Measuring the Cost of Living in Mexico and the US*

David Argente†
Pennsylvania State University

Chang-Tai Hsieh‡
University of Chicago

Munseob Lee§
University of California San Diego

July 2020

Abstract

Cross-country price indexes are crucial to compare living standards between countries and to measure global inequality. An accurate measurement of these price indexes has proven to be a difficult task because of the lack of accurate data on the consumption patterns of different countries. In this paper, we construct a unique data on prices and quantities for consumer packaged goods matched at the barcode-level across two countries, United States and Mexico. Using a non-homothetic price index we estimate that the Mexican real consumption is slightly larger relative to the United States than previously estimated. We identify heterogeneity in shopping behavior, quality of products and variety availability as important sources of bias in international price comparisons.

JEL Classification Numbers: E01, E31, O47
Keywords: Price Index, Purchasing Power Parity, Non-homothetic Demand

---

*We want to thank Esteban Rossi-Hansberg and Ina Simonovska for their comments and suggestions. Results are calculated based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen 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.

†Email: dargente@psu.edu. Address: 403 Kern Building, University Park, PA 16801.
‡Email: chsieh@chicagobooth.edu. Address: 5807 South Woodlawn Ave., Chicago, IL 60637.
§Email: munseoblee@ucsd.edu. Address: 9500 Gilman Drive #0519, La Jolla, CA 92093-0519.

1 Introduction

Indexes of prices across countries are a vital ingredient in estimates of standards of living and real output across countries. The most commonly used price indexes are those by the International Comparison Program (ICP). The ICP collects prices of more than a thousand specific products in each country, which it then aggregates to estimate price indexes of comparable bundles of product across countries.

There are at least three issues with the price data underlying the ICP's exercise. First, products differ in quality across countries. The ICP's goal is to find products that are "comparable" across countries but it is possible that the ICP matches lower quality items in one country with higher quality items in another country. Second, the ICP collects prices from large retail establishments, but the prices in these establishments may not be representative of prices most consumers pay. Finally, the ICP has no information on the availability of products in different countries, and therefore does not take into account differences in the varieties available in different markets.

There are also two important methodological issues with how the underlying prices are aggregated. First, the ICP aggregates the price of the 1000+ detailed products into a price index of 155 broad product categories ("basic headings"). Since the ICP does not have data on the expenditure share of the individual products, it simply takes the simple average of prices of the products in each basic heading. Second, as pointed out by Deaton and Heston (2010), the ICP's price indexes are best interpreted as an index derived from a model where individuals in all countries have identical homothetic preferences. However, there is abundant evidence that preferences are not homothetic.

In this paper, we use a new data set – the Nielsen Homescan Consumer Panel data for the US and Mexico – to try to improve upon the ICP's measurement exercise. The Nielsen Homescan data for the US tracks the shopping behavior of 40-60 thousand households in the US and 5 thousand households in Mexico. Households in the two countries either use in-home scanners or diaries to record their purchases. The Nielsen Homescan data thus has information on prices and quantities purchased by each household of specific products, identified by a 12-digit universal product code (UPC), for each shopping visit and each store where the purchase took place.

We begin by using the Nielsen data to calculate PPPs following exactly the ICP's methodology. This exercise has the same limitations as the ICP, but yields PPPs that are almost exactly the same as that provided by the ICP. We can then exploit the richness of the Nielsen data to improve upon the ICP's micro-data. First, the Nielsen data identifies products by barcodes. Instead of matching products manually, we use the barcodes to identify more than

5,000 identical products (with the same barcodes) in the two countries. Second, the Nielsen data has information on prices and quantities of all purchases for a representative sample of households in each country. We can thus measure the average price of each product from a representative sample of households (instead of a sample of retail stores). Third, since we observe all the purchases made by the households, we can estimate the importance of products available to Mexican consumers but not to American consumers, and vice versa.

The additional information available in the Nielsen data also allows us improve upon the ICP's methodology. First, since we have detailed demographic information of each household, we can estimate the parameters needed to measure a non-homothetic price index. Second, since the Nielsen data provides the expenditures shares of each product, we can use these shares to aggregate the price of each product into a price-index of each broad category (basic heading).

We measure the importance of these adjustments with a price index derived from a nested CES structure that allows for non-homothetic product demands. This price index adapts Feenstra (1994)'s correction for new goods for the non-homothetic utility function developed by Hanoch (1975) and Comin, Lashkari and Mestieri (2020). When we estimate this price index empirically with the Nielsen data, we find that cost of living in Mexico is 10% lower than suggested by the ICP. We further decompose the 10% difference between our price index and the ICP's price index into the contribution of three terms: sampling, quality, and Engel-curve variety bias.

Sampling bias reflects the effect of collecting prices from a sample of consumers (instead of stores) and the effect of using expenditure shares as weights (instead of using the simple average). Sampling bias lowers the ratio of prices in Mexico vs. the US by 18% compared to the ICP.

Quality bias quantifies the effect of matching products across countries with the same barcode instead of products that are roughly comparable as done by the ICP. When we compare products with the same barcodes in the two countries, the price gap between Mexico and the US is smaller compared to the gap for "comparable" products in the two countries. Specifically we find that the price gap between products in Mexico vs the US is 45% *higher* relative to the ICP.

Engel-Curve and Variety bias captures the effect of products available in Mexico but not the US, and products sold in the US and not in Mexico, adjusted for non-homethetic preferences. We find that Mexican varieties missing in the US market matter more than US varieties not sold in Mexico, and correcting this lowers the cost of living in Mexico relative to the US by 25%.

This paper is closely related to recent work that use alternative micro-data to estimate price indexes across countries. Specifically Cavallo, Diewert, Feenstra, Inklaar and Timmer (2018), Cavallo, Feenstra and Inklaar (2020) and Simonovska (2015) use online data and Feenstra, Xu and Antoniades (2019) use scanner data of toothpaste, laundry detergent, personal wash items and shampoo to estimate PPPs across countries. The main advantage of our data is that it contains detailed data on *all* nondurables from a representative sample of *consumers*. We can therefore go beyond the ICP's exercise and measure the effect of sampling, quality, and variety for all nondurable goods in Mexico and the US.

The rest of the paper is organized as follows. Section 2 presents the description of the data. In Section 3 we validate our linked data replicating the purchasing power parities produced by the ICP. In Section 4 we present three stylized facts about the prices in Mexico and the US. In Section 5 we develop our price index and the definition of the three biases we quantify. Section 6 presents our results. In Section 7, we discuss the potential implications of our findings and conclude.

2 Data Description

2.1 International Comparison Program (ICP)

We use the restricted ICP micro-data for Mexico and the United States from the 2011 ICP. The ICP is a worldwide statistical initiative led by the United Nations and the World Bank that collects and compares price data and GDP expenditures to estimate purchasing power parities (PPPs). The data is collected every 6 years through partnerships with statistical agencies in each country. The publicly available data contain information of 155 basic headings (e.g. "Appliances, articles, and products of personal care") that cover all the components of GDP. Approximately, 53 basic headings refer to goods of which 33 are non-durable goods. The restricted ICP data also includes information of the "items" (e.g. "Shampoo", "Razors", "Diapers") within each basic heading and contain the average national price of more than one thousand individual items.

In addition, the data contains the national accounts expenditures for each basic heading. Given that expenditure weights are not available at the item level, the data provide a classification of household consumption items as important or less-important based on national consumption patterns. In the 2011 ICP, importance is defined by reference to the expenditure share of the item within a basic heading. Products that are identified as important by a country were given more weight in calculating PPPs. Nonetheless, not all items are priced or available in each country; the number of items to be priced in each basic heading depends on the heterogeneity of the basic heading, the degree of overlap of products across countries, and the overlap of products each country labels as important to its economy.

After an item is selected, each country determines the sampling framework to sample its prices (e.g. the number of outlets, their type, and their location). The sample design must be such that the national annual average price of each item emerges from the data collection. An outlet is selected to be sampled based on its volume of sales and not on the volume of sales of individual products; because of cost considerations, price collectors often sample what is available in the outlet once they are there.

The ICP provides specifications about the products to be sampled for each item by the price collectors. These specifications include: 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, for the item "Baby Diapers", the price collector must sample products belonging to a well-known brand, containing between 18-24 pieces, either classic or basic type, with a size between 4 and 9.5 kg, and with a multi-pack package. Price collectors visit a sample of outlets, identifying products in the outlets that matched the product specifications on the product list and recording their prices. Although, some specifications are tightly defined, others may leave latitude to the price collector to determine which product to sample.

2.2 Nielsen Mexico

The Nielsen data for Mexico tracks the shopping behavior of 6,000 households for the years 2012-2013. The household sample is updated annually to be representative of all cities over 50,000 people and covers 55 cities in Mexico. Instead of using in-home scanners, households are visited biweekly to obtain complete consumption diary information about all products they purchased. Just as in the US, a UPC identifies each product. The data contain around 55,000 distinct products grouped into 100 product categories defined by Nielsen ranging from food to beauty aids. The categories cover 35-40 percent of all expenditures on goods in the Mexican CPI.

The data contain detailed information of each shopping trip (e.g. date, store, amount spent), transaction level information for each product purchased (e.g. quantity, price, deals, coupons), as well as detailed product level characteristics (e.g., brand, size, packaging, flavor) so that unit values can be computed. The data also include demographic variables at the household level such as the occupation of the household members, education, age, and family size.

Furthermore, households are classified in 7 socioeconomic levels using the index developed by the Mexican Association of Marketing Research and Public Opinion Agencies (AMAI), which is the agency responsible for maintaining the transparency and quality in market research in Mexico. AMAI provides a standardize criteria to define socioeconomic status (SES) in Mexico using demographic information such at the household level (i.e. income, education of the head of the households, number of employed members) as well as dwelling characteristics such as the number of bedrooms, number of bathrooms, and internet availability. In the Appendix C.2 we provide a more detailed description of each of the levels.

2.3 Nielsen US

The Nielsen Homescan tracks the shopping behavior of 40,000 to 60,000 households every year in 48 contiguous states plus Washington D.C. Each panelist uses in-home scanners to record their purchases. A 12-digit universal product code (UPC) identifies the items the panelists purchase. The data contain a few million distinct UPCs grouped into 1,235 product modules (product categories defined by Nielsen) that range from food to beauty aids to computer software. Our data cover around 40% of all of the expenditures on goods in the CPI.[1]

For each UPC, the data contain information on the brand, size, packaging, and a rich set of product features. If the panelist purchases the good at a store covered by Nielsen, the price is set automatically to the average price of the good at the store during the week when the purchase was made. If not, the panelist directly enters the price. Nielsen reports detailed transaction information for each product purchased (e.g., UPC code, quantity, price, deals, and coupons). We combine this information with the weight and volume of the product to compute unit values.

