American Economic Journal: Macroeconomics 2023, 15(3): 43–63
https://doi.org/10.1257/mac.20200486

Measuring the Cost of Living in Mexico and the United States†

By David Argente, Chang-Tai Hsieh, and Munseob Lee*

We use a dataset with prices and spending on consumer packaged goods matched at the bar code level across the United States and Mexico to measure the price index in Mexico relative to the United States. Mexican prices relative to the United States are 23 percent lower compared to the International Comparisons Project's (ICP) price index. We decompose the 23 percent gap into the biases from imputation, sampling, quality, and variety. Quality bias increases Mexican prices by 48 percent. Imputation, sampling, and variety bias lowers Mexican prices by 11 percent, 13 percent, and 33 percent, respectively. (*JEL* C43, E31, I31, O11, O12)

Indexes of prices across countries are a vital ingredient in estimates of standards of living and real output across countries. The most widely used price indexes are those by the International Comparison Program (ICP). The ICP collects prices of more than a thousand specific products ("items") in multiple countries, which it then aggregates into price indexes of 155 broad product categories ("basic headings").

Despite their widespread use, it is well-known that there are four potential biases in the prices provided by the ICP. The first is imputation bias. The ICP's surveyors are unable to collect the prices of many items. The ICP imputes prices for the items with missing prices using data from other countries. The second is sampling bias. The ICP calculates the price of an item as an average of store-level prices weighted by each store's total sales, but this procedure may not yield prices that reflect what consumers actually pay. Third, products differ in quality across countries, and it is possible that the ICP matches lower-quality items in poor countries with higher-quality ones in richer countries. Fourth, products available in one country are not available in other countries, and no adjustment is made for potential differences in the availability of products across countries.

*Argente: Pennsylvania State University (email: dargente@psu.edu); Hsieh: University of Chicago and NBER (email: chsieh@chicagobooth.edu); Lee: University of California San Diego (email: munseoblee@ucsd.edu). Giorgio Primiceri was coeditor for this article. We thank Robert Inklaar, Esteban Rossi-Hansberg, and Ina Simonovska for helpful comments. Researchers' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. Churn Ken Lee, Jake Orchard, and Jihoon Sung provided excellent research assistance.

†Go to https://doi.org/10.1257/mac.20200486 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

These potential biases have been brought up by many authors. For example, Deaton (2010) discusses how the ICP imputes missing prices, Deaton and Heston (2010) bring up the potential bias due to sampling and quality, and Feenstra, Xu, and Antoniades (2020) address the issue of variety bias in the ICP. However, although these biases are well understood conceptually, we have very little evidence on their empirical magnitude. Our goal in this paper is to use a new dataset to measure these biases empirically for the specific case of the price index of nondurable goods in Mexico relative to the United States.

Our data are the Nielsen Consumer Panel data for Mexico and the United States (the Nielsen data). The Nielsen data collect data on spending on nondurables for 40,000–60,000 households in the United States and 6,000 households in Mexico. Households in the two countries use in-home scanners or diaries to record their purchases of packaged goods. The Nielsen data include information on prices and quantities of nondurable goods (identified by a 12-digit bar code) purchased by each household on each shopping trip and from each retail store visited. The nondurables in the Nielsen data account for 60–65 percent of total nondurable spending in the two countries.

The richness of the Nielsen data allows us to measure potential biases in the ICP for nondurable goods. First, the Nielsen data have better coverage of items across both countries within the basic headings the data cover. As a result, we can construct a price index similar to the one produced by the ICP but that does not rely on price imputations. We find that the ICP's imputation of missing prices overstates prices by 11 percent in Mexico compared to the United States.

Second, the Nielsen data have information on prices and expenditures on all products for a representative sample of households in each country. The Nielsen data also identify the retail store where each purchase was made. We can therefore construct two price indexes, one using weights from household expenditures and another based on a store's total sales. Mexican prices (compared to the United States) aggregated from consumer expenditure weights are 13 percent lower compared to prices aggregated from weights that reflect a store's total sales. This gap comes from the fact that Mexican households shop more frequently and are more likely to purchase only lower-priced items in a given store compared to American households. Thus, using total sales in a store as weights overstates average prices in both countries, but more in Mexico compared to the United States.

Third, the Nielsen data include bar code information, which we use to identify more than 5,000 identical bar codes in the two countries. Mexican prices (compared to the United States) among products with the same bar code are 48 percent higher compared to the price index of "comparable" products calculated following the ICP's methodology. This suggests that, as conjectured by Deaton and Heston (2010), the ICP may match low-quality products in low-income countries with high-quality products in high-income countries.

Fourth, since we observe all the purchases made by households, we can measure the importance of products available to Mexican consumers but not to American consumers and vice versa. We find that Mexican varieties missing in the US market matter more than US products missing in Mexico. When we take into account the differences in the availability of varieties in the two countries, effective prices in

Mexico are 33 percent lower compared to the United States. The net effect of all four adjustments, for imputation, sampling, quality, and variety, lowers Mexican prices relative to the United States by 23 percent compared to the ICP.

This paper builds on recent work that uses alternative micro-data to estimate price indexes across countries. Specifically, Cavallo et al. (2018); Cavallo, Feenstra, and Inklaar (2021); and Simonovska (2015) use online data, and Feenstra, Xu, and Antoniades (2020) use scanner data of toothpaste, laundry detergent, personal wash items, and shampoo to estimate prices across countries. Our contribution is to bring new data to this long-standing question and to empirically measure multiple potential biases in the ICP using these new data. Specifically, our data contain detailed data on both prices and quantities on the majority of nondurables for a representative sample of consumers.

The four biases we document for the price index in Mexico relative to the United States are also potentially present in measures of inflation. For example, Nordhaus (1996); Bils and Klenow (2001); and Bils (2009) document the bias in price indexes over time because quality improvements are not fully taken into account. Broda and Weinstein (2010) quantify the size of the variety bias in the United States CPI over time. Coibion, Gorodnichenko, and Hong (2015) study the role of store-switching, and Chevalier and Kashyap (2019) document the importance of price discrimination in the bias in the CPI's sampling weights. We are not aware of a study that measures the importance of all these biases in the price index over time taken together in one consistent dataset.[1]

The rest of the paper is organized as follows. Section I describes the data. In Section II we develop our price index. Section III presents our results and decomposes the gap between our price index and the ICP into the contribution of imputation, sampling, quality, and variety. The last section concludes.

I. Data Description

In this section, we describe the data collected by the ICP, the Nielsen data, and the matched data we constructed from the Nielsen data matched to the items covered in the ICP.

A. *International Comparison Program*

The ICP is a statistical initiative that collects prices on more than a thousand detailed products ("items") around the world. The prices of different items are aggregated into 155 broad product categories ("basic headings") that cover all the components of GDP. Approximately 53 basic headings refer to goods, of which 33 are nondurables. The publicly available data have price indexes for the basic

[1] There is also a literature that uses Engel's law to measure price indexes over space and time without the need of price data. See Hamilton (2001); Costa (2001); and Nakamura, Steinsson, and Liu (2016), who quantify biases in the CPI of the United States and China. Almås (2012) and Atkin et al. (2020) apply this methodology to estimate welfare differences across income groups.

headings. We use the restricted 2011 ICP micro-data for Mexico and the United States, which have the prices at the item level.

The ICP specifies the characteristics of a "representative product" for each item. The specification of a representative product includes quantity and packaging (e.g., 250 milliliters of milk), source (e.g., produced domestically or imported), seasonal availability (e.g., year-round or only seasonal), product characteristics, and brand. For example, the representative product for the item "Baby Diapers" is a well-known brand, containing between 18 and 24 pieces, either classic or basic type, with a size between 4 and 9.5 kg, and with a multipack package.

