90)	4.21×10^{-3} (13.3)	1.01×10^{-2} (13.4)
Constant	7.80×10^{-3} (7.12)	1.40×10^{-2} (5.68)	2.67×10^{-3} (3.27)	3.62×10^{-3} (1.96)
Good 2 Dummy	-4.14×10^{-3} (-4.51)	-8.91×10^{-3} (-4.29)	-2.14×10^{-3} (-4.27)	-6.27×10^{-3} (-5.52)
Good 3 Dummy	-9.82×10^{-6} (-0.01)	1.67×10^{-3} (0.79)	2.48×10^{-3} (4.87)	6.15×10^{-3} (5.28)
Good 4 Dummy	-3.51×10^{-3} (-4.07)	-5.85×10^{-3} (-3.05)	1.12×10^{-3} (2.39)	3.05×10^{-3} (2.92)
Good 5 Dummy	1.24×10^{-2} (10.6)	3.28×10^{-2} (12.5)	1.74×10^{-2} (27.6)	4.27×10^{-2} (29.4)
Good 6 Dummy	-3.92×10^{-3} (-4.49)	-4.48×10^{-3} (-2.29)	3.68×10^{-4} (0.81)	3.10×10^{-3} (2.98)
Good 7 Dummy	6.13×10^{-3} (6.03)	1.66×10^{-2} (7.33)	1.67×10^{-2} (27.6)	4.12×10^{-2} (28.4)
Good 8 Dummy	1.81×10^{-2} (9.22)	4.40×10^{-2} (9.58)	3.60×10^{-2} (29.8)	8.62×10^{-2} (30.1)
Good 9 Dummy	1.16×10^{-2} (8.35)	3.06×10^{-2} (10.1)	2.37×10^{-2} (27.3)	5.80×10^{-2} (29.1)
Good 10 Dummy	1.33×10^{-3} (1.28)	6.97×10^{-3} (2.92)	9.38×10^{-4} (1.52)	3.79×10^{-3} (2.62)
Good 11 Dummy	1.61×10^{-2} (11.9)	3.76×10^{-2} (11.7)	2.67×10^{-2} (35.1)	5.94×10^{-2} (32.9)
Good 12 Dummy	1.44×10^{-3} (0.93)	-4.31×10^{-3} (-1.89)	1.14×10^{-3} (1.49)	-2.13×10^{-3} (-1.77)
Good 13 Dummy	-4.94×10^{-4} (-0.55)	1.98×10^{-3} (0.93)	4.54×10^{-3} (9.69)	1.15×10^{-2} (10.7)

Good 14 Dummy	-4.68×10^{-3} (-5.49)	-7.18×10^{-3} (-3.75)	-5.67×10^{-5} (-0.12)	1.49×10^{-3} (1.49)
Dummy for Even Month Cities	–	–	1.35×10^{-3} (4.35)	1.95×10^{-3} (2.69)
Dummy for Odd Month Cities	–	–	2.78×10^{-3} (8.16)	5.44×10^{-3} (6.94)
Dummy for U.S. Cities	1.08×10^{-2} (8.41)	2.83×10^{-2} (9.39)	1.09×10^{-2} (17.0)	2.96×10^{-2} (20.3)
Implied Width of the Border	7,412 mi.	6,953 mi.	7,182 mi.	8,103 mi.
Adjusted R^2	.67	.73	.74	.76

Notes:

Regression 1 uses the core cities. The dependent variable is the standard deviation of prices.

Regression 2 uses the core cities. The dependent variable is the spread between the 90th and 10th percentile of prices.

Regression 3 uses all cities. The dependent variable is the standard deviation of prices.

Regression 4 uses all cities. The dependent variable is the spread between the 90th and 10th percentile of prices.

The implied width of the border is calculated as:
(the coefficient on the border dummy) / (the coefficient on distance - (1230x2) times the coefficient on distance squared), where 1230 miles is the mean distance between cities in our sample.

T-statistics in parenthesis are calculated using White's (1980) correction for heteroscedasticity.

