Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 777

September 2003

REVISITING THE BORDER: AN ASSESSMENT OF THE LAW OF ONE PRICE USING VERY DISAGGREGATED CONSUMER PRICE DATA

Charles Engel, John H. Rogers and Shing-Yi Wang

Revisiting the Border: An Assessment of the Law of One Price Using Very Disaggregated Consumer Price Data

Charles Engel
University of Wisconsin

John H. Rogers
Board of Governors of the Federal Reserve System

Shing-Yi Wang
Yale University

June 23, 2003

Abstract

We reexamine the evidence for "border effects" in deviations from the law of one price, using data for consumer prices from Canadian and U.S. cities. The study parallels Engel and Rogers (1996), except that this study uses actual price data rather than price index data. We find evidence of border effects both in the levels of prices and the percentage change in prices. Even accounting for distance between cities and relative population sizes, we find that the absolute difference between prices in the U.S. and Canada in our data (annual from 1990 to 2002) is greater than 7 percent. This difference exists among tradables and nontradables, though for some categories of tradables (clothing and durables) the difference is smaller. The findings are similar for annual changes, though the magnitude is smaller: The border accounts for a difference in 1.5 percent in annual (log) price changes. Relative population sizes and distance are helpful in explaining price level differences (between Canadian and U.S. cities) for traded goods but are less helpful in explaining price level differences for nontraded goods or for accounting for differences in price changes for either traded or nontraded goods.

Keywords: exchange rates, prices
JEL classification: F3, F4

Addresses: Charles Engel, Department of Economics, University of Wisconsin, Madison, WI, 53711. e-mail: cengel@ssc.wisc.edu.

John H. Rogers, Division of International Finance, Federal Reserve Board, Washington, DC, 20551. e-mail: John.H.Rogers@frb.gov

Shing-Yi Wang, Department of Economics, 28 Hillhouse Ave., Yale University, New Haven, CT, 06511 e-mail: shing-yi.wang@yale.edu

The views expressed in this paper are those of the authors and do not necessarily represent the views of the Board of Governors of the Federal Reserve or the Federal Reserve System. Engel acknowledges research support from a grant from the National Science Foundation to the University of Wisconsin.

0. Introduction

Prices that households pay for consumer goods should not be very different across a pair of markets if those markets are well integrated. "Integration" means that barriers to commerce of all sorts – formal trade barriers, transportation costs, exclusivity of distribution networks, etc. – are low. Engel and Rogers (1996) (hereinafter referred to as ER) examined prices across a number of North American cities in an attempt to assess the integration of Canadian and American markets for goods. Their finding was that the markets were not as well integrated as one might have expected. Cities within each country showed much greater harmony in prices even if they were very distant markets compared to pairs of cities that lie across the U.S./Canada border, even if the cities were nearby geographically. There was, in the words of that study, a large "border" effect.

The literature suggests two ways in which this imperfect synchronization of prices might influence exchange-rate and monetary policy. On the one hand, following Mundell (1961), two countries that are highly integrated commercially are apt to be strong candidates for a common currency. One of the most powerful gains from a common currency is from lowering transactions costs for cross-border trade. Money eases trade, so a common money would ease trade across borders.[1] The more transactions that occur between economies, so the more integrated the goods markets, the greater the gains from a common currency.

On the other hand, short run deviations from the law of one price across national borders might reflect nominal exchange rate misalignment. That is, in each country nominal goods prices might be set in the local currency. Nominal exchange rates reflect not only current market conditions but also expectations of the future. As the nominal exchange rate fluctuates but goods prices adjust only slowly, there arise deviations of prices (expressed in a common currency)

[1] This simple idea finds empirical backing in the work of Rose (2000), who finds that adoption of a common currency greatly expands the volume of trade between nations.

across borders. That is, let $P_i^{US\$}$ be the U.S. dollar price of good *i* sold in the U.S., and $P_i^{CA\$}$ the Canadian dollar price of the same good sold in Canada. Both of these prices might adjust sluggishly to changes in demand or supply. As $S_{US\$/C\$}$, the U.S. dollar per Canadian dollar exchange rate, fluctuates as the market learns news of future economic conditions, there will be deviations from the law of one price condition, $P_i^{US\$} = S_{US\$/C\$} P_i^{CA\$}$. Devereux and Engel (2000) have argued that under these circumstances, there are gains to stabilizing nominal exchange rates. When there is local-currency pricing, changes in the nominal exchange rate do not change relative prices faced by consumers. Prices of foreign-produced and domestically-produced goods are both sticky in the local currency. There is no "expenditure switching" effect of exchange rate changes, so a flexible exchange rate does not help facilitate goods market adjustment. On the contrary, because short-run fluctuations in the nominal exchange rate induce price wedges between countries, they lead to inefficient allocation of resources. Exchange-rate stability can minimize these distorting deviations from the law of one price.