Prices of the representative product of each item are collected from a sample of retail establishments chosen based on their total sales. Prices are collected from products in a retail establishment that meet the specifications of the representative products. The ICP then calculates the average price of an item as a weighted average of the unit price of the representative product across all the sampled stores, where the weights are the total sales of each store.

There are two points to note about the ICP's sampling. First, we do not know the exact product chosen as the representative product in each store. Although the goal is to price the same product in multiple stores, the unavoidable problem is that stores differ in the products they sell. Therefore, it is likely that surveyors choose different products as the representative product of an item in different stores. Second, the weights used to aggregate prices in each store are calculated using the total sales of each store instead of the sales of an item in each store. The problem, as we will see later, is that total sales of a store as a share of total sales in all stores can be very different from weights calculated from the sales of an item in a store as a share of total sales of the item in all stores. For example, whole milk can account for a large share of sales in grocery stores but is relatively unimportant in sales for gas station convenience stores.

The ICP also collects national accounts expenditures for each basic heading but not for the items within each basic heading. These expenditures are used to aggregate basic headings into an aggregate price index. Within each basic heading, the ICP does not have expenditure weights at the item level. Instead, it classifies each item as "important" or "less important." For the nondurables in Mexico and the United States that we consider, the ICP classifies the majority of these items as "important."

B. *Mexico and US Nielsen*

The Nielsen data for the United States track the shopping behavior of 40,000 to 60,000 households in 48 contiguous states plus Washington, DC. Each household uses in-home scanners to record their purchases. The US data contain slightly under 1 million distinct 12-digit bar codes. For each bar code, the data contain information on the brand, size, packaging, and other rich sets of product features. We combine the information on the price paid by the consumer with Nielsen's data on the products' characteristics to calculate the unit price of each product (e.g., price per ounce). In what follows, we use "price" as a shorthand for a product's unit price.

The data also contain information on each purchasing trip the panelist makes, including information on the retailer, the retailer's location, the date of the transaction, and the expenditures and prices of each bar code purchased in each store. Furthermore, the data have demographic variables such as age, education, annual income, marital status, and employment that are updated annually based on surveys sent to the households. Nielsen constructs projection weights that make the sample representative of the US urban population that we use in our calculations.

The Nielsen data for Mexico track the shopping behavior of 6,000 households for the years 2012–2013.[2] The sample is representative of all cities over 50,000 people and covers 55 cities in Mexico. Instead of using in-home scanners, households record their spending in diaries that are collected biweekly by Nielsen. Nielsen's data for Mexico contain approximately 55,000 distinct bar codes.

The Mexican data contain detailed information on each shopping trip (date, store, amount spent), transaction-level information for each product purchased (quantity, price, deals, coupons), as well as detailed product-level characteristics (brand, size, packaging, flavor). As in the US data, we calculate the unit price from the products' characteristics and the price paid by the consumer. The data also include demographic variables at the household level, such as the occupation of the household members, education, age, and family size. As in the case of the United States, Nielsen constructs projection weights that make the sample representative of the Mexican urban population that we use in our calculations.

C. *Nielsen Matched to ICP Items*

We use the ICP's definition of an item along with Nielsen's description of the characteristics of each bar code to assign bar codes to items in the ICP. We match the bar codes in Nielsen in Mexico and the United States to 71 ICP items in 18 basic headings.[3] These 71 items account for 60 percent of aggregate expenditures on nondurables in Mexico and 65 percent in the United States. The basic headings of nondurables in the ICP not covered by the Nielsen data are products without bar codes, such as fresh meats and fruits.

In our matched sample of 71 ICP items, there are 45 items the prices of which are collected by the ICP in Mexico and the United States. Table 1 lists these 45 items along with the number of bar codes in Nielsen we classify under each item. The

[2] Since the ICP sampling is conducted every six years, and the Mexican scanner data are only available from 2012 to 2013, we focus on the cross-sectional patterns of the 2011 ICP estimates.

[3] There are a total of 117 items in the ICP in the 18 basic headings we match with the Nielsen data. We aggregated the ICP items that we cannot distinguish from the bar code description in the Nielsen data. For example, the description of the characteristics of bar codes of breads does not distinguish between "White Bread (Not Sliced)" and "Sliced White Bread" (two distinct items in the ICP), so we aggregate these two ICP items into one. This aggregation reduces the number of ICP items we match to Nielsen to 104. The "new" items after we aggregate are White Bread (not sliced and sliced); Dried/Instant Noodles (dried and instant noodles); Milk, unskimmed (pasteurized and ultra-pasteurized); Cheese (cheddar, processed, Camembert, and Gouda); Tomato Paste (small and large), Ice Cream (cornetto-type and packed), Black Tea (bags and loose leaves), Canned Beer (domestic and imported), Coffee Roasted (Arabica and Robusta), and Detergent Powder (washing machine and hand wash). Among these 104 items, 6 of them are not covered by Nielsen and 27 are priced only in the US Nielsen data but not in the Mexican data. We end up with a total of 71 ICP items for which we have Nielsen prices in Mexico and the United States. We find no significant difference in ICP's Mexican prices (compared to US prices) between matched and unmatched items.

Table 1—Items with Prices in the ICP

	No. bar codes MX	No. bar codes US	No. common
Cornflakes	1,312	5,813	263
White bread	159	1,683	42
Roll	77	4,662	16
Sandwich biscuits/cookies	568	709	3
Butter biscuits	18	26	0
Flavored biscuits/cookies sweet	1,803	13,124	134
Spaghetti	147	1,971	21
Dried/instant noodles	401	1,169	60
Canned tuna without skin	328	864	5
Milk, unskimmed	763	1,521	15
Milk, low-fat, pasteurized	46	3,486	9
Yogurt, plain	1,210	6,207	13
Sour cream	206	909	5
Cheese	1,445	15,713	99
Cream cheese	72	1,113	19
Potato chips	725	5,043	118
Tomato paste	104	1,022	20
Canned sweet corn/maize	186	877	37
White sugar	360	375	3
Strawberry/apricot jam	164	901	12
Orange marmalade	10	204	2
Chocolate bar	882	3,406	75
Ice cream	547	8,552	93
Cooking salt	235	1,468	13
Tomato ketchup	167	627	31
Chicken extract (bouillon/stock cube)	258	208	11
Baby food	45	1,148	4
Cocoa powder, tin	201	558	8
Instant coffee	362	463	31
Black tea	15	2,307	0
Mineral water	920	4,095	43
Carbonated soft drink (small)	267	4,376	47
Carbonated soft drink (large)	1,495	7,669	41
Apple juice	204	1,742	27
Canned beer	231	1,025	22
Bottled beer	289	2,196	23
Detergent powder	821	710	41
Liquid window cleaner	70	355	4
Kitchen paper roll	481	1,567	57
Dishwashing detergent	189	1,405	57
Toothpaste, tube	420	1,275	75
Shower gel	719	7,457	133
Regular sanitary pad/napkin	554	1,340	20
Shampoo	2,164	4,321	199
Toilet paper—multipack	1,029	2,042	49

Note: The table reports the number of bar codes in Mexico and the United States in the 45 items the prices of which are collected by the ICP.

2011 ICP imputes the prices for Mexico and/or the United States for the remaining 26 items. These items are listed in Table 2, again with the number of Nielsen bar codes we match to each item.