Table 4: Regressions Relating Price Volatility to Distance and the Border
(Prices measured as the log of the relative prices)

Indep. Variable	Regression 1	Regression 2	Regression 3	Regression 4
Distance	1.22×10^{-5} (2.50)	2.84×10^{-5} (2.23)	1.08×10^{-5} (3.39)	2.80×10^{-5} (3.29)
Distance^2	-4.10×10^{-9} (-2.25)	-9.62×10^{-9} (-2.01)	-3.85×10^{-9} (-3.32)	-1.01×10^{-8} (-3.26)
Border Dummy	4.12×10^{-2} (11.3)	0.11 (11.5)	1.78×10^{-2} (12.6)	4.78×10^{-2} (12.6)
Constant	6.52×10^{-3} (1.67)	1.32×10^{-2} (1.29)	-7.28×10^{-2} (-2.78)	-2.36×10^{-2} (-3.42)
Good 2 Dummy	1.88×10^{-2} (5.59)	5.74×10^{-2} (6.56)	2.76×10^{-2} (14.5)	8.14×10^{-2} (16.1)
Good 3 Dummy	3.64×10^{-2} (8.99)	0.11 (9.88)	6.29×10^{-2} (22.7)	0.18 (23.5)
Good 4 Dummy	3.48×10^{-2} (7.48)	9.16×10^{-2} (7.55)	4.35×10^{-2} (17.8)	0.11 (17.4)
Good 5 Dummy	5.38×10^{-2} (12.1)	0.15 (12.0)	6.88×10^{-2} (25.0)	0.18 (24.3)
Good 6 Dummy	7.15×10^{-3} (2.31)	2.38×10^{-2} (2.91)	2.53×10^{-2} (12.9)	6.95×10^{-2} (13.7)
Good 7 Dummy	2.38×10^{-2} (6.09)	6.14×10^{-2} (6.12)	3.85×10^{-2} (18.9)	9.96×10^{-2} (18.3)
Good 8 Dummy	6.42×10^{-2} (8.17)	0.17 (8.22)	7.71×10^{-2} (18.5)	0.20 (18.0)
Good 9 Dummy	2.93×10^{-2} (6.97)	8.08×10^{-2} (7.22)	5.12×10^{-2} (18.3)	0.14 (17.9)
Good 10 Dummy	1.53×10^{-2} (4.53)	4.85×10^{-2} (5.39)	2.64×10^{-2} (12.9)	7.68×10^{-2} (14.2)
Good 11 Dummy	3.44×10^{-2} (8.84)	8.99×10^{-2} (9.12)	6.92×10^{-2} (25.6)	0.18 (26.4)
Good 12 Dummy	2.87×10^{-2} (5.55)	7.61×10^{-2} (5.91)	3.25×10^{-2} (12.7)	8.82×10^{-2} (13.2)
Good 13 Dummy	1.19×10^{-2} (3.49)	3.70×10^{-2} (4.04)	2.89×10^{-2} (13.7)	7.96×10^{-2} (13.9)

Good 14 Dummy	-1.57×10^{-3} (-0.48)	1.31×10^{-3} (0.15)	1.40×10^{-2} (8.02)	4.20×10^{-2} (9.05)
Dummy for Even Month Cities	--	--	1.20×10^{-2} (7.08)	3.03×10^{-2} (6.72)
Dummy for Odd Month Cities	--	--	-2.12×10^{-2} (-1.25)	-6.06×10^{-3} (-1.34)
Dummy for U.S. Cities	1.93×10^{-2} (5.50)	5.04×10^{-2} (5.46)	3.92×10^{-2} (17.8)	0.10 (17.9)
Implied Width of the Border	19,570 mi.	23,383 mi.	13,185 mi.	14,959 mi.
Adjusted R^2	.59	.59	.49	.49

Notes:

Regression 1 uses the core cities. The dependent variable is the standard deviation of prices.

Regression 2 uses the core cities. The dependent variable is the spread between the 90th and 10th percentile of prices.

Regression 3 uses all cities. The dependent variable is the standard deviation of prices.

Regression 4 uses all cities. The dependent variable is the spread between the 90th and 10th percentile of prices.

The implied width of the border is calculated as:
(the coefficient on the border dummy) / (the coefficient on distance - (1230x2) times the coefficient on distance squared), where 1230 miles is the mean distance between cities in our sample.

T-statistics in parenthesis are calculated using White's (1980) correction for heteroscedasticity.