Thus if deviations from the law of one price are short-run, there may be a case for fixing nominal exchange rates, perhaps in the ultimate form of a common currency. On the other hand, if the deviations from the law of one price are large in the long run, then the markets are not well integrated, and they are poor candidates for a common currency in Mundell's framework.

The tests of ER do not permit the evaluation of long-run deviations from the law of one price. They use price index data. This means that ER cannot compare price levels in U.S. cities to price levels in Canadian cities. They are only able to compare rates of inflation. ER can only measure the extent of short-run deviations from the law of one price. That is, they can compare $\pi_{i,j}^{US\$}$ -- the inflation rate of good *i*, in U.S. dollars, in U.S. city *j* – to $\delta_{US\$/C\$} + \pi_{i,k}^{CA\$}$, where $\delta_{US\$/C\$}$ is the rate of depreciation of the U.S. dollar relative to the Canadian dollar, and $\pi_{i,k}^{CA\$}$ is the Canadian dollar inflation rate of good *i* in some Canadian city. ER use official consumer price data from Statistics Canada and the U.S. Bureau of Labor Statistics, which publish price data only in index form. Moreover, their data is disaggregated by categories of consumer goods, but not highly disaggregated. Their prices are subindexes of fairly broad categories of goods such as food at home, women's and girls' clothing, footwear, transportation, etc.

Here, we make use of data from the Economist Intelligence Unit (EIU) that includes actual prices of 100 consumer goods in 13 U.S. cities and 4 Canadian cities. The cities are listed in Table 1. The data is annual (recorded in December) from 1990 to 2002. The data is collected by EIU as a way to compare costs of living for cities throughout the world. The data is for a wide variety of products. There is heavy concentration on food items – 42 of the 100 goods are food or drink, such as tomatoes, ground beef, or six-year aged Scotch whiskey. There are 9 clothing items, such as women's cardigan sweaters. A half-dozen of the items are consumer durables, including a 2-slice electric toaster and a low priced car (900-1299 cc). Non-tradable services such as men's haircut (including tip) or one hour's babysitting constitute 21 of the items. The remaining 22 prices are for miscellaneous (tradable) products such as insect-killer spray (330 g) and aspirin (100 tablets). So the items are narrowly defined, and the EIU attempts to price comparable products across cities. They report prices according to type of outlet (supermarket, mid-priced store, etc.) Table 1 lists the products and outlets that we use in this study.[2]

Because we can compare actual price levels, we can investigate long-run differences in price levels among North American cities, as well as the behavior of short-run price changes. Our empirical work, therefore, estimates a simple model to explain price level differences between cities: the absolute value of the difference in the price between two cities is modeled as

[2] In a typical year, the EIU reports prices on many more products. We chose to work with these 100 items because there is price data for all of them for each city for every year, thus allowing us to use balanced panels.

a function of the log of distance between the cities, the absolute value of the population difference (since larger cities tend to have higher prices), a measure of the absolute value of the difference in sales taxes between cities, and a dummy variable that indicates whether or not the two cities are in different countries.[3] We use the same set of explanatory variables in a separate set of regressions in which the dependent variable captures the short-run movements in prices, and is thus similar to that used by ER: A typical observation might be $\left|\pi_{i,j}^{US\$} - (\delta_{US\$/C\$} + \pi_{i,k}^{CA\$})\right|$ if cities *j* and *k* are in different countries (where $|x|$ refers to the absolute value of x), or, for example, $\left|\pi_{i,j}^{US\$} - \pi_{i,k}^{US\$}\right|$ if cities *j* and *k* are both located in the U.S.

There are drawbacks both to our measurement of long-run price differences and short-run differences. Our data span only 13 years. If transitory price differences disappear slowly, then our 13-year sample might not be long enough to eliminate the effects of transitory deviations from the law of one price. Specifically, it may be the case that the U.S. dollar was "overvalued" compared to the Canadian dollar during a sizable fraction of our 13-year span, which would induce higher average prices in the U.S. that do not not reflect permanent barriers to integration. However, one might suspect that there must be some significant commercial barriers if transitory price differences can persist for years. On the other side of the coin, one-year changes might be too low frequency to capture the most significant transitory fluctuations in relative prices that emerge from volatile nominal exchange rates. But since the data is only annual, we cannot measure price changes at any higher frequency.