The data also contain information about each purchasing trip the panelist makes, such as the retailer, the location, and the date of the transaction. Further, 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 panelists. The surveys are sent in Q4 of each year and the variables are implemented in the first week of January of the following year. Nielsen provided 16 income bins top-coded at $100,000 up to 2005. After 2006, it has provided 20 income bins top-coded at $200,000. Nielsen asks panelists to report their combined total household annual income as of year-end of the previous calendar year. Nielsen

[1] Nielsen offers a variety of incentives to join and stay active such as monthly prize drawings, gift points, and regular sweepstakes. The incentives are designed to be non-biasing (i.e., Nielsen does not provide account-specific coupons out of concern for the potential effect on the natural purchase selection of outlets and products).

believes panelists are actually reporting their annualized estimated income as of the time of the survey and not referring to the previous years tax returns. Self-reported annual income is likely to be the total labor income. Nielsen constructs projection weights that make the sample representative of the US population that we use in all our calculations.[2]

2.4 Matched Data

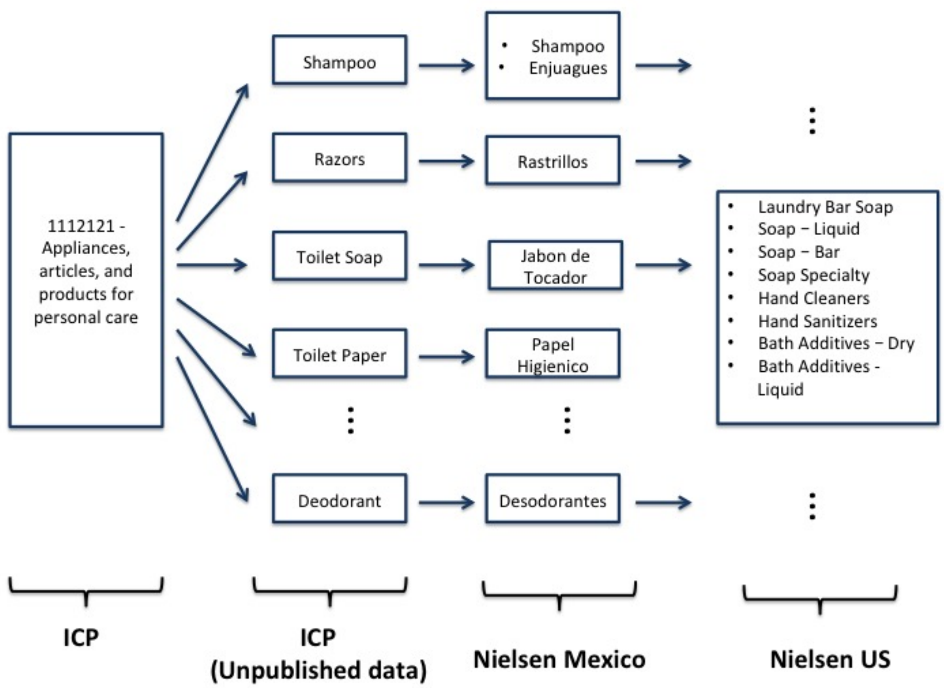
We manually link the ICP items to barcodes in the Nielsen data of both Mexico and the US. To do so, we first match the item (e.g. Diapers) to the closest category of barcodes in each of the Nielsen data sets. We then follow the specifications provided by the ICP to the price collectors to select the relevant barcodes for each item (e.g. well-known brand, between 18-24 pieces, classic or basie, medium, multipack, pull-ups). Importantly, given that several barcodes meet the specification of a given item, there are many barcodes in both countries matched to a given item. We focus on items available in both Mexico and the US. Since the categories in the Mexico Nielsen data are defined more broadly than those in the US Nielsen data, our linked data includes 267 product categories of the US Nielsen data and 47 categories of the Mexican data. The barcodes in these categories categories that meet the specifications of the items represent all items that are common across the two countries within 18 basic headings. Overall, our linked data represents 11% of GDP in Mexico and 5% in the US and accounts for 60% of non-durable goods expenditures in Mexico and 65% in the US; it is primarily for non-durable consumption that prices differ between poor and rich countries, other prices (e.g. capital goods, exports and imports) do not differ systematically across countries (Hsieh and Klenow, 2007). The non-durable consumption categories that we are unable to match are mainly those with products without barcodes such as poultry, fresh fish, and fresh fruit.

Furthermore, we match barcodes across Mexico and the US taking advantage of the fact that both the US and Mexico adopted the UPC system as their preferred standard. To obtain a UPC code, firms in both countries must first obtain a Global Standards One (GS1) company prefix. GS1 is the single official source of UPC codes and it has member organizations in over 100 countries. The company prefix begins with a two to three digit number that identifies the country where the barcode was issued and continues with five to ten digits that identify the firm and its products uniquely all over the world. The universal compatibility of the UPC system allows the movement of products across countries without requiring a different barcode in each country. GS1 US and GS1 Mexico issue authorized GS1

[2]Nielsen has a comprehensive program of dropping and replacing panelists that do not perform to minimum reporting standards. Currently, Nielsen retains about 80% of its active panel each year. Nielsen uses a stratified sampling design to ensure that the panel is demographically balanced.

barcodes for businesses in the US and Mexico beginning with 00-139 and 750 respectively.[3] In our data, more than that 5,000 barcodes are consumed by households both in the US and Mexico. Approximately, 80% of the matched barcodes have prefixes authorized by GS1 US. Figure 1 illustrates the structured of the matched data for the ICP item toilet soap. The item is first matched with "Jabon de Tocador" in the Mexican consumer panel data, then with multiple product modules in US consumer panel data, such as "Soap- Liquid," "Soap-Bar," and "Soap Specialty".

Figure 1: Structure of Matched Data



3 Replicating the ICP

In order to validate our linked data set, we use it to construct a price index that mimics the procedures followed by the ICP 2011. We then compare this index to the one published by the ICP. We begin by selecting a single representative product for a given item and store that

[3]The first digits of the GS1 prefixes identify only where the barcode was issued and not the country of origin for a given product. The prefixes simply provide number capacity to different countries for assignment from that location to companies who apply. Those companies in turn may manufacture products anywhere in the world. For example, if a Mexican company imports an item from a different country, then packaged and shipped that item to the US, the country code in the GS1 prefix would likely correspond to Mexico.

meets the specification defined by the ICP. Since it is not possible for us to know exactly the product that the price collector selected for each item, we choose the product that meets the specifications provided by the ICP with the largest volume of sales in each store.[4] We compute the average price of this product over our sample period. Since the ICP considers a single price observation for each item in each country, we average the prices of products to the item level using store-specific volume weights as it is done by the ICP.

Next, we follow the ICP and run a Country Product Dummy (CPD) regression for every basic heading as follows:

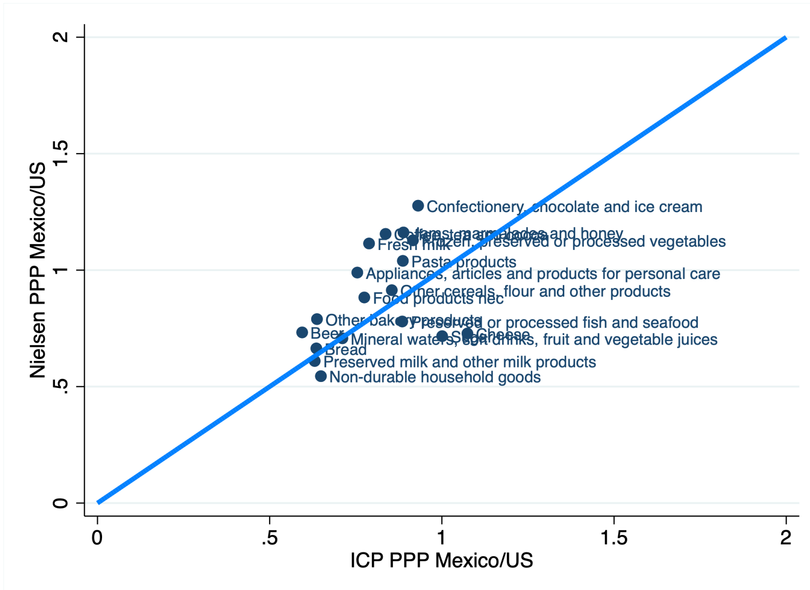
$$\log \bar{p}^c_{ib} = \eta^c_i + \eta^c_b + \epsilon^c_{ib} \tag{1}$$

where $\bar{p}^c_{ib}$ is the price of item i belonging to basic heading b in country c regressed on "item-country" fixed effects, η^c_i, and "basic heading-country" fixed effects, η^c_b. The estimated PPP, expressed as national currency per unit of the base country (in our case the US dollar), for a particular basic heading is $\bar{p}^c_b = \exp(\eta^c_b)$. Since all items in each basic heading are priced in both countries, the $\bar{p}^c_b$ is a form of a Jevons index.

In principle, there are many reasons to expect differences between the Nielsen-ICP index created using scanner data and the one produced by the ICP. First, our prices are collected directly from household panels and the ICP prices are collected through price surveys. Second, prices in the Nielsen data are collected in each transaction, while ICP prices are obtained once (or a few times) per year. Temporal aggregation could affect the comparison because PPPs can vary significantly within a year. Third, the CPD regressions by the ICP often include items that are not present in every country since the CPD model can be used with incomplete data as long as the price data is connected across countries. Our Nielsen-ICP is computed using complete information of the items present in the US and Mexico. Despite these differences, Figure 2 shows that Nielsen-ICP price level indexes computed with scanner data (adjusted by the Mexican peso nominal exchange rate with the US dollar) align well with those calculated by the ICP.

[4] Countries are asked to classify items as important or less important. If expenditure shares were available at the product level and they are thought to be large, the item is classified as important. However, since the basic headings are in fact defined as the most detailed level of expenditures for which countries are asked to supply expenditure shares, countries classify a product as important if it is in the CPI or if a local experts determine the product is important.

Figure 2: Nielsen vs ICP PPP



Note: The figure plots the ICP for Mexico and the United States against the Nielsen-ICP, which is an index generated using the same methodology as the ICP but using the Nielsen data.

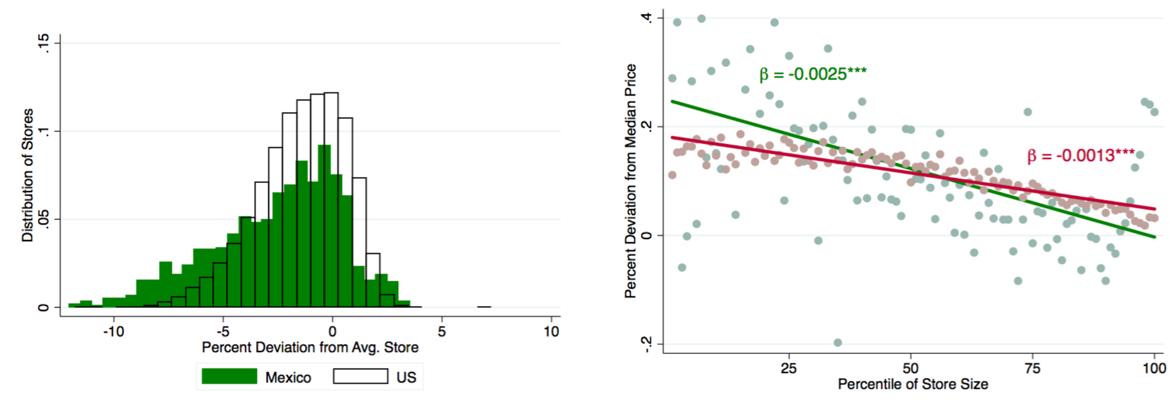
4 Stylized Facts

Fact 1: The size distribution of stores is more left skewed in Mexico. The prices set by smaller stores are higher on average.

We begin by comparing the size distribution of stores across the two countries. We approximate the size of the stores with the total revenue generated by the store in our sample period and across the match categories. Panel (a) in Figure 3 shows the size distribution of stores in the two countries. The figure shows the distribution of stores in each country relative to the size of the store, where for ease of comparison we plotted the size of the store as a percent deviation from the average store in each country. The figure shows that in both countries there are a few stores that generate very high revenue. In Mexico, however, the distribution has sizeable left tail indicating the presences of a large number of low revenue stores.