We also match bar codes in the United States and Mexico, using the fact that the two countries use the same bar code system.[4] The third column in Tables 1 and 2

[4] Both countries use the Universal Product Code (UPC), which consists of 12 numeric digits. The organization that assigns these bar codes (i.e., GS1) is present in over 115 countries.

TABLE 2—ITEMS WITH IMPUTED PRICES IN THE ICP

	No. bar codes MX	No. bar codes US	No. common
Wheat semolina (Suji)	8	2	0
Oats, rolled	222	1,898	19
Whole wheat bread	192	1,988	65
Pita bread	2	390	0
Salted crackers	225	2,961	32
Short pasta	210	50	0
Vermicelli (angel hair)	8	30	0
Macaroni	54	3,307	23
Milk, condensed	56	132	2
Milk, powdered	140	567	2
Green olives (with stones)	155	2,354	22
Canned green peas	71	709	4
Canned button mushrooms	101	618	4
Brown sugar	17	652	1
Pineapple jam	20	72	0
Natural honey, mixed blossoms	198	2,721	23
Thin soya sauce	73	337	13
Chili sauce	443	3,095	33
Coffee, roasted	319	5,574	43
Tea, green	12	1,281	0
Orange juice	285	2,634	33
Lemonade	52	1,029	3
All-purpose household cleaner	1,352	4,211	57
Deodorant, roll-on for men	638	119	13
Toilet soap	910	1,787	69
Baby diapers	957	2,399	45

Note: The table reports the number of bar codes in Mexico and the United States in the 26 items where the ICP imputes the price of the item in Mexico or the United States (or in both countries).

shows the number of matched bar codes for each ICP item. It is clear that the number of common bar codes is a small share of the bar codes sold in each country. The majority of bar codes sold in Mexico are not sold in the United States and vice versa.

II. An Exact Price Index

This section derives the ideal price index. We also describe how we use the Nielsen data to construct a price index that mimics the procedure followed by the ICP.

A. *Exact Price Index*

Following the ICP, we assume that utility is a function of consumption in the basic headings and that basic headings are in turn an aggregate of expenditures of specific items. Specifically, the utility of a representative household in a country is given by

$$U = \left[\sum_b \left(\sum_i C_{ib}^{\frac{\eta_b - 1}{\eta_b}} \right)^{\frac{\eta_b}{\eta_b - 1} \frac{\gamma - 1}{\gamma}} \right]^{\frac{\gamma}{\gamma - 1}}, \tag{1}$$

where C_{ib} is consumption of item i in basic heading b. When possible, we omit the country index in the notation. The parameters γ and η_b denote the elasticity of substitution across basic headings and items within basic headings, respectively. The set of basic headings and items are the same for all countries in the ICP. Therefore, the indexes for basic headings and items do not carry an index for the country.

The ICP assumes that an item consists of a single representative product, but there are multiple bar codes that satisfy the ICP definition of the representative product of an item. For example, there are 463 bar codes in the United States and 362 in Mexico that meet the ICP's definition of the representative product for "instant coffee" (see Table 1). To account for the varieties in each item, we assume consumption of an item in a country (say Mexico, indexed by m) is itself a CES aggregate of individual bar codes indexed by k:

$$C_{ib}^{m} = \left[\sum_{k \in m} \left(\varphi_{kib} C_{kib}^{m}\right)^{\frac{\sigma_{ib}-1}{\sigma_{ib}}}\right]^{\frac{\sigma_{ib}}{\sigma_{ib}-1}}. \tag{2}$$

Here, C_{kib}^{m} denotes total physical units and φ_{kib} the quality of bar code k, where the parameter σ_{ib} denotes the elasticity of substitution between the bar codes. Note that the aggregation over the bar codes is only over the set available in each country.[5]

The exact price index of item i in basic heading b in Mexico, taking the United States (indexed by u) as the numeraire, is then given by

$$EPI_{ib} = \left(\frac{\lambda_{ib}^{m}}{\lambda_{ib}^{u}}\right)^{\frac{1}{\sigma_{ib}-1}} \times \prod_{k \in common} \left(\frac{p_{kib}^{m}}{p_{kib}^{u}}\right)^{\omega_{kib}}. \tag{3}$$

The first term in equation (3) is the ratio of the share of total spending on the common bar codes in Mexico relative to the United States, where the spending shares are defined as

$$\lambda_{ib}^{m} \equiv \frac{\sum_{k \in common} p_{kib}^{m} C_{kib}^{m}}{\sum_{k \in m} p_{kib}^{m} C_{kib}^{m}} \quad \text{and} \quad \lambda_{ib}^{u} \equiv \frac{\sum_{k \in common} p_{kib}^{u} C_{kib}^{u}}{\sum_{k \in u} p_{kib}^{u} C_{kib}^{u}}.$$

The second term is the geometric mean of the ratio of the price in Mexico relative to the United States of bar code k for the bar codes common to the two countries, weighted by the logarithmic mean of the expenditure shares of the bar code.[6] The price of a bar code in equation (3) is a weighted average of the price of the same bar code in all the stores in a country,

$$p_{kib}^{m} = \prod_{s \in m} \left(p_{skib}^{m}\right)^{\phi_{skib}^{m}} \quad \text{and} \quad p_{kib}^{u} = \prod_{s \in u} \left(p_{skib}^{u}\right)^{\phi_{skib}^{u}},$$

[5] In online Appendix Section A, we extend the utility function at the item level to allow for nonhomothetic preferences. There we show that the bias in the ICP price index between Mexico and the United States is almost the same when we account for nonhomothetic preferences as in our baseline case with homothetic preferences.

[6] The logarithmic mean is $\omega_{kib} \equiv \frac{s_{kib}^{m} - s_{kib}^{u}}{\ln s_{kib}^{m} - \ln s_{kib}^{u}} \Big/ \sum_{k \in common} \frac{s_{kib}^{m} - s_{kib}^{u}}{\ln s_{kib}^{m} - \ln s_{kib}^{u}}$, where $s_{kib}^{m} \equiv \frac{p_{kib}^{m} C_{kib}^{m}}{\sum_{k \in m} p_{kib}^{m} C_{kib}^{m}}$ and $s_{kib}^{u} \equiv \frac{p_{kib}^{u} C_{kib}^{u}}{\sum_{k \in u} p_{kib}^{u} C_{kib}^{u}}$.

where p^m_{skib} is the price of bar code k in store s (in Mexico) and the weights are the share of spending on each bar code in the store.[7]

The exact price index of item i in basic heading b requires at least one common bar code between Mexico and the United States. For 63 of the 71 items in the matched data, we observe at least 1 common bar code. For the remaining eight items that do not have a common bar code, we account for these items with a variety correction term at the item level. Specifically, define I_C as the set of items with common bar codes. The exact price index for the basic heading can be defined as

$$EPI_b = \left(\frac{\lambda_b^m}{\lambda_b^u}\right)^{\frac{1}{\eta_b - 1}} \times \prod_{i \in I_C} EPI_{ib}^{\omega_{ib}^*}, \tag{4}$$

where ω_{ib}^* is logarithmic mean of the expenditure share of each item and the spending shares are defined as

$$\lambda_b^m \equiv \frac{\sum_{i \in I_C} p_{ib}^m C_{ib}^m}{\sum_i p_{ib}^m C_{ib}^m} \quad \text{and} \quad \lambda_b^u \equiv \frac{\sum_{i \in I_C} p_{ib}^u C_{ib}^u}{\sum_i p_{ib}^u C_{ib}^u}.$$

The second term in equation (4) is the weighted geometric average of the price index of the items that can be priced from equation (3). The first term in equation (4) is the ratio of the spending shares on items with common bar codes in the two countries. This term captures the weighted geometric average of the price index of the items that cannot be priced from equation (3).