There may be a large degree of measurement error in these prices. The EIU does not publish full details of its methodology, and one suspects that the prices are not as comparable as prices collected by the official agencies. However, the price data is used as the dependent

[3] There are also city dummies, and in some regressions, time dummies, as we explain in Section I.

variable in our regression, so any measurement error should not affect the consistency of our parameter estimates. There may be a lot of "noise" in the inter-city price comparisons for a particular item, which might make it difficult to assess the role of the border for comparisons of prices for a single good. But, we gain power by using panel estimation, assessing the role of the border for the entire collection of 100 goods. In addition, we estimate smaller panels for the different categories of goods described above: food, clothing, miscellaneous products, durables, and services.

Estimation of panels allows us to compare price levels of individual goods across countries, and reduce the problems of low power introduced by measurement error. Official statistical agencies do not make price data on individual goods publicly available, in part because of their concerns about measurement error. Instead they only report indexes, because the variance of the measurement error is reduced when the prices are averaged into an index. But once the data is averaged, we can no longer compare price levels of individual goods across locations.

We find significant evidence of border effects both in the levels of (logs of) prices and the percentage change in prices. Even accounting for distance between cities and relative population sizes, we find that the absolute difference between prices in the U.S. and Canada in our data (annual from 1990 to 2002) is greater than seven percent. This difference exists among tradables and nontradables, though for some categories of tradables (clothing and durables) the difference is smaller. The findings are similar for annual changes, though the magnitude is smaller – the border accounts for a difference in 1.5 percent in annual (log) price changes. Relative population sizes and distance are helpful in explaining price level differences (between Canadian and U.S. cities) for traded goods, but are less helpful in explaining price level differences for nontraded goods or for accounting for differences (between U.S. and Canadian cities) in price changes for either traded or nontraded goods.

What does all of this mean for the desirability of a common currency or fixed exchange rates for Canada and the U.S.? Probably nothing. In the first place, the adoption of a common currency is almost certainly a non-starter politically. Secondly, we have no standard by which to assess the magnitude of this border effect. Is a seven percent average difference in prices small or large? This study is not intended to yield a definitive conclusion, but instead is meant to encourage further study and to provide the starting point for a methodology that can assess the integration of markets. It is our hope that government and central bank researchers will work in cooperation with official statistical agencies to analyze very disaggregated price level data so that we can get a broader picture of the "border" effect among a collection of countries.

I. Estimation Strategy

Our measure of integration of two locations – the dependent variable in our regressions – is the absolute value of the log price difference of good *i* between locations *j* and *k*: $\left|p_{i,j,t} - p_{i,k,t}\right|$, where $p_{i,j,t}$ refers to the log of the price expressed in U.S. dollars of good *i*, in city *j*, at time *t*. Note that we express all prices in U.S. dollars so that we can compare prices across all cities.[4] The price data is annual, measured in December, for 1990-2002. The dimensions of our panel then are: 100 goods; 17 cities, which means 136 city pairs; and, 13 time periods. The panel consists of 176,800 observations. Prices are inclusive of tax.

[4] This means that the Canadian dollar price of goods sold in Canadian cities is converted into U.S. dollar values by multiplying by the U.S. dollar per Canadian dollar exchange rate. The EIU survey reports prices in U.S. dollar terms, converted using "the market exchange rate on the date of the survey".

When we consider changes in prices, the dependent variable is $\left|\pi_{i,j,t} - \pi_{i,k,t}\right|$, where $\pi_{i,j,t} \equiv p_{i,j,t} - p_{i,j,t-1}$. This data then runs from 1991-2002, for a time dimension of 12 periods. This panel has 163,200 observations.

The first explanatory variable in the regression is the log of the distance between locations *j* and *k*, $dist_{jk}$. Distance has proven to be a very useful explanatory variable for the volume of trade between two locations, as in the "gravity model" of trade. ER explain how it might also help explain deviations from the law of one price. The gravity model suggests that since transportation costs increase with distance, trade volumes will be greatest among nearby locations. When we consider the consumer prices of goods in two locations, it is very unlikely that either city is the exporter of the good. For example, we compare the price of olive oil between Washington, DC and Toronto, but neither city is known for its extensive groves of olive trees. Nonetheless, transportation costs might play a role in making prices more similar between nearby cities. The transport costs of olive oil from Greece to two close together cities is probably very similar, while it may be very different for two distant cities.

Distribution costs are a large component of the final consumer price. Distribution costs are more likely to be similar for neighboring locations. Distribution of some goods is very labor intensive, and labor markets may be more tightly integrated if they are nearer geographically.

Also, ER point out that the mark-up on certain products might be more similar for nearby communities, perhaps because of regional determinants of demand.