We then explore whether the prices set by the stores are correlated with their size. To do this, we first calculate the median price for each item across all stores in each country. Then, for each item and store, we calculate the percent deviation from the median price and average it across items to the store level. Panel (b) plots the percent deviation from the median price for each store against the percentiles of the size distribution of stores in each country. The figure shows that smaller stores tend to set higher prices in both countries, but particularly in Mexico. This is relevant since it indicates that sampling prices from large stores could lead to important biases, particularly in the case of Mexico where there size distribution of stores is more left-skewed and where the prices of smaller stores tend to be higher than those set in larger stores.

Figure 3: Size Distribution of Stores and Prices by Store Size



(a) Size Distribution (b) Prices and Store Size

Note: Panel (a) shows the size distribution of stores for Mexico and the US. Size is approximated using the total revenue of the stores. The x-axis is the percent deviation from the average store in each country. Panel (b) shows the relationship between prices and store size. The x-axis shows the percentiles of the store size distribution in 100 bins. The y-axis is the average deviation from the median price for each store. We calculate the median price for each item across all stores in each country. Then, for each item and store, we calculate the percent deviation from the median price and average it across items to the store level.

Fact 2: The distribution of prices in the US has a higher mean and a longer right tail, but these patterns are attenuated when we compare common goods.

Next, we compare the distribution of prices across the two countries within an item. Table 1 considers the categories that are common in Mexico and the US and shows statistics related to the shape of the distribution of prices in each country such as the 75%/25% ratio, the 90%/10%, the mean/50%, and the Kelly skewness. The statistics are computed within a category of goods as defined by Nielsen Mexico. The table contains information of all the barcodes belonging to the categories that are matched across the two countries and also reports the statistics for the set of common barcodes (matched sample). The table shows that the dispersion of prices is relatively similar across the two countries. If we consider all available barcodes in the two countries the dispersion in the distribution of prices is very similar; the 75%/25% ratio is 2.27 in the US and 2.13 in Mexico. Likewise, for the set of identical barcodes across the two countries, the 75%/25% ratio is between 1.96 and 2.16. Not surprisingly, for both countries and regardless of the sample, the median is lower than the average price indicating that the distribution of prices is right skewed. The last column of the table shows that the Kelly skewness of the distribution of prices in the US is higher than that in Mexico when all the products are considered. This indicates that in the US there is a longer right tail in the distribution of prices. When we consider only the set of barcodes that are common across the two countries both the ratio of the mean and the median and the Kelly skewness are very similar across countries.

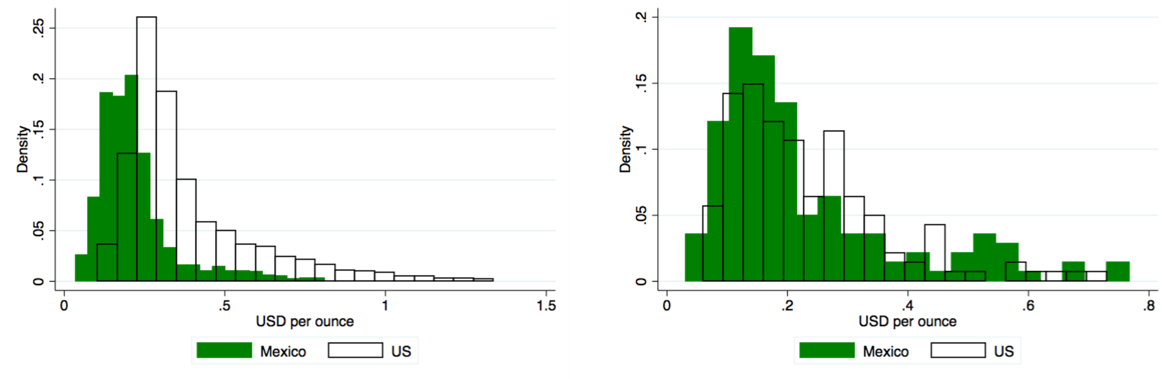
Table 1: Distribution of Prices

The table reports statistics related to the shape of the distribution of prices in Mexico and the US such as 75%/25% ratio, mean/50% ratio, and the Kelly skewness. It considers two samples: the set of all products available in the two countries within the common categories and the set of common products matched at the barcode level.

Country	Sample	75%/25%	Avg/50%	Kelly Skewness
US	All	2.27	1.55	0.48
Mexico	All	2.13	1.39	0.42
US	Matched	1.96	1.25	0.35
Mexico	Matched	2.16	1.39	0.34

An example of these patterns can be seen in the category "Cheese". Panel (a) in figure Figure 4 shows distribution of product prices in the US and Mexico. On average, products in the US are more expensive, and variance of price distribution is also high. This pattern changes when we compare prices of common products that are available in both countries. Panel (b) shows that for common products distributions of product prices look similar between two countries. These patterns could lead to a large quality bias in the construction of international price indexes. This is because, when comparing items that are roughly comparable instead of those that are exactly the same, statistical agencies are more likely to sample high price items from the US distribution than from the distribution of prices in Mexico.

Figure 4: Product Price Distribution for All and Common Products: Example of Cheese



(a) All Products (b) Common Products

Note: Panel (a) shows price distribution of products in cheese category for all products in the US and Mexico. Panel (b) describes price distribution of common products that are available in both countries.

Fact 3: Within category of products, richer households buy more expensive products.

Lastly, we explore the relationship between prices and income in the cross-section. This relationship is also known as the quality Engel curve. The curve is computed under the premise that across households, at a point in time, those paying higher unit prices are buying higher quality goods. If high-income households purchase different, more expensive, products within a category than low-income households we expect a positive relationship between prices and socioeconomic status. This positive relationship has been documented for US households and has been used to check whether the assumption of homothetic tastes holds in the data.[5]

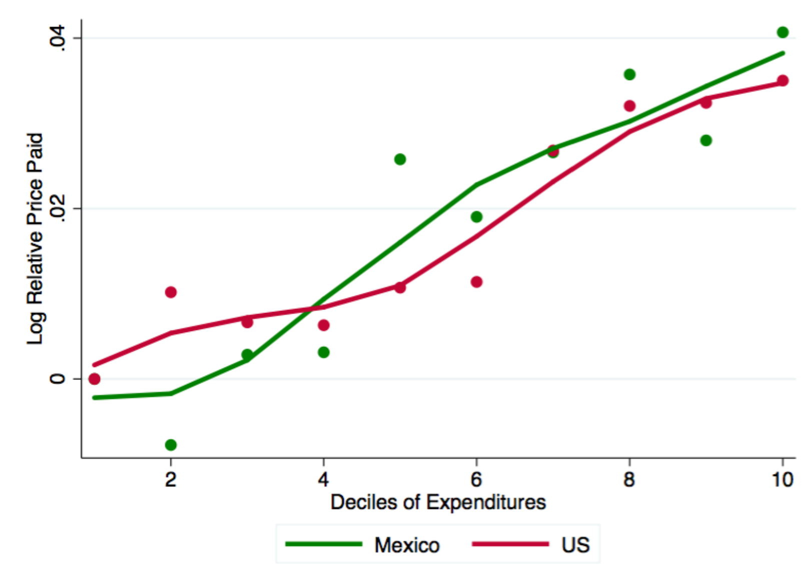
Although the US Nielsen data provide information of each household in income bins, the Mexican Nielsen data only reports the socio-economic level index for each households.[6] To approximate the permanent income level of the households in a comparable way across the two countries, we approximate income with the total level of expenditures of the households

[5] See Bils and Klenow (2001) for a detailed description and Broda, Leibtag and Weinstein (2009), Handbury (2019), Argente and Lee (forthcoming) and Faber (2014) for other applications.

[6] The socio-economic level index is a standardized index based on a statistical model that allows grouping households into seven levels considering the level of education of the head of the household, the people age 14+ who are employed, and other characteristics of the households such as the number of bedrooms, the number of bathrooms, number of, people age 14+- who are employed, number of vehicles, has access to internet. In Appendix C.2, we show the quality Engel Curve using this measure as a proxy of the income of the households.

over our sample period. Figure 5 plots the quality Engel curves for the US and Mexico. The figure shows how much more or less households pay per unit for products within a category. We considered the matched categories across the two data sets. In the figure, the relative prices are measured in a regression of the log unit price paid against each of the dummies for each of decile of the household expenditure distribution of each country and category, region, store, and quarter fixed effects. Each dot represents how much more high income households pay per unit for products within a category than low income households. Within a product category, high income households pay approximately 4 percent higher price than low income households. A substantial amount of difference comes from the fact that high income households buy higher quality products.

Figure 5: Quality Engel Curve - Mexico and the US



The figure plots the cross-sectional relationship between the relative prices paid and household income. The relative prices are measured in a regression of the log unit price paid against dummies for each decile of the household expenditures distribution of each country and category, region, store, and quarter fixed effects. Each dot represents how much more high income households pay per unit for products with respect to low income households.

5 Theoretical Framework

In this section, we introduce a non-homothetic CES framework and derive a price index, which has the homothetic exact price index developed by Feenstra (1994) as a special case. Then, we compare our price index to that of the ICP and provide an exact decomposition framework to quantify several potential biases in the purchasing power parties produced by the ICP.

5.1 Non-homothetic CES Preference

We use a multi-tiered CES aggregator to specify the utility function of a representative consumer in a country M (Mexico) and U (United States). Without loss of generality, we introduce the utility function in one country, Mexico. Later, we define a price index for Mexico having the United States as a numeraire.

The highest level of aggregation in the utility function is:

$$\mathbb{U}^M = \left(\sum_{b \in \Omega} \left(C_b^M \right)^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}} \tag{2}$$

where the basic headings are indexed by b, γ is the elasticity of substitution across basic headings, and Ω is the set of all basic headings common across the two countries. We model the next tier as:

$$C_b^M = \left(\sum_{i \in \Omega_b} \left(C_{ib}^M \right)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \tag{3}$$

where C_{ib}^M is the total quantity of item i that is consumed in basic heading b in Mexico M. η is the elasticity of substitution across items and Ω_b is the set of all items within a basic heading b that are common across countries. Both basic headings and items are common across countries in the ICP methodology.[7]

Based on the stylized facts documented in the previous section, we now introduce non-homotheticities using the non-separable class of CES functions in Sato (1975), Comin, Lashkari and Mestieri (2020), Matsuyama (2019) and Redding and Weinstein (2020), which satisfy implicit additivity in Hanoch (1975). The non-homothetic CES consumption index for item i, C_{ib}^M, is defined by the following implicit function:

$$\sum_{k \in \Omega_{ib}^M} \left(\frac{\varphi_{kib}^M C_{kib}^M}{\left(C_{ib}^M \right)^{(\epsilon_{kib} - \sigma_{ib})/(1-\sigma_{ib})}} \right)^{\frac{\sigma_{ib}-1}{\sigma_{ib}}} = 1 \tag{4}$$

where C_{kib}^M denotes total consumption of barcode k; Ω_{ib}^M is the set of barcodes in Mexico M within an item i and a basic heading b; σ_{ib} is the elasticity of substitution between barcodes;

[7] These first two layers of the multi-tiered CES aggregator provide theoretical guidance on how to aggregate the prices of items and basic headings when comparing the prices of two countries. In the ICP methodology, since they cannot favor a Paasche index or a Laspeyres index to aggregate across basic headings, they compute a Fisher index averaging over multiple indexes.

ϵ_{kib} is the constant elasticity of consumption of barcode k with respect to the consumption index (C_{ib}^M) which controls the income elasticity of demand for that barcode. Assuming that barcodes are substitutes ($\sigma_{ib} > 1$), it is required $\epsilon_{kib} < \sigma_{ib}$ for the consumption index to be globally monotonically increasing and quasi-concave, and therefore to correspond to a well-defined utility function. When $\epsilon_{kib} = 1$ for all $k \in \Omega_{ib}^M$, the utility function becomes homothetic. In Appendix A, we show that this non-homothetic CES specification is a special case of a nested logit model whose second-stage is deterministic, as shown for the homothetic CES case by Anderson, De Palma and Thisse (1987).