Table 5: Regressions Relating Price Volatility to Distance and the Border
(Prices measured as the first-difference of the log of the relative prices)

Indep. Variable	Regression 1	Regression 2	Regression 3	Regression 4
Distance	8.06×10^{-6} (4.89)	1.89×10^{-5} (4.80)	5.16×10^{-6} (4.98)	1.30×10^{-5} (5.50)
$Distance^2$	-2.96×10^{-9} (-4.87)	-6.78×10^{-9} (-4.63)	-1.96×10^{-9} (-5.35)	-4.88×10^{-9} (-5.79)
Border Dummy	5.09×10^{-3} (3.44)	1.59×10^{-2} (4.31)	3.16×10^{-3} (7.83)	8.30×10^{-3} (9.00)
Constant	7.72×10^{-3} (5.83)	1.32×10^{-2} (3.97)	3.53×10^{-3} (3.65)	3.57×10^{-3} (1.62)
Good 2 Dummy	-5.88×10^{-3} (-5.39)	-1.25×10^{-2} (-5.06)	-3.25×10^{-3} (-5.81)	-7.92×10^{-3} (-6.32)
Good 3 Dummy	-1.13×10^{-3} (-1.02)	2.04×10^{-4} (0.08)	1.75×10^{-3} (3.03)	4.60×10^{-3} (3.48)
Good 4 Dummy	-4.15×10^{-3} (-4.16)	-7.39×10^{-3} (-3.23)	2.92×10^{-4} (0.44)	1.38×10^{-3} (1.19)
Good 5 Dummy	2.01×10^{-2} (11.4)	4.81×10^{-2} (13.1)	2.46×10^{-2} (24.4)	5.95×10^{-2} (26.8)
Good 6 Dummy	-5.28×10^{-3} (-5.09)	-7.10×10^{-3} (-2.99)	-1.26×10^{-3} (-2.38)	3.36×10^{-4} (0.28)
Good 7 Dummy	9.98×10^{-3} (8.20)	2.81×10^{-2} (9.92)	1.96×10^{-2} (28.0)	4.75×10^{-2} (29.1)
Good 8 Dummy	2.63×10^{-2} (10.1)	6.75×10^{-2} (10.2)	4.72×10^{-2} (29.4)	0.11 (30.9)
Good 9 Dummy	1.40×10^{-2} (8.93)	3.61×10^{-2} (10.6)	2.49×10^{-2} (25.2)	6.09×10^{-2} (27.1)
Good 10 Dummy	8.10×10^{-4} (0.64)	4.66×10^{-3} (1.63)	1.06×10^{-3} (1.53)	4.83×10^{-3} (3.11)
Good 11 Dummy	2.24×10^{-2} (13.1)	5.61×10^{-2} (12.6)	3.31×10^{-2} (35.9)	7.87×10^{-2} (33.6)
Good 12 Dummy	9.50×10^{-4} (0.58)	-8.16×10^{-3} (-3.27)	1.71×10^{-3} (2.09)	-2.75×10^{-3} (-2.15)
Good 13 Dummy	-1.68×10^{-3} (-1.55)	1.01×10^{-4} (0.04)	2.69×10^{-3} (4.96)	7.43×10^{-3} (5.89)

Good 14 Dummy	-5.64×10^{-3} (-5.49)	-9.94×10^{-3} (-4.27)	-1.74×10^{-3} (-3.31)	-2.14×10^{-3} (-1.87)
Dummy for Even Month Cities	--	--	2.04×10^{-3} (4.83)	4.24×10^{-3} (4.23)
Dummy for Odd Month Cities	--	--	2.63×10^{-3} (5.48)	3.66×10^{-3} (3.36)
Dummy for U.S. Cities	1.34×10^{-2} (8.99)	3.32×10^{-2} (10.1)	1.50×10^{-2} (18.9)	3.95×10^{-2} (21.8)
Implied Width of the Border	6,481 mi.	7,062 mi.	9,136 mi.	8,030 mi.
Adjusted R^2	.68	.71	.74	.77

Notes:

Regression 1 uses the core cities. The dependent variable is the standard deviation of prices.

Regression 2 uses the core cities. The dependent variable is the spread between the 90th and 10th percentile of prices.

Regression 3 uses all cities. The dependent variable is the standard deviation of prices.

Regression 4 uses all cities. The dependent variable is the spread between the 90th and 10th percentile of prices.

The implied width of the border is calculated as:
(the coefficient on the border dummy) / (the coefficient on distance - (1230x2) times the coefficient on distance squared), where 1230 miles is the mean distance between cities in our sample.

T-statistics in parenthesis are calculated using White's (1980) correction for heteroscedasticity.