The second explanatory variable is the absolute value difference in the log of the population between cities *j* and *k*, $pop_{jk,t}$. This variable is included because larger cities tend to have higher prices. For the U.S. the data refer to Metropolitan (MSA) Population Data. For

Canada, the data are described as "Total Population, Census Div/Metro Areas." The population variable is time varying in the panel regressions, with data in each year from 1990-2002.[5]

We also introduce a measure of the difference in sales taxes between two locations as a possible explanatory variable for price differences. It is conceivable that markets are integrated to the extent that pre-tax prices are nearly equal but that differences in local sales taxes drives a wedge between prices in different locations. The tax rates used in the regressions are retail sales tax rates. For Canada, there are both national and provincial components to the rate. For the U.S., there is of course no national sales tax, so we simply use the state sales tax rates.[6] The absolute value of the tax rate difference between cities *j* and *k* is labeled tax_{jk}. This variable is not time varying because we use a single tax rate for each city for the entire period. We were not able to construct a full panel of tax rates, and so averaged the data we were able to compile for each city. There appears to be very little time variation in sales tax rates.

The variable that is meant to capture the degree of integration between U.S. and Canadian markets is $bord_{jk}$. This is a dummy variable that takes on the value of one if cities *j* and *k* lie on opposite sides of the national border between the U.S. and Canada. The coefficient on the border dummy captures the absolute average log price difference between U.S. and Canadian cities that is not explained by distance or city size (or one of the dummy variables described below.)

As in ER, we include dummy variables for each city, $citdum_j$. This variable takes on the value of one if one of the cities in the city pair is city *j*. It is intended to capture any idiosyncratic

[5] The U.S. data is from the Census Bureau, and the Canada data from Statistics Canada. The U.S. data was downloaded from the site: http://recenter.tamu.edu/Data/popm, and the Canadian data from Haver in the Cansimr database (Regional Canadian Economic Indicators).

[6] We do not include any measure of city-specific sales taxes for any U.S. cities. The data on sales taxes are compiled from a variety of sources: U.S. data (on-line): Center on Budget and Policy Priorities; and, Urban Institute State database. Canadian data: Canadian Tax Foundation's Finances of the Nation; Price Waterhouse; http://www.ca.taxnews.com/tnnpublic.nsf/notespages/4652A712797CB4AC85256959006AB77E/$file/FactsFigures2002.pdf ; http://www.bus.ualberta.ca/CIBS-WCER/WCER/pdf/43.pdf; and, http://freespace.virgin.net/john.cletheroe/usa_can/taxes/

aspects of the price of a given city that tends to make it different. We also performed regressions using time dummies, but the introduction of time dummies had little influence on our other parameter estimates. We also felt that there might be problems of interpretation when time dummies are included, so we report only results from regressions with no time dummies.

Thus, when we estimate equations for differences in price levels, our regression takes the form:

$$\left|p_{i,j,t} - p_{i,k,t}\right| = \beta_1 dist_{jk} + \beta_2 pop_{jk,t} + \beta_3 tax_{jk} + \beta_4 bord_{jk} + \sum_{h=1}^{N} \alpha_h citdum_h + u_{i,jk,t} \text{ .(1)}$$

For changes in prices, the equation is similar:

$$\left|\pi_{i,j,t} - \pi_{i,k,t}\right| = \eta_1 dist_{jk} + \eta_2 pop_{jk,t} + \eta_3 tax_{jk} + \eta_4 bord_{jk} + \sum_{h=1}^{N} \lambda_h citdum_h + u_{i,jk,t} \text{ . (2)}$$

As has been noted, we estimate these equations as a panel using all 100 goods. We also estimate using panels that have prices from each of 5 categories of goods: food, clothing, durables, miscellaneous products, and services.

II. Empirical Results

Table 2 reports regression results for equation (1) when the full sample of 100 items is used in the panel. The sales tax variable proved not to be statistically significant in our regressions, so we report results only for those specifications that drop that variable. The three remaining variables – $dist_{jk}$, $pop_{jk,t}$, and $bord_{jk}$ – are highly significant, and the coefficients all have the expected sign. The coefficients on $dist_{jk}$ and $pop_{jk,t}$ have interpretations as elasticities. A 10 percent increase in the distance between two cities *ceteris paribus* increases the absolute price difference between the cities by 3.2 one-hundredths of one percent. Similarly, the effect of a 10 percent increase in relative population between two cities is to increase the absolute value of the price differential by 9.5 one-hundredths of one percent.

The coefficient on the border gives us the absolute average difference in prices in the U.S. versus Canada, holding other explanatory effects constant. We see from Table 2A that the difference is 7.3 percent. Note also that the border effect is very precisely estimated, with a t-statistic over 40. While this magnitude of price difference appears to be large in economic terms, it is difficult to interpret it as a measure of economic integration without having similar statistics for other country pairs for comparison.