Lastly, the lowest tier aggregates purchases of a barcode from multiple stores into the total consumption of the barcode k in a country, C_{kib}^M:

$$C_{kib}^M = \prod_{s \in \Psi^M} \left(C_{skib}^M\right)^{\phi_{skib}^M} \tag{5}$$

where $\sum_{s \in \Psi^M} \phi_{skib}^M = 1$ for basic heading b, item i, barcode k, and store s. This differentiate the same barcode if it is sold in different stores allowing the differences of local retail services to play a role in the determination of prices.

5.2 Demand System and Price Index

We solve the expenditure minimization problem for a given barcode within an item and basic heading to obtain the following expressions for the price index (P_{ib}^M) dual to the consumption index (C_{kib}) and the expenditure share for a individual barcode k (s_{kib}^M):

$$P_{ib}^M = \left(\sum_{k \in \Omega_{ib}^M} (p_{kib}^M / \varphi_{kib}^M)^{1-\sigma_{ib}} (C_{ib}^M)^{\epsilon_{kib}-1} \right)^{\frac{1}{1-\sigma_{ib}}} \tag{6}$$

$$s_{kib}^M = \frac{(p_{kib}^M / \varphi_{kib}^M)^{1-\sigma_{ib}} (C_{ib}^M)^{\epsilon_{kib}-1}}{\sum_{l \in \Omega_{ib}^M} (p_{lib}^M / \varphi_{lib}^M)^{1-\sigma_{ib}} (C_{lib}^M)^{\epsilon_{lib}-1}} = \frac{(p_{kib}^M / \varphi_{kib}^M)^{1-\sigma_{ib}} (E_{ib}^M / P_{ib}^M)^{\epsilon_{kib}-1}}{(P_{ib}^M)^{1-\sigma_{ib}}} \tag{7}$$

Taking ratios of the shares of Mexico and the United States and rearranging , we obtain the following expression for the difference in the cost of living, which holds for each common barcode available in two countries (Ω_{ib}):

$$\frac{P_{ib}^M}{P_{ib}^U} = \frac{p_{kib}^M / \varphi_{kib}^M}{p_{kib}^U / \varphi_{kib}^U} \left(\frac{E_{ib}^M / P_{ib}^M}{E_{ib}^U / P_{ib}^U} \right)^{\frac{\epsilon_{kib}-1}{1-\sigma_{ib}}} \left(\frac{s_{kib}^M}{s_{kib}^U} \right)^{\frac{1}{\sigma_{ib}-1}}, \quad k \in \Omega_{ib} \tag{8}$$

Summing expenditures across common barcodes, we obtain the following expression for the aggregate share of common barcodes in total expenditure for Mexico and the United

States (λ_{ib}^M and λ_{ib}^U):

$$\lambda_{ib}^M \equiv \frac{\sum_{\Omega_{ib}} p_{kib}^M C_{kib}^M}{\sum_{\Omega_{ib}^M} p_{kib}^M C_{kib}^M} \quad \text{and} \quad \lambda_{ib}^U \equiv \frac{\sum_{\Omega_{ib}} p_{kib}^U C_{kib}^U}{\sum_{\Omega_{ib}^U} p_{kib}^U C_{kib}^U} \tag{9}$$

where Ω_{ib} is the set of common barcodes between the two countries and Ω_{ib}^M and Ω_{ib}^U represent the set of total barcodes in Mexico and the United States respectively.

Using this expression, the share of an individual barcode in total expenditure (s_{kib}^M) in equation (7) can be re-written as its share of expenditure on common barcodes (s_{kibt}^{M*}) times this aggregate share of common barcodes in total expenditure (λ_{ib}^M):

$$s_{kib}^M = \lambda_{ib}^M s_{kib}^{M*}, \quad k \in \Omega_{ib} \tag{10}$$

Taking logs to equation (8) and using equation (10), we obtain the following equation:

$$\log\left(\frac{P_{ib}^M}{P_{ib}^U}\right)^{1+\frac{\epsilon_{kib}-1}{1-\sigma_{ib}}} = \log\frac{p_{kib}^M/\varphi_{kib}^M}{p_{kib}^U/\varphi_{kib}^U} + \log\left(\frac{E_{ib}^M}{E_{ib}^U}\right)^{\frac{\epsilon_{kib}-1}{1-\sigma_{ib}}} + \log\left(\frac{s_{kib}^{M*}}{s_{kib}^{U*}}\right)^{\frac{1}{\sigma_{ib}-1}} + \log\left(\frac{\lambda_{ib}^M}{\lambda_{ib}^U}\right)^{\frac{1}{\sigma_{ib}-1}} \tag{11}$$

We define the ideal log-difference weights (ω_{kib}), the logarithmic mean of common variety expenditure shares, as follows:

$$\omega_{kib}^M = \frac{\dfrac{s_{kib}^{M*} - s_{kib}^{U*}}{\ln s_{kib}^{M*} - \ln s_{kib}^{U*}}}{\sum_{k\in\Omega_{ib}} \dfrac{s_{kib}^{M*} - s_{kib}^{U*}}{\ln s_{kib}^{M*} - \ln s_{kib}^{U*}}} \tag{12}$$

where

$$s_{kib}^{M*} = \frac{p_{kib}^M C_{kib}^M}{\sum_{k\in\Omega_{ib}} p_{kib}^M C_{kib}^M} \quad \text{and} \quad s_{kib}^{U*} = \frac{p_{kib}^U C_{kib}^U}{\sum_{k\in\Omega_{ib}} p_{kib}^U C_{kib}^U}$$

We introduce an assumption that tastes are the same between two countries for each common barcode ($\varphi_{kib}^M = \varphi_{kib}^U$ for all $k \in \Omega_{ib}$).

By multiplying the ideal log-difference weights (ω_{kib}) to equation (11), under the assumption that tastes are the same between two countries for each common barcode ($\varphi_{kib}^M = \varphi_{kib}^U$ for all $k \in \Omega_{ib}$), we can obtain the non-homothetic CES price index for item i by taking the arithmetic mean across common barcodes (Ω_{it}) and exponents on both sides:

$$\mathbb{NH}_{ib}^M \equiv \frac{P_{ib}^M}{P_{ib}^U} = \left(\prod_{k\in\Omega_{ib}} \left(\frac{p_{kib}^M}{p_{kib}^U}\right)^{\omega_{kib}} \times \left(\frac{\lambda_{ib}^M}{\lambda_{ib}^U}\right)^{\frac{1}{\sigma_{ib}-1}}\right)^{\frac{1}{1-\theta_{ib}}} \left(\frac{E_{ib}^M}{E_{ib}^U}\right)^{\frac{\theta_{ib}}{\theta_{ib}-1}} \tag{13}$$

where

$$\theta_{ib} \equiv \sum_{k \in \Omega_{ib}} \omega_{kib} \frac{\epsilon_{kib} - 1}{\sigma_{ib} - 1}$$

and the ratio of λ_{ib}^{M} and λ_{ib}^{U} represents the conventional variety correction term that accounts for different sets of goods available in two countries as in Feenstra (1994) and Broda and Weinstein (2006, 2010). $\mathbb{NH}_{ib}^{M}$ can be aggregated to the basic heading b level ($\mathbb{NH}_{b}^{M}$) and across basic headings by taking the weighted geometric mean and using the ideal log-difference weights from our CES specification defined as:

$$\omega_{ib} = \frac{\frac{s_{ib}^{M} - s_{ib}^{U}}{\ln s_{ib}^{M} - \ln s_{ib}^{U}}}{\sum_{i \in \Omega_b} \frac{s_{ib}^{M} - s_{ib}^{U}}{\ln s_{ib}^{M} - \ln s_{ib}^{U}}} \quad \text{and} \quad \omega_{b} = \frac{\frac{s_{b}^{M} - s_{b}^{U}}{\ln s_{b}^{M} - \ln s_{b}^{U}}}{\sum_{b \in \Omega} \frac{s_{b}^{M} - s_{b}^{U}}{\ln s_{b}^{M} - \ln s_{b}^{U}}}$$

Homothetic Case: The CES homothetic case is a special case of our non-homothetic CES framework. When $\epsilon_{kib} = 1$ for all $k \in \Omega_{kib}^{M}$, the total consumption of an item i in a basic heading b (C_{ib}^{M}) is the following:

$$C_{ib}^{M} = \left(\sum_{k \in \Omega_{ib}^{M}} \left(\varphi_{kib}^{M} C_{kib}^{M} \right)^{\frac{\sigma_{ib}-1}{\sigma_{ib}}} \right)^{\frac{\sigma_{ib}}{\sigma_{ib}-1}} \tag{14}$$

In this case the Exact Price Index is the same as the one developed by Feenstra (1994) and Broda and Weinstein (2006, 2010):

$$\mathbb{EPI}_{ib}^{M} \equiv \frac{P_{ib}^{M}}{P_{ib}^{U}} = \prod_{k \in \Omega_{ib}} \left(\frac{p_{kib}^{M}}{p_{kib}^{U}} \right)^{\omega_{kib}} \times \left(\frac{\lambda_{ib}^{M}}{\lambda_{ib}^{U}} \right)^{\frac{1}{\sigma_{ib}-1}} \tag{15}$$

Relationship between the non-homothetic and homothetic price indexes: The non-homothetic CES price index for item i ($\mathbb{NH}_{ib}^{M}$, equation 13) can be written as a function of the Exact Price Index ($\mathbb{EPI}_{ib}^{M}$, equation 15):

$$\begin{aligned} \mathbb{NH}_{ib}^{M} &= \left(\prod_{k \in \Omega_{ib}} \left(\frac{p_{kib}^{M}}{p_{kib}^{U}} \right)^{\omega_{kib}} \times \left(\frac{\lambda_{ib}^{M}}{\lambda_{ib}^{U}} \right)^{\frac{1}{\sigma_{ib}-1}} \right)^{\frac{1}{1-\theta_{ib}}} \left(\frac{E_{ib}^{M}}{E_{ib}^{U}} \right)^{\frac{\theta_{ib}}{\theta_{ib}-1}} \\ &= \mathbb{EPI}_{ib}^{M} \times \left(\frac{E_{ib}^{M}/E_{ib}^{U}}{\mathbb{EPI}_{ib}^{M}} \right)^{\frac{\theta_{ib}}{\theta_{ib}-1}} \end{aligned} \tag{16}$$

where θ_{ib} is defined as before. The first term is the Exact Price Index and the second is the ratio of real consumption across the two countries. This relationship will be important in the decomposition of the non-homothetic CES price index the we develop in the next section.

5.3 Decomposition of Non-homothetic CES Price Index

In order to compare and quantify the gap between the non-homothetic price index and the one constructed by the ICP, in this subsection we provide an exact decomposition of the non-homothetic price index that maps exactly to the one constructed by the ICP. In addition, we derive expressions for the sampling, quality, and Engel-curve variety biases.