Table 6: Regressions Relating Price Volatility to Distance and the Border
(Prices are measured as the forecast error from filter; log specification for distance)

Indep. Variable	Regression 1	Regression 2	Regression 3	Regression 4
Log(Distance)	6.11×10^{-4} (3.12)	1.31×10^{-3} (2.81)	8.43×10^{-4} (4.30)	1.84×10^{-3} (3.98)
Border Dummy	4.31×10^{-3} (13.6)	1.03×10^{-2} (13.7)	2.95×10^{-3} (7.96)	7.16×10^{-3} (8.08)
Constant	8.52×10^{-4} (0.57)	-1.40×10^{-4} (-0.04)	-7.10×10^{-4} (-0.48)	-3.74×10^{-3} (-1.07)
Good 2 Dummy	-2.14×10^{-3} (-4.29)	-6.26×10^{-3} (-5.55)	-2.14×10^{-3} (-4.25)	-6.26×10^{-3} (-5.49)
Good 3 Dummy	2.48×10^{-3} (4.86)	6.15×10^{-3} (5.27)	2.48×10^{-3} (4.80)	6.15×10^{-3} (5.22)
Good 4 Dummy	1.12×10^{-3} (2.40)	3.05×10^{-3} (2.92)	1.12×10^{-3} (2.39)	3.05×10^{-3} (2.92)
Good 5 Dummy	1.74×10^{-2} (27.3)	4.27×10^{-2} (29.0)	1.74×10^{-2} (27.9)	4.27×10^{-2} (29.7)
Good 6 Dummy	3.68×10^{-4} (0.81)	3.10×10^{-3} (2.99)	3.68×10^{-4} (0.80)	3.10×10^{-3} (2.96)
Good 7 Dummy	1.67×10^{-2} (27.5)	4.12×10^{-2} (28.2)	1.67×10^{-2} (28.0)	4.12×10^{-2} (28.8)
Good 8 Dummy	3.60×10^{-2} (29.9)	8.62×10^{-2} (30.1)	3.60×10^{-2} (30.1)	8.62×10^{-2} (30.4)
Good 9 Dummy	2.37×10^{-2} (27.2)	5.80×10^{-2} (28.9)	2.37×10^{-2} (27.9)	5.80×10^{-2} (29.6)
Good 10 Dummy	9.38×10^{-4} (1.51)	3.79×10^{-3} (2.62)	9.38×10^{-4} (1.50)	3.79×10^{-3} (2.59)
Good 11 Dummy	2.67×10^{-2} (35.0)	5.94×10^{-2} (32.9)	2.67×10^{-2} (34.9)	5.94×10^{-2} (32.7)
Good 12 Dummy	1.14×10^{-3} (1.48)	-2.13×10^{-3} (-1.77)	1.14×10^{-3} (1.47)	-2.13×10^{-3} (-1.76)
Good 13 Dummy	4.54×10^{-3} (9.69)	1.15×10^{-2} (10.7)	4.54×10^{-3} (9.62)	1.15×10^{-2} (10.6)
Good 14 Dummy	-5.67×10^{-5} (-0.13)	1.49×10^{-3} (1.50)	-5.67×10^{-5} (-0.13)	1.49×10^{-3} (1.49)

Dummy for Even Month Cities	1.53×10^{-3} (4.90)	2.34×10^{-3} (3.24)	-4.50×10^{-4} (-1.27)	-2.22×10^{-3} (-2.71)
Dummy for Odd Month Cities	2.80×10^{-3} (8.17)	5.49×10^{-3} (6.95)	8.75×10^{-4} (2.35)	1.06×10^{-3} (1.25)
Dummy for U.S. Cities	1.07×10^{-2} (16.7)	2.92×10^{-2} (20.0)	1.39×10^{-2} (18.7)	3.66×10^{-2} (21.4)
Dummy for all Pairs Involving NY or LA	—	—	-3.78×10^{-3} (-9.82)	-8.72×10^{-3} (-9.66)
Implied Width of the Border	8,666 mi.	9,663 mi.	4,310 mi.	4,779 mi.
Adjusted R^2	.74	.76	.75	.77

Notes:

Regression 1 uses all cities. The dependent variable is the standard deviation of prices.

Regression 2 uses all cities. The dependent variable is the spread between the 90th and 10th percentile of prices.

Regression 3 uses all cities. The dependent variable is the standard deviation of prices.