Finally, the aggregate exact price index is a weighted geometric average of the exact price index of each basic item:

$$EPI = \prod_b EPI_b^{\omega_b}, \tag{5}$$

where ω_b is the logarithmic mean of the expenditure shares of each basic heading and EPI_b is given by equation (4).

B. *A "Pseudo-ICP" Price Index*

There are two differences between the ideal price index we will calculate and the ICP's price index. First, the underlying data (the Nielsen data) are different from the data used by the ICP. Second, there is the difference in methodology. To isolate the effect of methodology, we use the Nielsen data to construct a price index that mimics the ICP. We call this a "Pseudo-ICP" price index.

We proceed in four steps. First, for each ICP item, we identify the set of bar codes in each store that meet the ICP's specifications for the representative product.[8] From

[7] Note that in the data we observe multiple prices for the same bar code purchased from multiple stores. Online Appendix Section B shows that the ideal price of a bar code can be derived as a CES price index from a consumer's discrete choice problem and that a first-order approximation to this index is the Cobb-Douglas price index.

[8] We also mimic the fact that the ICP survey only covers four cities in Mexico (Mexico City, Guadalajara, Monterrey, and Puebla), while it covers all urban areas in the United States.

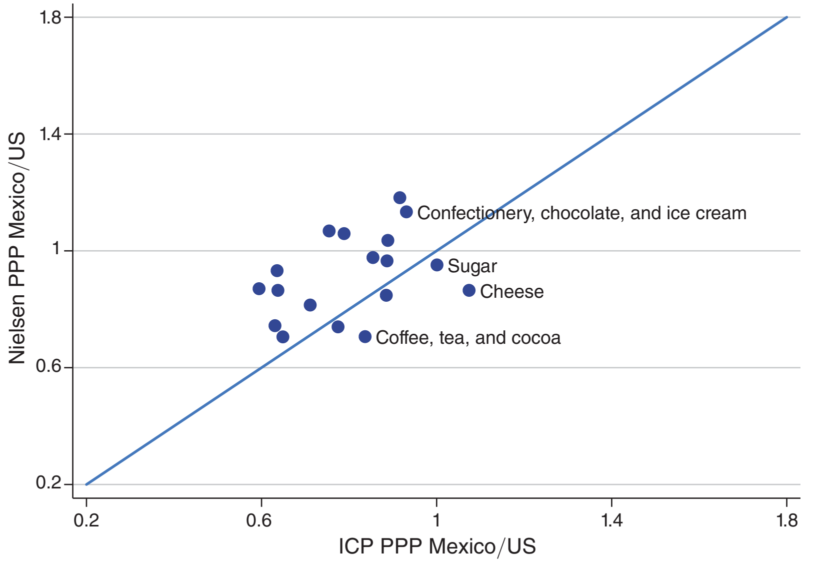

Figure 1. Pseudo- versus Actual ICP Price Index

Note: The figure plots the Pseudo-ICP index of each of the 18 basic headings against the index published by ICP.

this set of bar codes, we pick the bar code with the largest volume of sales in each store as the representative product of the item in the store. The representative product of an item is therefore not the same in each store. We obviously do not know the exact product in each store surveyed by the ICP's price surveyors, but we believe this procedure is a good approximation of how the ICP chooses the representative product of an item in each store. The price of an item is then the geometric mean of the price of the store-specific representative product weighted by the store's total sales.

Second, the ICP does not collect prices for the 26 items in Table 2. Instead, it imputes the prices for these items using data from other countries. We estimate the imputed prices by comparing the basic heading price indexes reported by the ICP with the geometric mean of the item-level prices they do collect. If, for example, a basic heading price reported by the ICP is higher than the geometric mean of item-level observed prices, we infer that the imputed prices for the average item with a missing price are higher than those of other items in the same basic heading.

Third, the price index of a basic heading is an equally weighted geometric average of the price of the items in each basic heading.[9] Figure 1 shows the scatterplot of the Pseudo-ICP price index of a basic heading (Mexico relative to the United States) versus the price index published by the ICP. The correlation is not one, but the Pseudo-ICP price index generally aligns with the ICP's price index. The median of

[9] In principle, the ICP classifies items as "important" and "less important" and gives "important" items more weight when aggregating to the basic heading level. However, most of the nondurable ICP items in Mexico and the United States that we match to the Nielsen data are classified as "important."

the Pseudo-ICP price index at the basic heading level is 0.90, whereas the median of the actual ICP price index is 0.81. The regression coefficient is 1.12 and statistically significant at the 1 percent level. At the item level, the median of the Pseudo-ICP price index is 0.85, whereas the median of the numbers published by the ICP is 0.73. A linear regression across the prices at the item level yields coefficient 1.10 that is also statistically significant at the 1 percent level.

The final step is to aggregate the price of a basic heading into an aggregate price index. The ICP calculates the geometric average of the price index of a basic heading across all basic headings. In the case of the price index of Mexico versus the United States, the price indexes of the basic headings are first averaged using Mexico's weights, then averaged using the US weights, and then the geometric mean of the two is taken. The result is a Fisher index, which facilitates multilateral comparisons since such indexes are transitive (i.e., price comparisons between two countries are the same whether it is computed directly or indirectly through a third country) and base country invariant (i.e., price comparisons between two countries are the same regardless of the choice of base country). Since we are only comparing two countries, we aggregate the price index at the basic heading level using the logarithmic weight of each basic heading. This is the theoretically consistent way to aggregate prices with a CES utility function.

III. The Price of Nondurables in Mexico versus the United States

In this section, we calculate the exact price index from the Nielsen data. We decompose the gap between the exact and the Pseudo-ICP price index into the bias due to imputation, sampling, quality, and variety.

We need the elasticity of substitution between bar codes within each item. We follow Feenstra's (1994) procedure as extended by Broda and Weinstein (2006, 2010). The procedure consists of estimating a demand and supply equation for each bar code using the information on prices and quantities of each bar code from the Nielsen data.[10] The average of the elasticity of substitution across bar codes we obtain is 6.56, with standard deviation of 3.13.[11]

We use these estimates of the elasticity of substitution between bar codes, along with the Nielsen data, to estimate the exact price index of each item from equation (3). We then use equation (4) to aggregate the price of an item using logarithmic weights of each item in a basic heading, which we then aggregate into an aggregate price index using logarithmic weights for the basic heading as equation (5). Column 1 in Table 3 shows that the exact price index is 0.65, which implies that Mexican prices are 35 percent lower than in the United States.

The second column shows the Pseudo-ICP price index. Here, we take the weighted average of the Pseudo-ICP price of the basic heading (shown in Figure 1) using the same logarithmic weights for each basic heading that we used to aggregate the exact

[10] See online Appendix Section C for more details.

[11] We check the robustness of these results by calculating price indexes and biases with the common elasticity of substitution, 6. The estimated exact price index (0.64) is very close to the item-specific elasticity of substitution (0.65). See online Appendix Section D for more details.

Table 3—Cost of Living in Mexico versus the United States

Aggregate price index		Bias due to				Aggregate bias
Exact	Pseudo-ICP	Imputation	Sampling	Quality	Variety	
0.65	0.86	0.89	0.87	1.48	0.67	0.77

Notes: The table reports the aggregate exact price index, Pseudo-ICP price index, and the gap between the exact and ICP index due to imputation, sampling, quality, and variety. Aggregate bias is the product of the bias due to imputation, sampling, quality, and variety.

price index of each basic heading. The aggregate Pseudo-ICP price index, shown in column 2 in Table 3, is 0.86, which suggests that Mexican prices are only 14 percent lower than in the United States. The exact price index calculated from the same data suggests that Mexican prices are 35 percent lower, which indicates an aggregate bias of 23 percent.