We begin by formalizing the analysis in Section 3 to construct a price index that mimics that of ICP. We follow the methodology by the ICP using the Nielsen Consumer Panel data. First, we select a single representative product for a given item and store, by choosing barcode that meets the specifications provided by the ICP with the largest volume of sales in each store, since this would be an item classified as important within a basic heading. Next, we aggregate the price of this item using store specific weights as done by the ICP.

$$\bar{p}_{ib}^{M} = \prod_{s \in \Psi^{M}} \left(\bar{p}_{sib}^{M}\right)^{\phi_s^M}$$

Note that, due to data limitations, the ICP uses weights based on the total store sales (ϕ_s^M) instead of the sales of individual items in the store (ϕ_{skib}^M) to compute the average national price of a given item.[8] Lastly, we take the geometric average across items within a basic heading using equal weights, since the ICP only has expenditures at basic heading level and not at the item level. This is the result of the Country Product Dummy (CPD) methodology implemented by equation 1. As a result, we define the ICP price index for basic heading b ($\mathbb{ICP}_b^M$) as follows:

$$\mathbb{ICP}_b^M = \left(\prod_{i \in \Omega_b} \frac{\bar{p}_{ib}^M}{\bar{p}_{ib}^U}\right)^{\frac{1}{N_b}} \tag{17}$$

where N_b is the number of items within a basic heading b.

The non-homothetic CES price index at the basic heading level can be written as a function of the price index developed by the ICP. The relationship between the two indexes can be written as:

$$\mathbb{NH}_b^M = \Theta_b^M \times \mathbb{ICP}_b^M \tag{18}$$

$$\Theta_b^M \equiv \mathbb{S}_b^M \times \mathbb{Q}_b^M \times \mathbb{V}_b^M$$

[8]Each country samples representative stores based on a register of all outlets available to consumers and their respective volume of sales.

where Θ_b^M is the aggregate bias correction term and $\mathbb{S}_b^M$, $\mathbb{Q}_b^M$ and $\mathbb{V}_b^M$ represent the sampling bias, the quality bias, and the Engel-curve variety bias, respectively.

Sampling Bias, $\mathbb{S}_b^M$**:** Due to data limitations, the ICP faces two potential biases in its sampling procedure. First, at the item level, the ICP uses an equal-weights geometric average since they lack data on expenditures at the item level. Second, the sample probabilities are proportionate to the store volume of sales, and sales may not reflect the sales of individual items. After a store is selected to be sampled, it is good practice that the price collectors gather the prices of as many items are available in the store to reduce the collection costs even if the sales of an individual item are not well represented in the stores sampled. Therefore, the sampling bias ($\mathbb{S}_b^M$) is the product of these two biases and is defined as follows:

$$\mathbb{S}_b^M \equiv \left(\left(\prod_{i\in\Omega_b} \frac{\bar{p}_{ib}^M}{\bar{p}_{ib}^U} \right)^{\frac{-1}{N_b}} \times \prod_{i\in\Omega_b} \left(\frac{\bar{p}_{ib}^M}{\bar{p}_{ib}^U} \right)^{\omega_{ib}} \right) \times \left(\prod_{i\in\Omega_b} \left(\frac{\hat{p}_{ib}^M/\bar{p}_{ib}^M}{\hat{p}_{ib}^U/\bar{p}_{ib}^M} \right)^{\omega_{ib}} \right) \tag{19}$$

where

$$\bar{p}_{ib}^M = \prod_{s\in\Psi^M} \left(\bar{p}_{sib}^M\right)^{\phi_s^M} \quad \text{and} \quad \hat{p}_{ib}^M \equiv \prod_{s\in\Psi^M} \left(\bar{p}_{sib}^M\right)^{\phi_{sib}^M}$$

where as before $\bar{p}_{ib}^M$ is the item level price computed using the ICP methodology (using ϕ_s^M as weights) and $\hat{p}_{ib}^M$ is computed aggregating prices from multiple stores using weights defined at the item $\times$ store level (ϕ_{sib}^M).

The first component of the sampling bias captures the bias due to the fact that the ICP does not have information of the expenditures of each item. The size of this term, for a country c, depends on the covariance of the expenditure weights of each item and their prices $\text{cov}(\omega_{\mathbf{b}}, \ln(\bar{\mathbf{p}}_{\mathbf{b}}^{\mathbf{c}}))$ where $\omega_{\mathbf{b}}$ is a vector of item-specific expenditure weights in basic heading b and $\ln\left(\bar{\mathbf{p}}_{\mathbf{b}}^{\mathbf{M}}\right)$ is the vector of log prices.[9] The covariance is negative if people spend more on items with lower prices.

The second component of the sampling bias captures the importance of aggregating prices using weights reflecting the importance of each item in each store. For example, a fruit market may also sell milk, but its volume of sales is more reflective of the fruit's sales. This bias is larger than 1 for a country c if $\text{cov}(\phi_{\mathbf{ib}}^{\mathbf{c}}, \ln(\bar{\mathbf{p}}_{\mathbf{ib}}^{\mathbf{c}})) > \text{cov}(\phi^{\mathbf{c}}, (\bar{\mathbf{p}}_{\mathbf{ib}}^{\mathbf{c}}))$ where $\phi_{\mathbf{ib}}^{\mathbf{c}}$ is the vector of expenditure weights that vary by item and basic heading, (i.e. $\{\phi_{sib}^c\}_{s\in\Psi^c}$), $\phi^{\mathbf{c}}$ is the vector of weights that only vary at the store level (i.e. $\{\phi_s^c\}_{s\in\Psi^c}$), and $\ln\left(\bar{\mathbf{p}}_{\mathbf{ib}}^{\mathbf{M}}\right)$ is the vector of log prices (i.e. $\{\ln(\bar{p}_{sib}^c)\}_{s\in\Psi^c}$).[10] Given that we expect $\text{cov}(\phi_{\mathbf{ib}}^{\mathbf{c}}, \ln(\bar{\mathbf{p}}_{\mathbf{ib}}^{\mathbf{c}}))$ to be negative, the second term of the sampling bias, all else equal, will be less than 1 if the consumers in Mexico

[9]Propositin 1 in the Appendix shows formally this relationship.
[10]See Proposition 2.

tend to buy a larger share of their products at stores where they are cheaper relative to the consumers in the US.

Quality Bias, $\mathbb{Q}_b^M$: Lower quality products in Mexico might get matched with higher quality products in the United states due to the fact that the distribution of prices in the US within an item has a longer right tail. This could lead to an understatement of the price levels in Mexico and comes from the fact that price collectors identify products in the outlets that match the product specifications given by the ICP but these specification are often on purpose not very narrow in order to increase the overlap of items across countries. We measure the extent of the quality bias by comparing the prices collected by the ICP (under the aggregation that already corrects for the sampling weights) and the average price of an item after considering all the barcodes available in a given country:

$$\mathbb{Q}_b^M \equiv \prod_{i \in \Omega_b} \left(\left(\frac{\hat{p}_{ib}^M}{\hat{p}_{ib}^U} \right)^{-1} \times \prod_{k \in \Omega_{ih}} \left(\frac{p_{kib}^M}{p_{kib}^U} \right)^{\omega_{kib}} \right)^{\omega_{ib}} \tag{20}$$

The finer level of disaggregation in our data allows us to compare the price differences across the two countries comparing products with the exact same physical attributes. This is because any change in an attribute of a good (e.g. form, size, package, formula) results in a new barcode. Note that the quality bias would be bigger than one if the ICP is matching lower quality items in Mexico to higher quality items in the US.

Engel Curve Variety Bias, $\mathbb{V}_b^M$: This last bias measures both the cross-country differences in availability of products and the differences in real consumption across countries and is defined as follows:

$$\mathbb{V}_b^M = \widetilde{\mathbb{V}}_b^M \times \prod_{i \in \Omega_b} \left[\left(\frac{E_{ib}^M / E_{ib}^U}{\mathbb{EPI}_{ib}^M} \right)^{\frac{\omega_{ib}\theta_{ib}}{\theta_{ib}-1}} \right] \tag{21}$$

where

$$\widetilde{\mathbb{V}}_b^M \equiv \prod_{i \in \Omega_b} \left(\frac{\lambda_{ib}^M}{\lambda_{ib}^U} \right)^{\frac{\omega_{ib}}{\sigma_{ib}-1}} \tag{22}$$

where $\widetilde{\mathbb{V}}_b^M$ is the ratio of the share of common products in Mexico relative to the share of common products in US and is analogous to the one developed by Feenstra (1994). In our context, if this term is less than one, it indicates that Mexican consumers spend a lower share of their expenditures in goods that are common in the two countries.

We called the second term Engel curve adjustment since it captures the differences in real consumption across the two countries and depends on the income elasticity of demand of the common barcodes across the two countries. When not all products are common across countries, this term serves as an adjustment to the standard variety bias since the share of common barcodes across countries naturally depends on their income differences. However, even if all barcodes across the two countries are common, the Engel Curve adjustment corrects the price index for the relative importance of each product as the relative income of the countries change. If the elasticities of consumption of each barcode with respect to the consumption index equal to one, we are back to the homothetic preferences case. In this case, θ_{ib} is equal to zero and the Engel-curve variety bias becomes $\widetilde{\mathbb{V}}_b^M$.

6 Results

In this section, we describe the procedure to estimate the elasticity of substitution between barcodes within each item and the elasticity of consumption for each barcode. We do so using data from the United States. We then use our decomposition to quantify the importance of each the biases.

6.1 Parameter Estimation

We first estimate the elasticity of substitution between barcodes (σ_{ib}) for each item. Then, taking those estimates as given, we estimate the constant elasticity of consumption (ϵ_{kib}) for each barcode k with respect to the consumption index (C_{kib}^M).

In order to estimate σ_{ib}, we rely on the method developed by Feenstra (1994) and extended by Broda and Weinstein (2006) and Broda and Weinstein (2010). The procedure consists of estimating a demand and supply equation for each barcode by using only the information on prices and quantities. For this estimation, we face the standard endogeneity problem for a given barcode. Although we cannot identify supply and demand, the data do provide information about the joint distribution of supply and demand parameters.

We first model the supply and demand conditions for each barcode within a product category. Specifically, we estimate the demand elasticities by using the following system of differenced demand and supply equations as in Broda and Weinstein (2006):

$$\Delta^{\underline{k},t} \ln S_{kibt} = (1 - \sigma_{ib}) \Delta^{\underline{k},t} \ln P_{kibt} + \iota_{kibt} \tag{23}$$

$$\Delta^{\underline{k},t} \ln P_{kibt} = \frac{\delta_{ib}}{1 + \delta_{ib}} \Delta^{\underline{k},t} \ln S_{kibt} + \kappa_{kibt} \tag{24}$$

Note that when the inverse supply elasticity is zero (i.e. δ_{ib}=0), the supply curve is horizontal and there is no simultaneity bias in σ_g. Equations 23 and 24 are the demand and supply equations of barcode k in an item i differenced with respect to a benchmark barcode in the same item. The k^{th} good corresponds to the largest selling barcode in each item. The k-differencing removes any item level shocks from the data.

The identification strategy relies on two important assumptions. First, we assume that ι_{kibt} and κ_{kibt} are uncorrelated (i.e., $\mathbb{E}_t(\iota_{kibt}\kappa_{kibt}) = 0$). Because we already removed any item level shocks, we are left with within item variation that is likely to render independence of the barcode-level demand and supply shocks within an item. Second, we assume that σ_{ib} and ω_{ib} are restricted to be the same over time and for all barcodes in a given item.

To take advantage of these assumptions, we define a set of moment conditions for each item i in a basic heading b as below:

$$G(\beta_{ib}) = E_T[\nu_{kibt}(\beta_{ib})] = 0 \tag{25}$$

where $\beta_{ib} = [\sigma_{ib}, \delta_{ib}]'$ and $\nu_{kibt} = \iota_{kibt}\kappa_{kibt}$.