Regression 4 uses all cities. The dependent variable is the spread between the 90th and 10th percentile of prices.

The implied width of the border is calculated as:

1230*(the coefficient on the border dummy) / (the coefficient on log-distance), where 1230 miles is the mean distance between cities in our sample.

T-statistics in parenthesis are calculated using White's (1980) correction for heteroscedasticity.

Table 7a: Regressions Relating Price Volatility to Distance and the Border
Separate Regressions for Each Good
(Prices measured as forecast error from filter)

Gd.	Distance	$Distance^2$	Border Dummy	Constant	U.S. Cities Dummy	Even Cities Dummy	Odd Cities Dummy	Adj R^2
1	3.00×10^{-6} (3.25)	-8.77×10^{-10} (-2.69)	7.77×10^{-3} (19.5)	1.47×10^{-2} (20.3)	-2.47×10^{-3} (-4.40)	-1.03×10^{-3} (-2.16)	-6.22×10^{-4} (-1.26)	.68
2	2.24×10^{-6} (2.54)	-6.37×10^{-10} (-1.86)	8.70×10^{-3} (22.8)	9.16×10^{-3} (18.8)	-5.00×10^{-4} (-1.48)	3.29×10^{-4} (1.18)	2.14×10^{-3} (4.88)	.81
3	5.91×10^{-6} (5.38)	-1.90×10^{-9} (-4.59)	8.67×10^{-3} (20.3)	1.11×10^{-2} (13.9)	-1.14×10^{-4} (-0.23)	1.38×10^{-3} (3.43)	2.08×10^{-3} (3.78)	.75
4	1.25×10^{-6} (0.88)	-3.90×10^{-10} (-0.76)	6.34×10^{-3} (10.3)	1.03×10^{-2} (9.67)	1.29×10^{-3} (1.67)	6.02×10^{-3} (8.56)	3.44×10^{-3} (4.61)	.59
5	1.77×10^{-5} (7.49)	-5.66×10^{-9} (-6.64)	3.63×10^{-3} (3.43)	1.68×10^{-2} (9.19)	3.12×10^{-3} (2.66)	4.99×10^{-3} (4.10)	4.08×10^{-3} (3.26)	.40
6	8.96×10^{-7} (0.85)	-2.81×10^{-10} (-0.76)	6.19×10^{-3} (15.1)	1.03×10^{-2} (13.8)	1.15×10^{-3} (2.33)	4.07×10^{-3} (8.82)	4.52×10^{-3} (8.51)	.71
7	4.36×10^{-6} (1.64)	-1.90×10^{-9} (-2.05)	3.42×10^{-3} (3.22)	1.80×10^{-2} (9.17)	4.71×10^{-3} (3.29)	1.05×10^{-2} (8.05)	1.49×10^{-2} (11.2)	.58
8	-1.02×10^{-5} (-1.65)	3.75×10^{-9} (1.71)	-1.42×10^{-3} (-0.59)	3.01×10^{-2} (7.17)	1.64×10^{-2} (6.15)	1.64×10^{-3} (6.08)	2.28×10^{-2} (7.14)	.49
9	3.73×10^{-6} (0.89)	-2.14×10^{-9} (-1.53)	-9.66×10^{-4} (-0.56)	2.37×10^{-2} (7.35)	8.91×10^{-3} (4.01)	1.16×10^{-2} (5.40)	1.60×10^{-2} (7.31)	.41
10	5.51×10^{-6} (6.46)	-1.91×10^{-9} (-6.45)	1.26×10^{-2} (35.8)	1.21×10^{-2} (18.3)	-1.37×10^{-3} (-2.87)	-2.04×10^{-3} (-4.68)	-3.26×10^{-3} (-7.63)	.89
11	1.07×10^{-5} (3.90)	-3.07×10^{-9} (-3.03)	1.46×10^{-2} (13.0)	1.52×10^{-2} (8.55)	8.10×10^{-3} (5.53)	7.18×10^{-3} (5.35)	1.14×10^{-2} (9.31)	.73
12	8.65×10^{-6} (3.30)	-3.26×10^{-9} (-3.59)	9.26×10^{-3} (8.91)	1.29×10^{-2} (5.24)	-1.81×10^{-3} (-1.04)	-2.89×10^{-3} (-2.19)	-8.41×10^{-4} (-0.63)	.27
13	1.84×10^{-6} (1.52)	-7.29×10^{-10} (-1.68)	4.36×10^{-3} (8.51)	1.47×10^{-2} (18.5)	5.91×10^{-4} (1.10)	5.36×10^{-3} (9.61)	6.58×10^{-3} (10.5)	.62
14	-1.56×10^{-6} (-1.34)	6.67×10^{-10} (1.62)	6.24×10^{-3} (13.2)	1.04×10^{-3} (10.9)	1.36×10^{-3} (2.21)	5.69×10^{-3} (10.3)	3.87×10^{-3} (6.76)	.69