Recall that we use the same logarithmic weights at the basic heading level to aggregate the exact and Pseudo-ICP price indexes for each basic heading. Therefore, the difference between the aggregate exact and the ICP price index in Table 3 comes from the price index at the basic heading level. In turn, the gap between the two indexes at the basic heading level is the product of the biases due to sampling, quality, and variety for each basic heading:

$$\frac{EPI_b}{ICP_b} = Imputation\ Bias_b \times Sampling\ Bias_b \times Quality\ Bias_b \times Variety\ Bias_b.$$

We now quantify each of these biases.

Imputation Bias.—Out of the 71 items in our sample, the ICP does not collect prices for 26 of them from either Mexico or the United States. The ICP runs a set of country-product dummy (CPD) regressions to impute the missing values for items the prices of which cannot be found (Deaton 2010). We revisit this imputation and quantify the size of the bias. As discussed in the previous section, we back out imputed prices by comparing a basic heading price reported by the ICP and geometric mean of item-level prices within a basic heading.

The imputation bias is defined as follows:

$$\text{(6)} \qquad Imputation\ Bias_b \equiv \left[\prod_i \left(\tilde{p}^m_{ib}/\tilde{p}^u_{ib}\right)^{\frac{1}{N_b}}\right] / ICP_b,$$

where ICP_b is the Pseudo-ICP that uses imputed prices for items not collected from either Mexico or the United States and $\tilde{p}_{ib}$ is the price of the item computed from the Nielsen data. There is no imputation bias if the ICP collects prices for all the items.

Figure 2 plots the ratio of US to Mexican prices of a basic heading, with imputed item-level prices on the y-axis against the price ratio calculated from the Nielsen data on the x-axis. The observations on the 45-degree line are those for which the ICP has prices for all items within the basic heading. For these basic headings, the imputation bias is one. However, on average the majority of the observations are

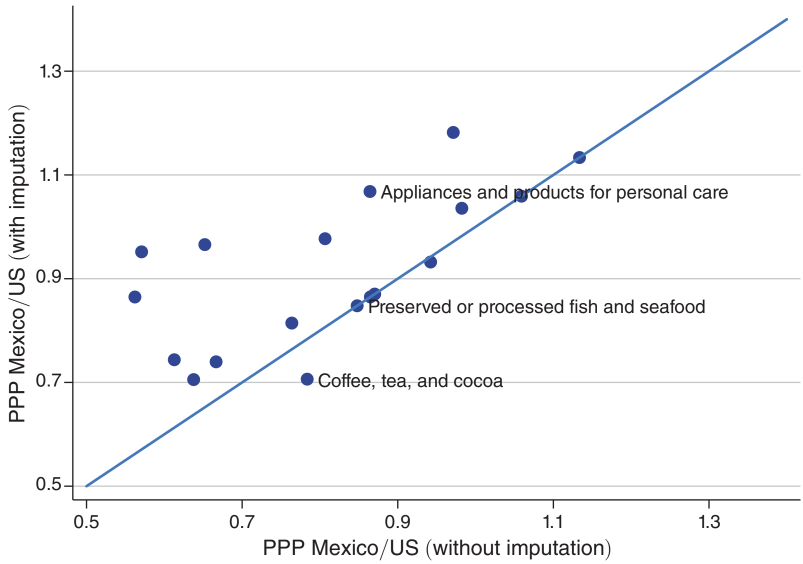

FIGURE 2. IMPUTATION BIAS FOR EACH BASIC HEADING

Note: The figure plots the PPP with and without item-level imputation for each of the 18 basic headings.

above the 45-degree line: the ratio of Mexican to US prices with imputed item-level prices is typically higher than the price ratio calculated from actual price data.

Table 3 shows the net effect of imputation on the aggregate price index. Specifically, the third column shows the imputation bias calculated from equation (6). Imputation lowers Mexican prices by 11 percent compared to the ICP's index.

Sampling Bias.—There are three sources of sampling bias. First, the ICP samples items only from the four cities in Mexico (i.e., Mexico City, Guadalajara, Monterrey, and Puebla). Second, at the item level, the ICP uses an equal-weights geometric average because they lack data on expenditures at the item level. Third, the ICP aggregates prices from each store using total sales of the store instead of the sales of the item in the store as weights. The sampling bias is the product of these three biases:

$$(7)\quad Sampling\ Bias_b \equiv \left[\prod_i \left(\frac{\bar{p}_{ib}^m/\bar{p}_{ib}^u}{\tilde{p}_{ib}^m/\tilde{p}_{ib}^u}\right)^{\frac{1}{N_b}}\right] \times \left[\frac{\prod_i \left(\frac{\bar{p}_{ib}^m}{\bar{p}_{ib}^u}\right)^{\omega_{ib}}}{\prod_i \left(\frac{\bar{p}_{ib}^m}{\bar{p}_{ib}^u}\right)^{\frac{1}{N_b}}}\right] \times \left[\prod_i \left(\frac{\hat{p}_{ib}^m/\bar{p}_{ib}^m}{\hat{p}_{ib}^u/\bar{p}_{ib}^u}\right)^{\omega_{ib}}\right],$$

where $\tilde{p}_{ib}^m$ is the average price of an item computed as the weighted average of store-level prices, where the weights are the total sales of the store considering only the main cities in Mexico; $\bar{p}_{ib}^m$ is the same weighted average considering all cities; and $\hat{p}_{ib}^m$ is the weighted average where the weights are the sales of the particular item in the store.

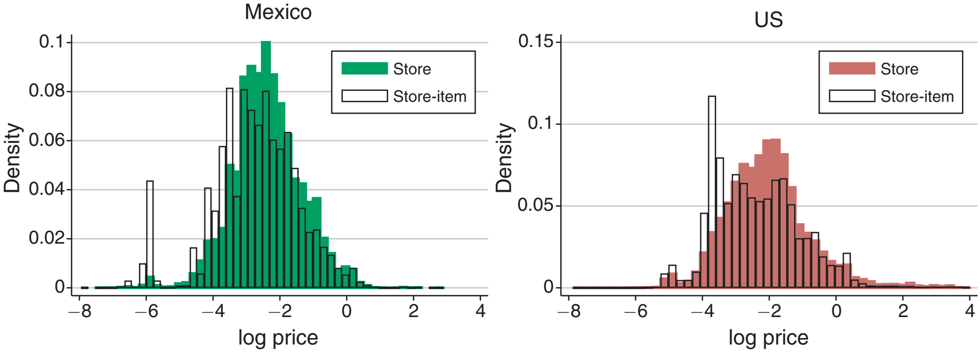

Figure 3. Distribution of Prices Using Store versus Store-Item Weights

Note: The figure shows the distribution of the price of an item calculated as a geometric average of the price of the item at the store level using either the store's total sales or the sales of the item in the store as weights.

The first term denotes the bias from the geographical coverage of the ICP survey in Mexico. By sampling only from the 4 major cities in Mexico, the ICP underestimates the average price of items by approximately 1 percent.[12] The four cities the ICP samples in Mexico are among the largest in terms of population, and their average household income is close to that of the median city in the country (i.e., choosing these four cities is close to choosing a representative city for Mexico.) This, combined with the fact that the ICP only samples urban cities in the United States as well, yields a small geographic bias.