To get a sense of the usefulness of panel estimation, we can compare the findings from the panel with our findings when we estimate equation (1) for each item individually. We find that out of 100 individual regressions, the coefficient on distance was significant at the 5 percent level and correctly signed for only 23 items; on relative population for 27 items; and, on the border dummy for 70 goods. (There were 8 items for which the distance variable was significant but incorrectly signed, 5 in which population was significant but with the wrong sign, and zero such cases for border.) At the 10 percent significance level, the number of significant and correctly signed coefficients were: 27 for distance, 30 for population and 72 for the border (with 10 incorrectly signed significant coefficients on distance, 11 on population, but none on border).

The estimated coefficients for equation (2) when all items are included in the panel are reported in Table 2B. As one should expect, all of the coefficients are smaller in magnitude when these short-run changes are examined. While the border dummy and relative population are still statistically significant, distance no longer is. That is, changes in the absolute price differences are not significantly linked to distance, which contrasts with the finding of ER.

The coefficient on the border dummy tells us that, *ceteris paribus*, the influence of the border effect is to increase the absolute value of the difference in price changes in U.S. cities relative to Canadian cities by 1.4 percentage points. Again, more data from other countries are needed before we can assess the economic significance of this finding.

We also estimated regression (2) individually for each of the 100 items, with this outcome: The coefficient on distance was significant at the 5 percent level and of the correct sign for 1 item (5 at the 10 percent level; relative population for 5 (6) items; border for 50 (56) items. For no items was a variable significant but of the incorrect sign.

The panel estimation is restrictive in that it imposes the same coefficients in regressions (1) and (2) for all items. Tables 3-8 report results for regressions estimated on smaller panels.

Tables 3A, 4A, 5A, and 6A report results of estimation of equation (1) on price level differences for four mutually exclusive groupings of items: food, miscellaneous products, clothing, and durables, respectively. The border dummy is the only variable that was significant and of the correct sign in all four regressions. Its magnitude varies across categories. The border effect implies approximately an 8 percent difference in prices of food items, and a 9 percent difference in prices of miscellaneous products; but only a 2 percent difference in the price of clothing and a 3 percent difference in the price of durable items. Population is only significant and of the correct sign for food items. Apparently prices are higher in larger cities only for food. Similarly, distance is only significant and of the correct sign for food items. This suggests that perhaps shipping costs are important in determining price differences, since these costs are apt to be a relatively high fraction of total value for food.

Table 7A reports the results of regression (1) for a panel that only includes prices of services. The border coefficient and relative population are significant and of the correct sign. The border accounts for an 8 percent difference in prices across the U.S./Canadian border, *ceteris paribus*. The fact that both the border coefficient and relative population are significant explainers of price differences for services most likely is accounted for by the differences in labor markets across locations. One can surmise that wages are relatively high in large cities, and in the U.S. compared to Canada, thus making services prices higher.

It is interesting to compare the findings in Table 7A to those in Table 8A. The latter table reports the results of a panel regression in which only goods – food, miscellaneous products, clothing, and durables – are included. First, note that distance is significant in explaining price differences for goods but not services. This seems to indicate that shipping costs are an important reason why distance matters for prices, since the effect is restricted to goods that are traded. (And, recall that this effect mostly arises from food items.)

Second, the border effect is quite similar in magnitude for both services and goods. Perhaps this represents the influence of higher wages in the U.S. compared to Canada. This might push up the cost of all products, including goods (because of the labor input into distribution) in the U.S. Alternatively, it may be that the mark-up is higher in the U.S. as a consequence of price discrimination by sellers. But another possibility that we cannot rule out is that this difference represents a persistent overvaluation of the U.S. dollar relative to the Canadian dollar in the 1990s. That is, perhaps the price wedge is not the result of equilibrium factors, but instead arises because of a very long-lived, but transitory disequilibrium.

We also note that the magnitude of the relative population variable is almost identical for goods and for services, which may lend support for the notion that the local wage is a large determinant of retail prices even for goods.

Tables 3B, 4B, 5B, and 6B report the results of estimation of equation (2) for price changes for food, miscellaneous products, clothing, and durables, respectively. The border dummy is significant and of the correct sign for all of these categories except clothing. As we found with price levels, the border effect on price changes is largest for food items and miscellaneous products, and is especially small for clothing. Distance is never a significant explanatory variable in these regressions, and population is significant only for food items.