Notes: The regressions above use all cities. The dependent variable is the standard deviation of prices. T-statistics in parenthesis are calculated using White's (1980) correction for heteroscedasticity.

Table 7b: Regressions Relating Price Volatility to Distance and the Border
Separate Regressions for Each Good
(Prices measured as forecast error from filter; log specification used for distance)

Good	Log (Distance)	Border Dummy	Constant	U.S. Cities Dummy	Even Cities Dummy	Odd Cities Dummy	Adj. R^2
1	7.45×10^{-5} (3.17)	7.78×10^{-3} (19.2)	1.15×10^{-2} (6.90)	-2.48×10^{-3} (-4.38)	-9.78×10^{-4} (-2.02)	-6.39×10^{-4} (-1.29)	.67
2	6.58×10^{-4} (4.00)	8.68×10^{-3} (22.6)	6.11×10^{-3} (5.39)	-4.83×10^{-4} (-1.45)	3.45×10^{-4} (1.23)	2.13×10^{-3} (4.87)	.81
3	1.29×10^{-3} (5.61)	8.64×10^{-3} (19.5)	5.66×10^{-3} (3.20)	-9.14×10^{-5} (-0.18)	1.45×10^{-3} (3.84)	2.03×10^{-3} (3.60)	.74
4	4.38×10^{-4} (1.41)	6.26×10^{-3} (10.1)	8.04×10^{-3} (3.49)	1.34×10^{-3} (1.75)	6.00×10^{-3} (8.65)	3.43×10^{-3} (4.59)	.59
5	3.81×10^{-3} (7.00)	3.58×10^{-3} (3.24)	9.06×10^{-4} (0.23)	3.16×10^{-3} (2.75)	5.24×10^{-3} (4.12)	3.94×10^{-3} (3.07)	.37
6	1.17×10^{-4} (0.41)	6.23×10^{-3} (15.2)	1.00×10^{-2} (4.97)	1.12×10^{-3} (2.24)	4.11×10^{-3} (8.86)	4.52×10^{-3} (8.53)	.72
7	-1.15×10^{-4} (-0.17)	3.54×10^{-3} (3.28)	2.03×10^{-2} (4.25)	4.65×10^{-3} (3.23)	1.06×10^{-2} (8.21)	1.48×10^{-2} (11.0)	.58
8	-1.23×10^{-3} (-0.82)	-1.50×10^{-3} (-0.62)	3.37×10^{-2} (3.15)	1.64×10^{-2} (6.06)	1.61×10^{-2} (5.95)	2.29×10^{-2} (7.16)	.48
9	-1.00×10^{-3} (-0.98)	-8.25×10^{-4} (-0.47)	3.08×10^{-2} (4.11)	8.86×10^{-3} (4.06)	1.18×10^{-2} (5.47)	1.59×10^{-2} (7.17)	.40
10	8.97×10^{-4} (4.43)	1.26×10^{-2} (34.6)	8.81×10^{-3} (5.82)	-1.38×10^{-3} (-2.67)	-1.93×10^{-3} (-4.17)	-3.31×10^{-3} (-7.43)	.88
11	2.76×10^{-3} (4.65)	1.47×10^{-2} (12.9)	3.14×10^{-3} (0.77)	8.06×10^{-3} (5.40)	7.35×10^{-3} (5.51)	1.14×10^{-2} (9.19)	.73
12	6.66×10^{-4} (1.31)	9.46×10^{-3} (8.76)	1.24×10^{-2} (2.83)	-1.93×10^{-3} (-1.09)	-2.62×10^{-3} (-1.92)	-9.33×10^{-4} (-0.68)	.25
13	5.73×10^{-5} (0.18)	4.41×10^{-3} (8.56)	1.51×10^{-2} (7.00)	5.57×10^{-4} (1.04)	5.42×10^{-3} (9.82)	6.56×10^{-3} (10.4)	.61
14	1.76×10^{-5} (0.06)	6.20×10^{-3} (12.9)	9.72×10^{-3} (4.40)	1.38×10^{-3} (2.22)	5.64×10^{-3} (10.1)	3.89×10^{-3} (6.78)	.69

Notes: The regressions above use all cities. The dependent variable is the standard deviation of prices. T-statistics in parenthesis are calculated using White's (1980) correction for heteroscedasticity.