The second term in equation (7) is the sampling bias discussed by Deaton and Heston (2010). It denotes the bias from using the simple average instead of the weighted average of the item-level price, and its magnitude depends on the covariance of the expenditure weights of each item and its price.[13] The covariance is negative if people spend more on items with lower prices. In the Nielsen data, the difference in these covariances between Mexico and the United States is essentially zero. Thus, at least in the case of Mexico relative to the United States, there is little bias from not using expenditure weights at the item level.

The third term in equation (7) is the bias from aggregating store-level prices using the total sales of the store as weights instead of the sales of the item in the store. For example, a fruit vendor may also sell milk, but the vendor's total sales depend mostly on fruit sales instead of milk. The resulting bias in the price index depends on whether consumers in Mexico purchase more products at stores where these products are cheaper compared to consumers in the United States.

Figure 3 shows that this is indeed the case. It plots the distribution of prices for each item, where the items' average price is averaged over the store-specific price

[12]Consistent with the coverage of our data, the ICP samples items from urban areas in all US states (i.e., $\tilde{p}^u_{ib} = \bar{p}^u_{ib}$).

[13]Online Appendix Section E formally shows this relationship and empirically estimates the covariance using a reduced-form approach.

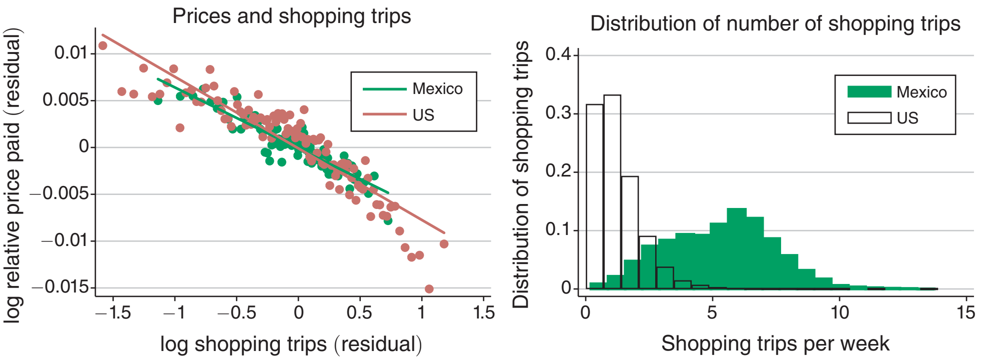

FIGURE 4. SHOPPING INTENSITY AND PRICES IN MEXICO AND THE UNITED STATES

Notes: The left panel plots the de-meaned price paid for a product versus the de-meaned number of shopping trips. The de-meaned product price is the residual from a regression of the log price of a bar code on category, store, and quarter fixed effects. The de-meaned number of shopping trips is the residual of the log number of shopping trips on category, store, and quarter fixed effects. The right panel shows the distribution of the number of shopping trips per household. Number of shopping trips is the number of stores visited by a household in a week.

using the stores' total sales and the sales of each item in the store. For both countries, the distribution weighted using store-item weights is shifted to the left compared to the distribution using store-level weights. Furthermore, the gap between the two distributions is larger for Mexico compared to the United States. Weighting store-specific prices using a store's total sales overstates the average price paid by consumers in the two countries, but more so in Mexico.

Figure 4 shows why the gap between the two price distributions is larger for Mexico. The left panel of Figure 4 plots the de-meaned price of a bar code paid by a household in a shopping visit against the number of shopping visits of the household in the Nielsen data.[14] Households in Mexico and the United States that shop more frequently pay lower prices for the same product compared to households that shop less frequently. The right panel of Figure 4 shows the distribution of the number of shopping trips per household. A typical US household makes one shopping trip per week, whereas in Mexico the number is five. The average Mexican household shops more intensely than the typical American household and thus, buys more of the cheaper products in a given store.

Table 3 shows the effect of the difference in shopping behavior between Mexico and the US on the price index. Specifically, the fourth column shows the weighted mean of the sampling bias at the basic heading level calculated from equation (7), where the weights are the logarithmic share of each basic heading. The difference in shopping intensity between Mexico and the United States lowers Mexican prices by 13 percent compared to the ICP because the ICP's sampling weights ignore the effect of shopping intensity.

[14] The number of shopping visits is the number of retail outlets a household makes purchases from in a week.

Table 4—Price Gap between Mexico and United States for All versus Common Bar Codes

Sampling	All bar codes	Common bar codes
Representative product	−28.0%	−0.5%
All products in item	−42.1%	0.5%

Notes: The table reports average percent difference in item-level prices between Mexico and the United States. We take the weighted average with item-level expenditure weights. Column 1 considers all bar codes available in the two countries within each item. Column 2 restricts the sample to bar codes sold in Mexico and the United States. Row 1 chooses one representative product per store. Row 2 considers all the bar codes and calculates the weighted average of the prices of chosen products.

Quality Bias.—In 63 out of the 71 items in our matched sample, there is at least 1 common bar code sold in the 2 countries. With at least one common bar code within an item, we can calculate quality and variety biases separately.[15] The quality bias in the ICP reflects the effect of matching high-quality goods in one country with a low-quality good in another country. We measure quality bias of an item as

$$\text{(8)} \qquad Quality\ Bias_{ib} \equiv \prod_{k \in common} \left(\frac{p^m_{kib}}{p^u_{kib}}\right)^{\omega_{kib}} \Big/ \left(\frac{\hat{p}^m_{ib}}{\hat{p}^u_{ib}}\right).$$

The denominator in equation (8) is the price gap between Mexico and the United States at the item level computed as the average price of the "representative product" in each store. The numerator is the price gap where the "representative product" in a store is chosen from the set of bar codes sold in both Mexico and the United States.[16] Quality bias is greater than one if the representative product of an item in a store in Mexico is of lower quality than the representative products of the same item in the United States.

Table 4 shows that this is indeed the case. The first row follows the ICP's methodology and chooses one representative product per store to calculate the average price of an item. The first column shows the average percent difference between Mexico and the United States in the price of an item, which is 28.0 percent. The second column picks the representative product of an item in the store chosen from the bar codes common to the two countries. The average price gap using products with the same bar code is close to zero. The second row provides a similar number but where the average price of an item is calculated as a weighted average of the price of all the bar codes in the item. The first column considers all the bar codes in an item; the second column only considers bar codes sold in the two countries. Average prices calculated from all the bar codes are 42.1 percent lower in Mexico.

[15] The relative magnitude of the variety bias and the quality bias could be affected if we identify too few bar codes in common, particularly if there are a significant number of products among the unmatched bar codes that are in fact the same products. Nonetheless, the aggregate bias is the product of the variety and the quality biases. Therefore, as long as products with the same bar code are in fact the same product in the two countries, the aggregate bias is not affected by the possibility that some of the unmatched products may in fact be the same product.

[16] Both the numerator and the denominator in equation (8) weight the price of the item in each store by the sales of the item in the store.

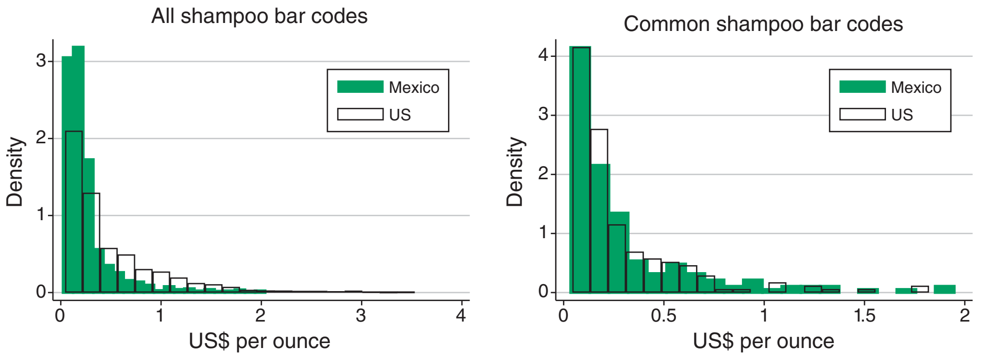

Figure 5. Price Distribution for All and Common Shampoo Bar Codes

Notes: The left panel shows the price distribution of bar codes of all shampoo products in the United States and Mexico. The right panel plots the price distribution of shampoo products sold in both countries.