For each item i, all the moment conditions that enter the GMM objective function can be combined to obtain Hansen (1982)'s estimator:

$$\hat{\beta_{ib}} = \arg\min_{\beta_{ib}\in B} G^*(\beta_{ib})'WG^*(\beta_{ib}) \quad \forall\, i \in \omega_b \tag{26}$$

where $G^*(\beta_{ib})$ is the sample analog of $G(\beta_{ib})$, W is a positive definite weighting matrix, and B is the set of economically feasible β_{ib} (i.e., $\sigma_{ib} > 0$).

As Argente, Lee and Moreira (2020), we estimate σ_{ib} and δ_{ib} jointly with Nielsen Retail Scanner data from 2006 to 2016, which covers approximately 90 participating retail chains across all US markets.[11] If the procedure renders imaginary estimates or estimates of the wrong sign, we use a grid search to evaluate the GMM objective function above.

[11] For more details on the Nielsen Retail Scanner data, refer to Argente, Lee and Moreira (2018, 2020).

Table 2: Descriptive Statistics of Estimated Parameters

	mean	std. dev.	10th-percentile	median	90th-percentile
σ_{ib}	9.29	3.42	5.61	8.73	12.49
ϵ_{kib}	0.82	1.79	-1.58	0.89	2.99
θ_{ib}	0.02	0.13	-0.13	0.03	0.16
E^M_{ib}/E^U_{ib}	0.72	0.60	0.14	0.53	1.47

Note: This table reports descriptive statistics for the elasticity of substitution (σ_{ib}), the elasticity of consumption of barcode k with respect to the consumption index (ϵ_{kib}), parameter in the non-homothetic CES price index (θ_{ib}) and nominal expenditure ratio (E^M_{ib}/E^U_{ib}).

The first row of Table 2 reports descriptive statistics for σ_{ib}. The average of the elasticity of substitution is 9.29 with standard deviation of 3.42. This is qualitatively consistent with other papers estimating elasticity of substitution with scanner data (Hottman, Redding and Weinstein, 2016; Argente, Lee and Moreira, 2020).

Given σ_{ib} estimates, we use the Engel curve to estimate ϵ_{kib} as in Comin, Lashkari and Mestieri (2020)[12]:

$$\ln \frac{s^h_{kibt}}{s^h_{\mathbf{K}ibt}} - (1-\sigma_{ib}) \ln \frac{p^h_{kibt}}{p^h_{\mathbf{K}ibt}} = (\epsilon_{kib} - 1) \left(\ln \frac{E^h_{ibt}}{p^h_{\mathbf{K}ibt}} + \frac{1}{(1-\sigma_{ib})} \ln s^h_{\mathbf{K}ibt} \right) + \psi^h_t + \epsilon^h_{kibt} \tag{27}$$

where $\mathbf{K}$ is the benchmark barcode, which corresponds to the largest selling barcode in each item, and ψ^h_t is the set of fixed effects and controls. Equation 27 is estimated with quarter fixed effects, state fixed effects, and other observable characteristics at the household level (e.g. age, income). Note that because barcodes are substitutes within all items ($\sigma_{ib} > 1$), it is required $\epsilon_{kib} < \sigma_{ib}$ for the consumption index to be globally monotonically increasing and quasi-concave, and therefore to correspond to a well-defined utility function.[13]

The second row of Table 2 reports the descriptive statistics for ϵ_{kib}. Our estimates for this parameter have a mean of 0.82 and a standard deviation of 1.79. The third row of Table 2 reports the descriptive statistics for θ_{ib}. Recall that this parameter is the equally weighted average of $\frac{\epsilon_{kib}-1}{\sigma_{ib}-1}$ across the common barcodes within an item across the two countries. Our estimates of this parameter have a mean of 0.02 and a standard deviation of 0.13. Note that when θ_{ib} is close to zero, the expenditure ratio plays a small role in the price index. Lastly, we report other informative moments for the quantification of the Engel-curve variety bias

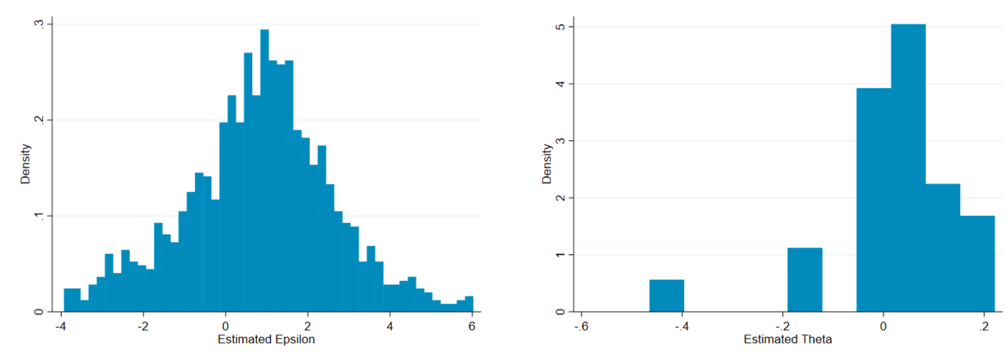
[12] Find Appendix section D for the derivation of estimation equation

[13] In less than one percent of the cases, $\epsilon_{kib} >= \sigma_{ib}$. In these cases, we impute $\epsilon_{kib} = \sigma_{ib} - 0.01$.

such as the nominal expenditure ratio. As the last row of Table 2 shows, the ratio varies across items. It has a median of 0.53 and a mean of 0.72 (with standard deviation of 0.60), which indicates that the distribution is skewed to the right.

Figure 6 shows distribution of estimated ϵ_{kib} and θ_{ib}. Mean of ϵ_{kib} is close to one. Therefore, mean of θ_{ib} is around zero.

Figure 6: Distribution of Estimated ϵ_{kib} and θ_{ib}



ϵ_{ki} θ_i

Note: Panel (a) and (b) show the distribution of estimated ϵ_{kib} and θ_{ib}, respectively.

6.2 Decomposition Results

Table 3 reports the price index developed by the ICP ($\mathbb{ICP}^M$) and the non-homothetic CES price index ($\mathbb{NH}^M$) after aggregating across basic headings. Recall that the ICP aggregates across basic headings using a Fisher index. In the table we report both the results after aggregating basic headings using a Fisher index and our using our theoretical framework which imply a CES aggregation in the first layer. Equation 18 indicates that the gap between the two price indexes can be decomposed into the sampling bias ($\mathbb{S}^M$), the quality bias ($\mathbb{Q}^M$), and the Engel-curve variety bias ($\mathbb{V}^M$). The aggregate bias (Θ^M) is estimated to be 0.90.

Table 3: Decomposition Results for the Non-Homothetic CES Price Index

	Price Indexes		Biases			Agg. Bias
	$\mathbb{ICP}^M$	$\mathbb{NH}^M$	$\mathbb{S}^M$	$\mathbb{Q}^M$	$\mathbb{V}^M$	Θ^M
CES	0.80	0.72	0.82	1.45	0.75	0.90
Fisher	0.82	0.74	0.84	1.45	0.74	0.90

Note: The table reports ICP-style price index ($\mathbb{ICP}^M$) and the non-homothetic CES price index ($\mathbb{NH}^M$). By equation 18, the gap between two price indexes can be decomposed into sampling bias ($\mathbb{S}^M$), quality bias ($\mathbb{Q}^M$), and Engel-curve variety bias ($\mathbb{V}^M$). Θ^M is the aggregate bias defined as $\mathbb{S}^M \times \mathbb{Q}^M \times \mathbb{V}^M$. We report two different way of aggregations at the top level of aggregation: CES aggregation consistent to the model and Fisher aggregation as a robustness check.

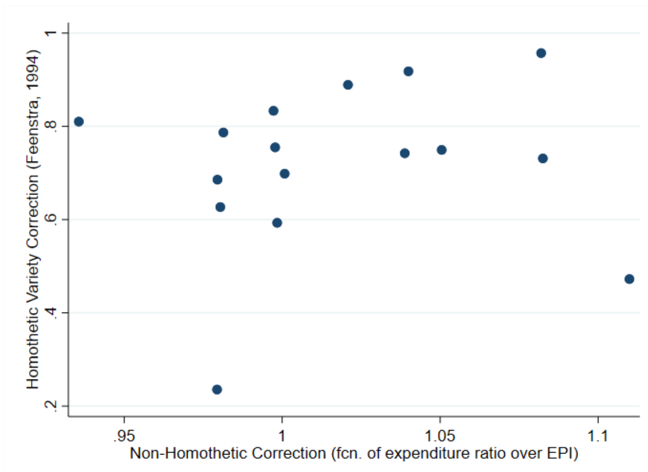
The sampling bias is less than 1 (0.82). We do not find a significant difference across the two countries in the first term of the sampling bias, which is related to the covariance of the expenditure weights at the item level and their prices within a basic heading. This is due to the fact that there are not many items within a basic heading, so the difference between aggregating with expenditure weights or with equal weights within a basic heading is not significantly different across the two countries.

The direction of the sampling bias is mainly due to the importance of the second term, which is related to the fact that the sample probabilities are proportionate to the store volume of sales and do not reflect the sales of individual products. This is particularly relevant in Mexico, where we find that consumers tend to spend more at stores where specific items have lower prices. In other words, we find that $\text{cov}(\phi^{\mathbf{M}}_{\mathbf{ib}}, \ln(\bar{\mathbf{p}}^{\mathbf{M}}_{\mathbf{ib}})) < \text{cov}(\phi^{\mathbf{U}}_{\mathbf{ib}}, \ln(\bar{\mathbf{p}}^{\mathbf{U}}_{\mathbf{ib}}))$ after controlling for item and country effects. In Appendix B.2 we provide more evidence of this relationship using a reduced form and show it is robust to several specifications.

The quality bias (1.45) constitutes a downward bias to the original ICP estimates; an understatement of the price level in Mexico. This is because low quality varieties in Mexico are matched to high quality varieties in the United States, even after sampling items that meet the the same ICP specifications across the two countries.

However, Mexico has important gains from Engel-curve variety (0.75) reflected in the fact that the consumption shares of common goods across the two countries is lower than that of the United States. Engel-curve variety bias in equation 21 can be decomposed into two components: (i) homothetic variety correction as in Feenstra (1994) and (ii) non-homothetic component as a function of nominal expenditure ratio over the exact price index. Figure 7 shows two components for each basic heading. Homothetic variety correction is below one for all basic headings, with a mean around 0.72, while the non-homothetic correction doesn't vary much across basic headings, with a mean around 1.02. Therefore, most of variation in the Engel-curve variety bias comes from traditional variety correction, while contribution of non-homothetic correction is marginal.

Figure 7: Two Components in Engel-curve Variety Bias



Note: Engel-curve variety bias can be decomposed into two components: (i) homothetic variety correction as in Feenstra (1994) and (ii) non-homothetic component as a function of nominal expenditure ratio over the exact price index. The figure plots two components for each basic heading.

Note that if the Engel-curve variety term is not included, as is the case of the commonly used Sato-Vartia price index, the aggregate bias would 1.19 which would imply that the ICP estimates are overestimated instead of underestimated.[14]

Table 4 reports the same decomposition assuming that preferences are homothetic (i.e. $\mathbb{EPI}^M$). In this case, the aggregate bias (Θ^M) is estimated to be 0.88. The homothetic variety bias ($\widetilde{\mathbb{V}}^M$) is 0.74 which is very similar to the Engel-curve variety bias (0.75).

[14]In Appendix Section E, we report decomposition results in the case with common elasticity of substitution $\sigma_{ib} = 9$. Results do not vary much.