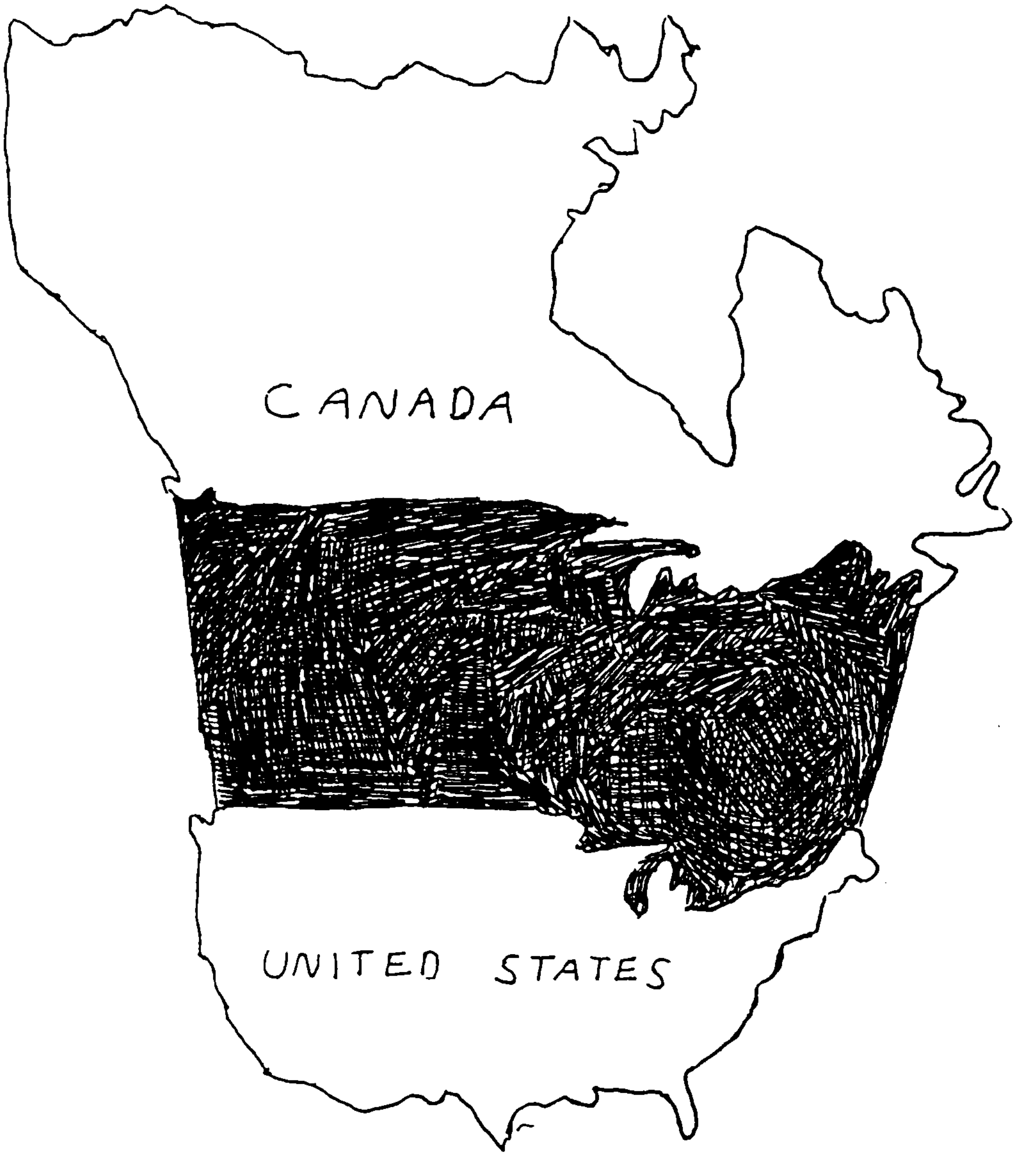

FIGURE 1