The second column restricts to bar codes sold in the two countries. Here, the average price in Mexico is very close to the average price in the United States.

Figure 5 shows the distribution of the price of shampoo. The left panel shows distribution of the unit price of shampoo in the United States and Mexico for all shampoo bar codes. The right panel shows the distribution of the unit price of shampoo only for shampoo bar codes sold in both countries. As can be seen, the average shampoo in the United States is more expensive than in Mexico, but the difference is much smaller for shampoos sold in both countries.

Table 3 shows the aggregate effect of the patterns shown in Figure 5 and Table 4 on aggregate quality bias (for the 63 items for which this calculation is possible). The effect is large: quality bias increases Mexican prices by 48 percent relative to the United States.

Variety Bias.—We have already seen that the majority of Mexican bar codes are not available in the US market and vice versa. The ICP price index is biased due to missing varieties if the importance of Mexican bar codes not available in the United States is different from the importance of US bar codes not available in Mexico. We measure variety bias (of an item) by

$$\text{(9)} \qquad Variety\ Bias_{ib} \equiv \left(\frac{\lambda_{ib}^{m}}{\lambda_{ib}^{u}}\right)^{\frac{1}{\sigma_{ib}-1}}.$$

This is the ratio of the expenditure share of common bar codes in Mexico relative to the expenditure share of common products in the United States, as in Feenstra (1994).

Figure 6 plots the distribution of share of spending on common bar codes relative to the spending share on the average bar code in the country. US households spend more on typical common bar codes, while the opposite is true in Mexico. This means that missing Mexican varieties are more important compared to missing US varieties.

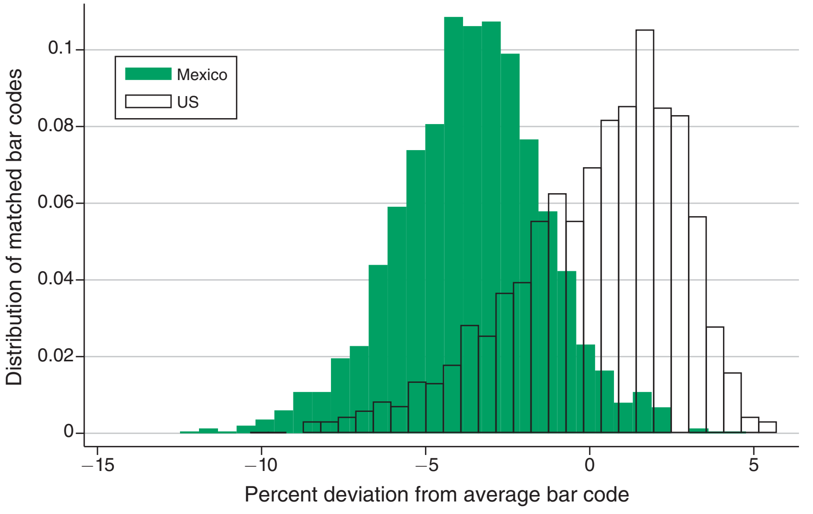

Figure 6. Distribution of Revenue of Bar Codes Sold in Both Countries

Note: The figure shows the distribution of revenue generated by the common bar codes relative to the average product in each country.

Bar codes that are present in Mexico and not present in the United States are both very popular in Mexico (in terms of expenditures) and also very cheap relative to those sold in both countries. In the case of cereals, for instance, unmatched bar codes in Mexico are mainly produced by Mexican firms at a significantly lower price per unit.

Table 3 summarizes the effect of variety bias on the price index (for items with common bar codes in the two countries). We estimate that varieties only available to Mexican consumers lower the effective prices paid by Mexican consumers by 33 percent compared to the United States.[17]

Bias from Items without Common Bar Codes.—Finally, 8 of the 71 items in our sample do not have common bar codes across the two countries. For these items, we cannot separately calculate quality and variety biases, but we can measure the product of these two biases on the price of a basic heading from the following equation:

$$\text{(10)} \quad \textit{Quality and Variety Bias From Items Without Common Bar codes}_b$$

$$= \left(\frac{\lambda_b^m}{\lambda_b^u}\right)^{\frac{1}{\eta_b - 1}} \times \frac{\prod_{i \in I_C} \left(\frac{\hat{p}_{ib}^m}{\hat{p}_{ib}^u}\right)^{\omega_{ib}^*}}{\prod_i \left(\frac{\hat{p}_{ib}^m}{\hat{p}_{ib}^u}\right)^{\omega_{ib}}},$$

[17] Our results from comparing Mexico and the United States differ from those found in studies using within-country variation (e.g., Handbury and Weinstein 2015). Intuitively, the variety bias captures the relative share of expenditures in common goods across locations. Within countries, households residing in a low-income state consume many of the same goods that are available in higher-income states. Across countries, variety bias is driven by country-specific factors, and, in the case of Mexico and the United States, the importance of local brands in Mexican households' consumption is remarkable. As a result, the share of expenditures in common goods for Mexico is substantially lower than that of the United States, even if the United States has a higher average household income.

where as before, I_C is the set of items with common bar codes. In our data, equation (10) is very close to one for all the basic headings with at least one item that does not have a common bar code in the two countries. So at least for these eight items for which we need to calculate equation (10), the net effect of quality and variety biases does not change the price index of Mexican goods relative to the United States. Note that this is not the case for the items for which we can separately estimate the effect of quality and variety biases. For these 63 items, the net effect of these 2 biases decreases Mexican prices by 1 percent $(1.48 \times 0.67 = 0.99)$ compared to the United States.

Price Index Bias for Basic Headings.—The implication from Table 3 is that the ICP overstates Mexican prices because of imputation, sampling, and variety bias but understates Mexican prices because of quality bias. The net effect—aggregate bias in Table 3—is that the exact price index is about 23 percent lower compared to the Pseudo-ICP index.

Table 5 shows the bias in the ICP for the basic headings.[18] The first message is that there is a large amount of heterogeneity in the aggregate bias. The exact price index is similar to the ICP for "sugar" and "coffee, tea, and cocoa." At the other extreme, the ICP price index is significantly lower than the ideal price index for "food products nec" and "nondurable household goods." This heterogeneity suggests that one should be careful about drawing strong inferences from prices obtained from a narrow set of products.

The relative importance of the four biases also differs quite a bit across products. Imputation bias is large for basic headings with a large number of missing prices in the ICP, such as "Pasta products," where four out of six items are missing either from Mexico or the United States. Sampling bias is very large for "appliances and products for personal care" and virtually one for "fresh milk." Mexican households appear to be more price sensitive in where they buy appliances and articles for personal care compared to fresh milk. Quality bias is basically one for fresh milk, sugar, and beer: the price of the average fresh milk, sugar, and beer product is about the same as fresh milk, sugar, and beer sold both in Mexico and the United States. On the other hand, quality bias is very large for cereals, drinks and juices, nondurable household goods, and personal care products. Finally, Mexican varieties are more important than US varieties for Mexican households in almost every basic item (the variety bias term is less than one), but the exception is sugar. For this basic heading, US varieties not sold in Mexico are slightly more important than Mexican bar codes not sold in the United States.