Table 4: Decomposition Results for the Homothetic Exact Price Index

	Price Indexes		Biases			Agg. Bias
	$\mathbb{ICP}^M$	$\mathbb{EPI}^M$	$\mathbb{S}^M$	$\mathbb{Q}^M$	$\widetilde{\mathbb{V}}^M$	$\widetilde{\Theta}^M$
CES	0.80	0.71	0.82	1.45	0.74	0.88
Fisher	0.82	0.74	0.84	1.45	0.74	0.90

Note: The table reports ICP-style price index ($\mathbb{ICP}^M$) and the homothetic Exact Price Index ($\mathbb{EPI}^M$). By equation 18, the gap between two price indexes can be decomposed into sampling bias ($\mathbb{S}^M$), quality bias ($\mathbb{Q}^M$), and homothetic variety bias ($\widetilde{\mathbb{V}}^M$). $\widetilde{\Theta}^M$ is the aggregate bias defined as $\mathbb{S}^M \times \mathbb{Q}^M \times \widetilde{\mathbb{V}}^M$. We report two different way of aggregations at the top level of aggregation: CES aggregation consistent to the model and Fisher aggregation as a robustness check.

Table 5 reports all price indexes: (i) ICP ($\mathbb{ICP}^M$), (ii) Sato-Vartia price index ($\mathbb{SV}^M$), (iii) homothetic Exact Price Index ($\mathbb{EPI}^M$), and (iv) the non-homothetic CES price index ($\mathbb{NH}^M$). Compared to the ICP index, the Sato-Vartia price index is higher after correcting sampling and quality biases. The homothetic Exact Price Index and non-homothetic CES price index are slightly lower than the ICP-style price index, after correcting homothetic and Engel-curve variety bias. Our results indicate that the real income inequality across the US and Mexico is lower than that predicted by the ICP estimates and have implications for the comparison of standards of living for the calculation of poverty counts across the two countries.

Table 5: Multiple Price Indexes

	Price Indexes			
	$\mathbb{ICP}^M$	$\mathbb{SV}^M$	$\mathbb{EPI}^M$	$\mathbb{NH}^M$
CES	0.80	0.95	0.71	0.72
Fisher	0.82	1.00	0.74	0.74

Note: The table reports ICP-style price index ($\mathbb{ICP}^M$), Sato-Vartia price index ($\mathbb{SV}^M$), homothetic Exact Price Index ($\mathbb{EPI}^M$), and the non-homothetic CES price index ($\mathbb{NH}^M$). We report two different way of aggregations at the top level of aggregation: CES aggregation consistent to the model and Fisher aggregation as a robustness check.

7 Conclusion

The construction of cross-country price indexes is a difficult task and yet it is of crucial importance to compare living standards between countries and to measure global inequality. The ICP faces several challenges both of data availability and also methodologically. In this paper, we construct a data set for two countries that allows us to address some of the data limitations faced by the ICP; namely, the fact that they cannot compare exactly the same item across countries and that they do not have expenditure information to sample items appropriately. We also develop a non-homothetic CES index to be able to account for differences in the availability of varieties across countries as well as differences in real consumption. Using this index we estimate that the Mexican real consumption is larger relative to the United States than previously estimated. We identify heterogeneity in shopping behavior, quality of products and variety availability as important sources of bias in international price comparisons.

References

Anderson, Simon P, André De Palma, and Jacques-Francois Thisse, "The CES is a discrete choice model?," *Economics Letters*, 1987, *24* (2), 139–140.

Argente, David and Munseob Lee, "Cost of Living Inequality during the Great Recession," *Journal of the European Economic Association*, forthcoming.

_ , _ , and Sara Moreira, "Innovation and product reallocation in the great recession," *Journal of Monetary Economics*, 2018, *93*, 1–20.

_ , _ , and _ , "The Life Cycle of Products: Evidence and Implications," 2020.

Bils, Mark and Peter J Klenow, "Quantifying Quality Growth," *American Economic Review*, 2001, pp. 1006–1030.

Broda, Christian and David E Weinstein, "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, 2006, *121* (2), 541–585.

_ and _ , "Product creation and destruction: Evidence and price implications," *American Economic Review*, 2010, *100* (3), 691–723.

_ , Ephraim Leibtag, and David E Weinstein, "The role of prices in measuring the poor's living standards," *Journal of Economic Perspectives*, 2009, *23* (2), 77–97.

Cavallo, Alberto, Robert C. Feenstra, and Robert Inklaar, "Product Variety, Trade Costs and the Standard of Living Across Countries," 2020.

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

Comin, Diego A, Danial Lashkari, and Martí Mestieri, "Structural Change with Long-run Income and Price Effects," *NBER Working Paper No. 21595*, 2020.

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

Faber, Benjamin, "Trade Liberalization, the Price of Quality, and Inequality: Evidence from Mexican Store Prices," 2014.

Feenstra, Robert C, "New product varieties and the measurement of international prices," *American Economic Review*, 1994, pp. 157–177.

— , Mingzhi Xu, and Alexis Antoniades, "What is the Price of Tea in China? Goods Prices and Availability in Chinese Cities," *The Economic Journal*, 2019.

Handbury, Jessie, "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across US Cities," *NBER Working Paper No. 26574*, 2019.

Hanoch, Giora, "Production and demand models with direct or indirect implicit additivity," *Econometrica*, 1975, *43* (3), 395.

Hansen, Lars Peter, "Large sample properties of generalized method of moments estimators," *Econometrica*, 1982, pp. 1029–1054.

Hottman, Colin J, Stephen J Redding, and David E Weinstein, "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, 2016, *131* (3), 1291–1364.

Hsieh, Chang-Tai and Peter J Klenow, "Relative prices and relative prosperity," *American Economic Review*, 2007, *97* (3), 562–585.

Matsuyama, Kiminori, "Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade," *Econometrica*, 2019, *87* (2), 497–528.

Redding, Stephen J and David E Weinstein, "Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences," *The Quarterly Journal of Economics*, 2020, *135* (1), 503–560.

Sato, Ryuzo, "The most general class of CES functions," *Econometrica*, 1975, *43* (5-6), 999.

Simonovska, Ina, "Income differences and prices of tradables: Insights from an online retailer," *The Review of Economic Studies*, 2015, *82* (4), 1612–1656.

APPENDIX

A Non-homothetic CES as a Discrete Choice Model

Consider an individual who has to choose a certain amount of one good among n possible goods. His choice is seen as a two-stage process as in Anderson, De Palma and Thisse (1987): i) he chooses which good to buy, and ii) the quantity of that good.

Second Stage

Suppose that the outcome of the first stage is good i. Then the individual's conditional direct utility is assumed to be:

$$u_i = \ln q_i$$

where q_i is the quantity of good i. Maximizing under the budget constraint yields the demand

$$q_i^* = \frac{y}{p_i}$$

The conditional indirect utility is therefore

$$V(p_i) = -\ln p_i + \ln y$$

First Stage

Assume that the choice of good i follows

$$U_i = V(p_i) + \mu(\epsilon_i - 1)\ln q + \mu\gamma_i$$

where U_i is the stochastic utility associated with i, μ is a positive constant, ϵ_i is a constant, and γ_i is a random variable with zero mean and unit variance. The probability for the individual to choose i is given by

$$s_i = \text{Prob}\left[U_i = \max_{j=1,\ldots,n} U_j\right]$$

assuming that γ_i's are identically, independently Gumbel distributed, s_i becomes the multinomial logit:

$$s_i = \frac{e^{\frac{V(p_i)+\mu(\epsilon_i-1)\ln q}{\mu}}}{\sum_{j=1}^{n} e^{\frac{V(p_j)+\mu(\epsilon_j-1)\ln q}{\mu}}}$$

using the definition of $V(p_i)$ and letting $\mu = \frac{1}{\sigma-1}$ we find

$$s_i = \frac{p_i^{1-\sigma} q^{\epsilon_i-1}}{\sum_{j=1}^{n} p_j^{1-\sigma} q^{\epsilon_j-1}}$$

which is the same share as in Comin, Lashkari and Mestieri (2020). For $\epsilon_i = 1$ for all i, the expression is equivalent to the CES utility function case.

B Sampling Bias

B.1 Proofs of Propositions

Proposition 1. If the number of basic headings $N_b \to \infty$, the first term of the sampling bias is larger than 1 if $\text{cov}(\omega_{\mathbf{b}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{b}}^{\mathbf{M}}\right)) > \text{cov}(\omega_{\mathbf{b}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{b}}^{\mathbf{U}}\right))$.

Proof. The bias on the first term for country M is:

$$\prod_{i\in\Omega_b} \frac{\left(\bar{p}_{ib}^{M}\right)^{\omega_{ib}}}{\left(\bar{p}_{ib}^{M}\right)^{\frac{1}{N_b}}}$$

This ratio is greater than one if and only if:

$$\sum_{i\in\Omega_b}\left(\omega_{ib} - \frac{1}{N_b}\right)\ln\left(\bar{p}_{ib}^{M}\right) > 0 \tag{28}$$

We want to show that this term is equivalent to $\text{cov}(\omega_{\mathbf{b}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{b}}^{\mathbf{M}}\right))$ where $\omega_{\mathbf{b}}$ is a vector of weights in basic heading b and $\ln\left(\bar{\mathbf{p}}_{\mathbf{b}}^{\mathbf{M}}\right)$ is the vector of log prices. By definition:

$$\text{cov}(\omega_{\mathbf{b}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{b}}^{\mathbf{M}}\right)) = \lim_{N_b\to\infty} \frac{1}{N_b - 1}\sum_{i\in\Omega_b}\left(\omega_{ib} - \frac{1}{N_b}\right)\left(\ln\left(\bar{p}_{ib}^{M}\right) - \frac{1}{N_b - 1}\sum_{i\in\Omega_b}\ln\left(\bar{p}_{ib}^{M}\right)\right)$$

Using that

$$\lim_{N_b\to\infty} \frac{1}{N_b - 1}\sum_{i\in\Omega_b}\left(\omega_{ib} - \frac{1}{N_b}\right)\left(\frac{1}{N_b - 1}\sum_{i\in\Omega_b}\ln\left(\bar{p}_{ib}^{M}\right)\right) = 0$$

Then $\text{cov}(\omega_{\mathbf{b}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{b}}^{\mathbf{M}}\right))$ is equivalent to $\frac{1}{N_b-1}\sum_{i\in\Omega_b}\left(\omega_{ib} - \frac{1}{N_b}\right)\ln\left(\bar{p}_{ib}^{M}\right)$.