Comparing the effects of distance, relative population, and the border on relative price changes between cities for services and goods (from Tables 7B and 8B), we again find little difference. The magnitude of the border coefficient is very similar – it accounts for a 1.7 percentage difference in price changes for services, and 1.4 percentage point difference for goods. The coefficients on distance are small and insignificant in both panels. One slight distinction is that the coefficient on relative population is slightly larger and significant in the goods panel.

III. Caveats and Conclusions

We have found that distance and relative population play a significant role in explaining price level differences between the U.S. and Canada, but only a minor and usually insignificant role in the regressions explaining differences in price changes. The major exception to this is that distance does not play an important part in explaining price level differences for service items. But the border dummy is almost universally significant, both statistically and apparently in economic magnitude. The price differences across borders exist among tradables (i.e., goods) and nontradables (services), though for some categories of tradables (clothing and durables) the difference is smaller. Roughly, the magnitude of the border effect is a 7 percent difference in the absolute prices between Canada and the U.S., and a 1.5 percent difference in price changes.

We do not view the findings of this study as conclusive. The precision of our estimates is limited by the precision of the measurement of prices; the lack of availability of prices for more than four cities in Canada; the number of goods for which we have a full time series from 1990-2002 of prices (only 100 goods); the frequency of observation of prices (annual); and, the time span of the data (only the most recent thirteen years.) It would also be helpful to be able to use data on other explanatory variables for price differences, such as wages in the service sector by city. And, as we have noted, this study only examines price differences for one pair of countries. There is no set of results for other countries to use as a gauge for comparisons. Most of the data for more refined study probably lies in the files of national statistical agencies. There are significant potential benefits to analyzing that data as a way of measuring the economic integration of economies and the significance of short-run fluctuations in exchange rates.

References

Devereux, Michael B., and Charles Engel, 2000, "Monetary Policy in the Open Economy Revisited: Exchange Rate Flexibility and Price-Setting Behavior," National Bureau of Economic Research, working paper no. 7665.

Engel, Charles, and John H. Rogers, 1996, "How Wide is the Border?" American Economic Review 86, 1112-1125.

Mundell, Robert, 1961, "A Theory of Optimum Currency Areas," American Economic Review 51, 657-665.

Table 1

List of Cites, Goods and Type of Retail Outlet

U.S. Cities Atlanta, Boston, Chicago, Cleveland, Detroit, Houston, Los Angeles, Miami, New York, Pittsburgh, San Francisco, Seattle, Washington DC

Canadian Cites Calgary, Montreal, Toronto, Vancouver

Goods (Type of Retail Outlet) ("average" refers to the average of mid-priced and discount outlets, as reported by EIU.)