IV. Conclusion

The construction of cross-country price indexes is of crucial importance to compare living standards between countries and to measure global inequality. The ICP has taken on the important and heroic exercise of measuring these prices but faces

[18] The product of the quality and variety terms from items without common bar codes is very close to one for all basic headings.

Table 5—Biases in ICP for Each Basic Heading

	Biases				
Basic heading	Imputation	Sampling	Quality	Variety	Aggregate bias
Other cereals, flour, and other products	0.83	0.53	1.56	0.55	0.38
Bread	1.01	0.95	0.89	0.64	0.55
Other bakery products	0.65	1.27	1.77	0.46	0.67
Pasta products	0.68	1.02	1.41	0.83	0.80
Preserved or processed fish and seafood	1.00	0.94	0.94	0.21	0.19
Fresh milk	1.00	1.04	0.98	0.38	0.38
Preserved milk and other milk products	0.82	1.02	1.41	0.67	0.79
Cheese	1.00	0.88	1.46	0.58	0.75
Frozen, preserved, or processed vegetables	0.82	0.95	0.84	0.60	0.39
Sugar	0.60	1.44	1.10	1.06	1.01
Jams, marmalades, and honey	0.95	0.69	1.05	0.69	0.47
Confectionery, chocolate, and ice cream	1.00	0.75	1.24	0.92	0.86
Food products nec	0.90	0.94	1.61	0.98	1.34
Coffee, tea, and cocoa	1.11	0.97	1.74	0.56	1.05
Mineral waters, soft drinks, and juices	0.94	0.73	1.71	0.56	0.65
Beer	1.00	0.96	1.06	0.87	0.88
Nondurable household goods	0.90	0.99	2.81	0.60	1.50
Appliances and products for personal care	0.81	0.70	1.54	0.66	0.57

Notes: The table reports imputation bias, sampling bias, quality bias, and variety bias for each basic heading. Aggregate bias is defined as the product of the imputation, sampling, quality, and variety biases.

severe data limitations. In this paper, we construct a dataset for two countries that allows us to address some of these data limitations, namely, the fact that the ICP has incomplete information on items across countries, that they do not have expenditure information to weight items appropriately, that they cannot compare exactly the same item across countries, and that they do not have information on differences on the set of products available in each country. Using our alternative data, we estimate that Mexican real consumption is larger relative to the United States than was previously estimated. We identify the imputation of prices and the heterogeneity in shopping behavior, quality of products, and variety availability as important sources of bias in international price comparisons. Overall, our results show that the real nondurable consumption inequality across the United States and Mexico is 23 percent lower than that predicted by the ICP estimates.

There are several generalizable lessons from our study for international price comparisons. First, aggregate bias must be estimated considering all biases jointly. Biases often imply adjustments in opposite directions, so addressing them in isolation could lead to drastically different conclusions about the comparison of the standards of living across countries. Second, because average prices are correlated with income per capita, the magnitude of the quality bias increases when comparing unequal countries. Third, the relevance of the variety bias is not correlated with income but instead with the importance of domestic varieties in overall consumption. For this reason, using within-country variation yields the opposite results to using cross-country variation. Fourth, the bias from the limited geographical coverage of the ICP survey in Mexico is not large, indicating that coverage is less relevant if the sampling for each country only focuses on urban areas, as it does in the case of Mexico and the United States. Fifth, the bias from using the simple average instead of the weighted average of the item-level price is not large; the covariance of the expenditure weights of each item and its price is not large for either the United States or Mexico. Sixth, a country's shopping intensity is correlated with the relevance of the sampling bias. This is because weighting store-specific prices using a store's total sales overstates the average price paid by consumers in all countries. Finally, we find a great deal of heterogeneity in the size of aggregate bias across different product categories. As more product-level data become available, we will be able to get better estimates for aggregate price index beyond consumer packaged goods.

REFERENCES

Almås, Ingvild. 2012. "International Income Inequality: Measuring PPP Bias by Estimating Engel Curves for Food." *American Economic Review* 102 (2): 1093–1117.

Argente, David, Chang-Tai Hsieh, and Munseob Lee. 2023. "Replication data for: Measuring the Cost of Living in Mexico and the United States." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E153241V1.

Atkin, David, Benjamin Faber, Thibault Fally, and Marco Gonzalez-Navarro. 2020. "Measuring Welfare and Inequality with Incomplete Price Information." NBER Working Paper 26890.

Bils, Mark. 2009. "Do Higher Prices for New Goods Reflect Quality Growth or Inflation?" *Quarterly Journal of Economics* 124 (2): 637–75.

Bils, Mark, and Peter J. Klenow. 2001. "Quantifying Quality Growth." *American Economic Review* 91 (4): 1006–30.

Broda, Christian, and David E. Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121 (2): 541–85.

Broda, Christian, and David E. Weinstein. 2010. "Product Creation and Destruction: Evidence and Price Implications." *American Economic Review* 100 (3): 691–723.

Cavallo, Alberto, W. Erwin Diewert, Robert C. Feenstra, Robert Inklaar, and Marcel P. Timmer. 2018. "Using Online Prices for Measuring Real Consumption across Countries." *AEA Papers and Proceedings* 108: 483–87.

Cavallo, Alberto, Robert C. Feenstra, and Robert Inklaar. 2021. "Product Variety, the Cost of Living and Welfare Across Countries." NBER Working Paper 28711.

Chevalier, Judith A., and Anil K. Kashyap. 2019. "Best Prices: Price Discrimination and Consumer Substitution." *American Economic Journal: Economic Policy* 11 (1): 126–59.

Coibion, Olivier, Yuriy Gorodnichenko, and Gee Hee Hong. 2015. "The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications." *American Economic Review* 105 (3): 993–1029.

Costa, Dora L. 2001. "Estimating Real Income in the United States from 1888 to 1994: Correcting CPI Bias Using Engel Curves." *Journal of Political Economy* 109 (6): 1288–1310.

Deaton, Angus. 2010. "Price Indexes, Inequality, and the Measurement of World Poverty." *American Economic Review* 100 (1): 5–34.

Deaton, Angus, and Alan Heston. 2010. "Understanding PPPs and PPP-Based National Accounts." *American Economic Journal: Macroeconomics* 2 (4): 1–35.

Feenstra, Robert C. 1994. "New Product Varieties and the Measurement of International Prices." *American Economic Review* 84 (1): 157–77.

Feenstra, Robert C., Mingzhi Xu, and Alexis Antoniades. 2020. "What is the Price of Tea in China? Goods Prices and Availability in Chinese Cities." *Economic Journal* 130 (632): 2438–67.

Hamilton, Bruce W. 2001. "Using Engel's Law to Estimate CPI Bias." *American Economic Review* 91 (3): 619–30.

Handbury, Jessie, and David E. Weinstein. 2015. "Goods Prices and Availability in Cities." *Review of Economic Studies* 82 (1): 258–96.

Nakamura, Emi, Jón Steinsson, and Miao Liu. 2016."Are Chinese Growth and Inflation Too Smooth? Evidence from Engel Curves." *American Economic Journal: Macroeconomics* 8 (3): 113–44.

Nordhaus, William D. 1996. "Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not." In *Economics of New Goods*, edited by Timothy F. Bresnahan and Robert J. Gordon, 27–70. Chicago, IL: University of Chicago Press.

Simonovska, Ina. 2015. "Income Differences and Prices of Tradables: Insights from an Online Retailer." *Review of Economic Studies* 82 (4): 1612–56.