White bread, 1 kg (supermarket)	Drinking chocolate (500 g) (supermarket)	'Women's dress, ready to wear, daytime (chain store)
Butter, 500 g (supermarket)	Coca-Cola (1 l) (supermarket)	Women's shoes, town (chain store)
Margarine, 500g (supermarket)	Tonic water (200 ml) (supermarket)	Women's cardigan sweater (chain store)
White rice, 1 kg (supermarket)	Mineral water (1 l) (supermarket)	Women's tights, panty hose (chain store)
Spaghetti (1 kg) (supermarket)	Wine, common table (750 ml) (supermarket)	Child's shoes, sportswear (chain store)
Flour, white (1 kg) (supermarket)	Beer, local brand (1 l) (supermarket)	Girl's dress (chain store)
Sugar, white (1 kg) (supermarket)	Beer, top quality (330 ml) (supermarket)	Hourly rate for domestic cleaning help (average)
Cheese, imported (500 g) (supermarket)	Scotch whisky, six years old (700 ml) (supermarket)	Babysitter's rate per hour (average)
Cornflakes (375 g) (supermarket)	Soap (100 g) (supermarket)	Compact disc album (average)
Milk, pasteurised (1 l) (supermarket)	Laundry detergent (3 l) (supermarket)	Television, colour (66 cm) (average)
Olive oil (1 l) (supermarket)	Toilet tissue (two rolls) (supermarket)	Kodak colour film (36 exposures) (average)
Peanut or corn oil (1 l) (supermarket)	Dishwashing liquid (750 ml) (supermarket)	Cost of developing 36 colour pictures (average)
Potatoes (2 kg) (supermarket)	Insect-killer spray (330 g) (supermarket)	Daily local newspaper (average)
Tomatoes (1 kg) (supermarket)	Batteries (two, size D/LR20) (supermarket)	Paperback novel (at bookstore) (average)
Oranges (1 kg) (supermarket)	Frying pan (Teflon or good equivalent) (supermarket)	Three-course dinner at top restaurant for four people (average)
Apples (1 kg) (supermarket)	Electric toaster (for two slices) (supermarket)	Four best seats at cinema (average)
Lemons (1 kg) (supermarket)	Laundry (one shirt) (mid-priced outlet)	Low priced car (900-1299 cc) (low)
Bananas (1 kg) (supermarket)	Dry cleaning, man's suit (mid-priced outlet)	Family car (1800-2499 cc) (low)
Lettuce (one) (supermarket)	Dry cleaning, woman's dress (mid-priced outlet)	Deluxe car (2500 cc upwards) (low)
Peas, canned (250 g) (supermarket)	Dry cleaning, trousers (mid-priced outlet)	Cost of a tune up (but no major repairs) (low)
Peaches, canned (500 g) (supermarket)	Aspirins (100 tablets) (supermarket)	Hilton-type hotel, single room, one night including breakfast (average)
Sliced pineapples, canned (500 g) (supermarket)	Razor blades (five pieces) (supermarket)	Moderate hotel, single room, one night including breakfast (average)
Beef: steak, entrecote (1 kg) (supermarket)	Toothpaste with fluoride (120 g) (supermarket)	One drink at bar of first class hotel (average)
Beef: stewing, shoulder (1 kg) (supermarket)	Facial tissues (box of 100) (supermarket)	Two-course meal for two people (average)
Beef: roast (1 kg) (supermarket)	Hand lotion (125 ml) (supermarket)	Simple meal for one person (average)
Beef: ground or minced (1 kg) (supermarket)	Lipstick (deluxe type) (chain store)	Regular unleaded petrol (1 l) (average)
Pork: chops (1 kg) (supermarket)	Man's haircut (tips included) (average)	Taxi: initial meter charge (average)
Pork: loin (1 kg) (supermarket)	Woman's cut & blow dry (tips included) (average)	Taxi rate per additional kilometre (average)
Ham: whole (1 kg) (supermarket)	Cigarettes, Marlboro (pack of 20) (supermarket)	Taxi: airport to city centre (average)
Chicken: fresh (1 kg) (supermarket)	Cigarettes, local brand (pack of 20) (supermarket)	International foreign daily newspaper (average)
Fresh fish (1 kg) (supermarket)	Electricity, monthly bill for family of four (average)	International weekly news magazine (Time) (average)
Instant coffee (125 g) (supermarket)	Men's business suit, two piece, medium weight (chain store)	One good seat at cinema (average)
Ground coffee (500 g) (supermarket)	Men's business shirt, white (chain store)	
Tea bags (25 bags) (supermarket)	Socks, wool mixture (chain store)	

Table 2A

Panel Regression, Levels, All Items

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.003208	0.000941	3.41
$bord_{jk}$	0.073104	0.001736	42.12
$pop_{jk,t}$	0.009451	0.001597	5.92

Notes: The equation was estimated using the full panel of 100 items, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|p_{i,j,t} - p_{i,k,t}\right|$, the absolute value of the difference in the log of the price (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. $dist_{jk}$ is the log of the distance (measured in miles as the great circle distance) between cities *j* and *k*. $bord_{jk}$ is a dummy variable that takes on the value of one if the two cities *j* and *k* are in different countries. $pop_{jk,t}$ is the absolute value of the difference in the logs of the populations of cities *j* and *k* in the year 2000. Also included in the regression, but not reported, are dummy variables for each city. Std. err. denotes Huber-White robust standard errors. Number of observations = 176,800.

Table 2B

Panel Regression, First Differences, All Items

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.000857	0.000579	1.48
$bord_{jk}$	0.014425	0.001026	14.07
$pop_{jk,t}$	0.002255	0.000938	2.40

Notes: The equation was estimated using the full panel of 100 items, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|\pi_{i,j,t} - \pi_{i,k,t}\right|$, the absolute value of the difference in the log of the inflation (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 163,200.

Table 3A

Panel Regression, Levels, Food Items Only

Dependent variable	Coefficient	Std. Err.	t-stat
--------------	------------	-----------	--------
$dist_{jk}$	0.005839	0.001477	3.95
$bord_{jk}$	0.079617	0.002508	31.74
$pop_{jk,t}$	0.018186	0.002493	7.29

Notes: The equation was estimated using the 42 items that are food items, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|p_{i,j,t} - p_{i,k,t}\right|$, the absolute value of the difference in the log of the price (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 74256.

Table 3B

Panel Regression, First Differences, Food Items Only

Dependent variable	Coefficient	Std. Err.	t-stat
--------------	------------	-----------	--------
$dist_{jk}$	0.001037	0.001028	1.01
$bord_{jk}$	0.014520	0.00181	8.02
$pop_{jk,t}$	0.003325	0.001697	1.96

Notes: The equation was estimated using the 42 items that are food items, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|\pi_{i,j,t} - \pi_{i,k,t}\right|$, the absolute value of the difference in the log of the inflation (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 68544.

Table 4A

Panel Regression, Levels, Miscellaneous Products Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	-0.000040	0.002026	-0.02
$bord_{jk}$	0.089489	0.004126	21.69
$pop_{jk,t}$	0.006415	0.003344	1.92

Notes: The equation was estimated using the 22 items that are miscellaneous products, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|p_{i,j,t} - p_{i,k,t}\right|$, the absolute value of the difference in the log of the price (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 38896.

Table 4B

Panel Regression, First Differences, Miscellaneous Products Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.000973	0.001072	0.91
$bord_{jk}$	0.017147	0.001984	8.64
$pop_{jk,t}$	0.002032	0.001727	1.18

Notes: The equation was estimated using the 22 items that are miscellaneous products, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|\pi_{i,j,t} - \pi_{i,k,t}\right|$, the absolute value of the difference in the log of the inflation (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 35904.

Table 5A
Panel Regression, Levels, Clothing Items Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.004589	0.002740	1.67
$bord_{jk}$	0.019125	0.006292	3.04
$pop_{jk,t}$	-0.010780	0.004587	-2.35

Notes: The equation was estimated using the 9 items that are clothing items, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left| p_{i,j,t} - p_{i,k,t} \right|$, the absolute value of the difference in the log of the price (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 15912.

Table 5B

Panel Regression, First Differences, Clothing Items Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.000838	0.001607	0.52
$bord_{jk}$	0.004700	0.003315	1.42
$pop_{jk,t}$	0.002460	0.002656	0.93

Notes: The equation was estimated using the 9 items that are clothing items, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left| \pi_{i,j,t} - \pi_{i,k,t} \right|$, the absolute value of the difference in the log of the inflation (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 14688.

Table 6A

Panel Regression, Levels, Durables Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.000036	0.002962	0.01
$bord_{jk}$	0.029864	0.005344	5.59
$pop_{jk,t}$	-0.009670	0.005046	-1.92

Notes: The equation was estimated using the 6 items that are durables, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|p_{i,j,t} - p_{i,k,t}\right|$, the absolute value of the difference in the log of the price (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 10608.

Table 6B

Panel Regression, First Differences, Durables Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	-0.000380	0.001685	-0.22
$bord_{jk}$	0.008923	0.003289	2.71
$pop_{jk,t}$	-0.00014	0.002671	-0.05

Notes: The equation was estimated using the 6 items that are durables, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|\pi_{i,j,t} - \pi_{i,k,t}\right|$, the absolute value of the difference in the log of the inflation (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 9792.

Table 7A

Panel Regression, Levels, Services Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.001664	0.002078	0.80
$bord_{jk}$	0.078399	0.003627	21.61
$pop_{jk,t}$	0.009293	0.003696	2.51

Notes: The equation was estimated using the 21 items that are services, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|p_{i,j,t} - p_{i,k,t}\right|$, the absolute value of the difference in the log of the price (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 37128.

Table 7B

Panel Regression, First Differences, Services Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.000734	0.001008	0.73
$bord_{jk}$	0.017125	0.001468	11.67
$pop_{jk,t}$	0.000943	0.001499	0.63

Notes: The equation was estimated using the 21 items that are services, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|\pi_{i,j,t} - \pi_{i,k,t}\right|$, the absolute value of the difference in the log of the inflation (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 34272.

Table 8A
Panel Regression, Levels, Goods Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.003619	0.001050	3.45
$bord_{jk}$	0.071696	0.001968	36.44
$pop_{jk,t}$	0.009493	0.001762	5.39

Notes: The equation was estimated using the 79 items that are goods, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|p_{i,j,t} - p_{i,k,t}\right|$, the absolute value of the difference in the log of the price (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 139672.

Table 8B

Panel Regression, First Differences, Goods Only

Dependent variable	Coefficient	Std. Err.	t-stat
$dist_{jk}$	0.000889	0.000673	1.32
$bord_{jk}$	0.013707	0.001215	11.28
$pop_{jk,t}$	0.002603	0.001105	2.36

Notes: The equation was estimated using the 79 items that are goods, for 136 city-pairs, with annual data for 1990-2002. The dependent variable is $\left|\pi_{i,j,t} - \pi_{i,k,t}\right|$, the absolute value of the difference in the log of the inflation (expressed in U.S. dollars) of good *i*, between cities *j* and *k*, at time *t*. The regression uses the same independent variables as the regression reported in Table 2A. See that table for definitions. Sample size equals 128928.