Proposition 2. If the number of stores $S \to \infty$, the second term of the sampling bias is larger than 1 if $\text{cov}(\phi_{\mathbf{ib}}^{\mathbf{M}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{ib}}^{\mathbf{M}}\right)) - \text{cov}(\phi_{\mathbf{ib}}^{\mathbf{U}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{ib}}^{\mathbf{U}}\right)) > \text{cov}((\phi^{\mathbf{U}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{ib}}^{\mathbf{U}}\right)) - \text{cov}((\phi^{\mathbf{M}}, \ln\left(\bar{\mathbf{p}}_{\mathbf{ib}}^{\mathbf{M}}\right))$

Proof. Before aggregating across items, the second term of the sample bias for a country c is:

$$\frac{\hat{p}^c_{ib}}{\bar{p}^c_{ib}} = \frac{\prod_{s\in\Psi^c} \left(\bar{p}^c_{sib}\right)^{\phi^c_{sib}}}{\prod_{s\in\Psi^c} \left(\bar{p}^c_{sib}\right)^{\phi^c_{s}}}$$

This ratio is larger than 1 for country c if

$$\sum_{s\in\Psi^c} \phi^c_{sib} \ln\left(\bar{p}^c_{sib}\right) > \sum_{s\in\Psi^c} \phi^c_{s} \ln\left(\bar{p}^c_{sib}\right) \tag{29}$$

Let $\phi^{\mathbf{c}}_{\mathbf{ib}}$ be the vector of expenditure weights that vary by item and basic heading, (i.e. $\{\phi^c_{sib}\}_{s\in\Psi^c}$), $\phi^{\mathbf{c}}$ be the vector of weights that only vary at the store level (i.e. $\{\phi^c_{s}\}_{s\in\Psi^c}$), and $(\bar{\mathbf{p}}^{\mathbf{c}}_{\mathbf{ib}})$ be the vector of log prices (i.e. $\{\ln\left(\bar{p}^c_{sib}\right)\}_{s\in\Psi^c}$). Then equation 29 can be written as:

$$\lim_{S\to\infty} \frac{1}{S-1} \sum_{s\in\Psi^U} \phi^U_{sib} \ln\left(\bar{p}^U_{sib}\right) > \lim_{S\to\infty} \frac{1}{S-1} \sum_{s\in\Psi^U} \phi^U_{s} \ln\left(\bar{p}^U_{sib}\right)$$

Using the definition of covariance on both sides

$$\begin{aligned}\lim_{S\to\infty} \frac{1}{S-1} \sum_{s\in\Psi^c} \phi^c_{sib} \times \frac{1}{S-1} \sum_{s\in\Psi^c} \ln\left(\bar{p}^c_{sib}\right) + \frac{1}{S-1} \sum_{s\in\Psi^c} \left(\phi^c_{sib} - \bar{\phi}^c_{ib}\right) \left(\ln\left(\bar{p}^c_{sib}\right) - \overline{\ln\left(\bar{p}^c_{sib}\right)}\right) > \\ \lim_{S\to\infty} \frac{1}{S-1} \sum_{s\in\Psi^c} \phi^c_{s} \times \frac{1}{S-1} \sum_{s\in\Psi^c} \ln\left(\bar{p}^c_{sib}\right) + \frac{1}{S-1} \sum_{s\in\Psi^c} \left(\phi^c_{s} - \bar{\phi}^c\right) \left(\ln\left(\bar{p}^c_{sib}\right) - \overline{\ln\left(\bar{p}^c_{sib}\right)}\right)\end{aligned}$$

Taking the limit as $\lim_{S\to\infty}$ on both sides, we find that $\mathrm{cov}(\phi^{\mathbf{c}}_{\mathbf{ib}}, \bar{\mathbf{p}}^{\mathbf{c}}_{\mathbf{ib}}) > \mathrm{cov}((\phi^{\mathbf{c}}, \bar{\mathbf{p}}^{\mathbf{c}}_{\mathbf{ib}})$.

B.2 Empirical Tests

To quantify the importance of the first term of the sampling bias, we rely on Proposition 1 and test whether $\mathrm{cov}(\omega_{\mathbf{b}}, \ln\left(\bar{\mathbf{p}}^{\mathbf{M}}_{\mathbf{b}}\right)) > \mathrm{cov}(\omega_{\mathbf{b}}, \ln\left(\bar{\mathbf{p}}^{\mathbf{U}}_{\mathbf{b}}\right))$. To do so we realy on the following specification:

$$\omega_{ib} = \alpha + \beta\ \ln(\bar{p}^c_{ib}) \times \mathbb{1}\left\{c = \text{Mexico}\right\} + \lambda^c + \theta_b + \epsilon^c_{ib}$$

where the dependent variable are the Sato-Vartia weights for each item. The coefficient of interest is β which indicates whether there is a difference between the covariance between the weights and the prices of items across the two countries; Table B1 shows that we do not find a significant difference indicating that the first term of the sampling bias is close to 1.

Table B1: Sampling Bias First Term: Expenditure Weights and Prices

	(1)	(2)	(3)	(4)
$\ln(\bar{p})$	-0.010	-0.044	-0.010	-0.039
	(0.071)	(0.051)	(0.071)	(0.059)
$\ln(\bar{p}) \times$ Mexico	-0.002	0.004	-0.002	-0.006
	(0.095)	(0.023)	(0.095)	(0.056)
Observations	58	58	58	58
R-squared	0.001	0.775	0.001	0.775
Basic Heading	N	Y	N	Y
Country	N	N	Y	Y

Note: The table shows the relationship between the expenditure weights at the item level and the prices of items within a basic heading. Column (2) includes basic heading effects, Column (3) includes country effects, and Column (4) both.

To quantify the size of the second term of the sampling bias, we rely on Proposition 2.

In order to compare the magnitude of $\text{cov}(\phi^{\mathbf{M}}_{\mathbf{ib}}, \ln\left(\bar{\mathbf{p}}^{\mathbf{M}}_{\mathbf{ib}}\right))$ relative to $\text{cov}(\phi^{\mathbf{U}}_{\mathbf{ib}}, \ln\left(\bar{\mathbf{p}}^{\mathbf{U}}_{\mathbf{ib}}\right))$ we estimate the following specification:

$$\phi^c_{sib} = \alpha + \beta \ln(\bar{p}^c_{sib}) \times \mathbb{1}\left\{c = \text{Mexico}\right\} + \theta^c + \lambda_i + \epsilon^c_{sib}$$

where the dependent variable are the country-specific expenditure weights at the store-item level and the independent variable are the log prices at the same level. We include country and item effects in the specification. Table B2 presents the results. It shows that the covariance of expenditure weights and prices is strongly negative for items and stores in Mexico. The results are robust after controlling for store, item, and country effects simultaneously.

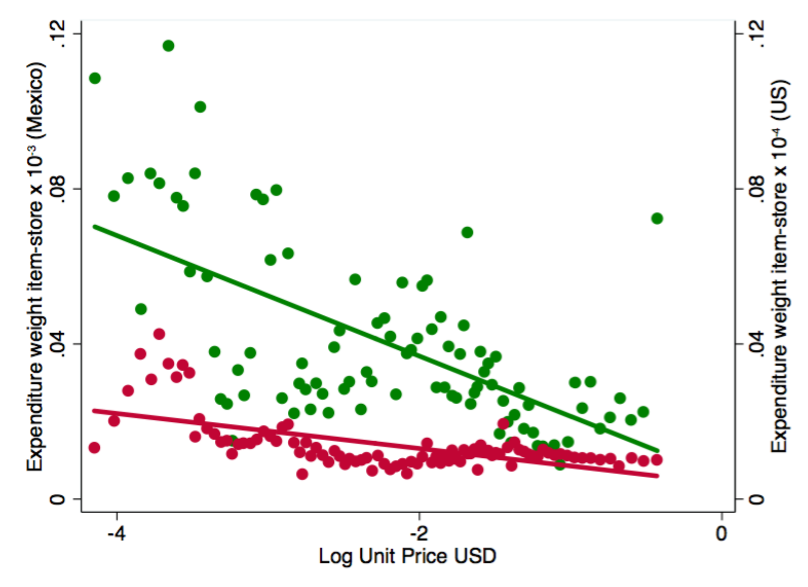
Table B2: Decomposition Results for the Homothetic Exact Price Index

	(1)	(2)	(3)
$\ln(\bar{p})$	-0.0004	0.0032	0.0039
	(0.004)	(0.005)	(0.004)
$\ln(\bar{p}) \times$ Mexico	-0.1528***	-0.0756**	-0.1147**
	(0.000)	(0.004)	(0.005)
Observations	764,419	764,419	761,751
R-squared	0.028	0.033	0.216
Store	N	N	Y
Item	N	Y	Y
Country	N	Y	Y

Note: The table shows the results of estimating the relationship between the country-specific expenditure weights at the store-item level and the log prices at the same level. The dependent variable is multiplied times 10^4. Column (2) includes item and country effects, Column (3) includes the same controls in addition to store effects.

Figure B1 shows the same relationship graphically. The green line in the figure shows that the relationship between the expenditure weights and the log prices at the store-item level is substantially more negative for Mexico than for the US (red line). This shows that the second term of the sampling bias, given that the covariance of weights and prices is more negative for Mexico, is less than 1.

Figure B1: Weights and Log(prices)



Note: The figure show the relationship between the expenditure weights at the store-item level and the log prices at the same level for each country. The green line shows this relationship for Mexico and the red line for the US.

C Other Stylized Facts

C.1 Law of One Price

There are large deviations from the law of one price within and across countries. We measure price dispersion across all the stores in our sample, focusing on the price gap between stores located in the same country versus the price gap between stores located in different countries during the first three months of 2012. A store in our data is a city$\times$retailer combination. We have a total of 5,591,169 store pairs in the United States, 73,876 store pairs in Mexico, and 220,915 cross-border store pairs. For all the matched barcodes in each store pair, we compute the difference in the log price between the two stores. Table C1 presents statistics across store pairs on the mean, median, and maximum of the absolute price gap for store-pairs located in the United States (US-US), Mexico (Mexico-Mexico), and across the border (US-Mexico). The median price gap across store pairs is 22 percent for US store-pairs, 12 percent for Mexico store pairs, and 30 percent for cross-border pairs; the dispersion in cross-border price gaps vastly exceeds that of within country price gaps. The price gaps for the same good might reflect differences in the quality of local retail services and highlights the importance of the sampling methodology to compute national average prices.

Table C1: Law of One Price (US and Mexico)

The table reports price dispersion across stores in our sample, focusing on the price gap between cities located in the same country versus the price gap between stores located in different countries during our sample period. The statistics reported are the mean absolute, median absolute, and max absolute (log) price gap within store pairs for the first three months of 2012.

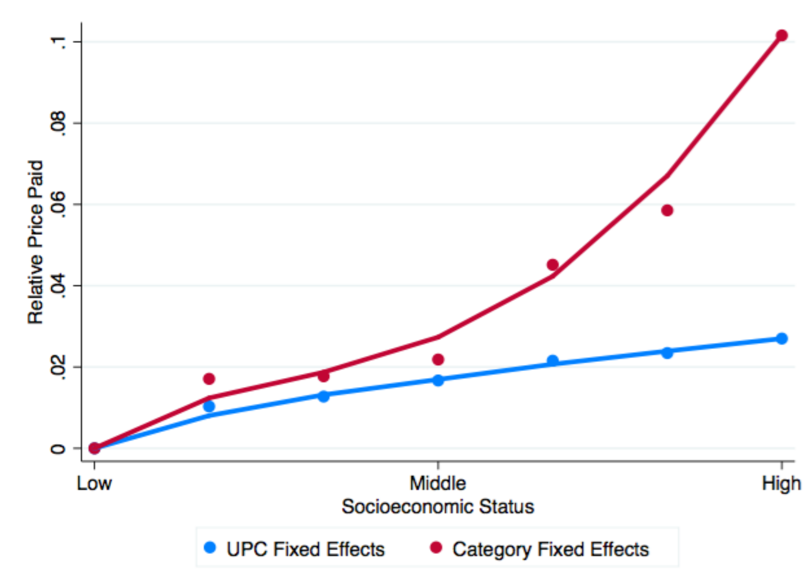
US-US	Mean Absolute	Median Absolute	Max Absolute
Mean	.309	.284	.57
Median	.255	.221	.441
Std. Dev.	.272	.274	.563
Mexico-Mexico	**Mean Absolute**	**Median Absolute**	**Max Absolute**
Mean	.158	.151	.226
Median	.134	.122	.191
Std. Dev.	.127	.128	.193
US-Mexico	**Mean Absolute**	**Median Absolute**	**Max Absolute**
Mean	.521	.494	.705
Median	.331	.308	.389
Std. Dev.	.722	.718	.984

C.2 Quality Engel Curve Using Socio-economic Status

Socio-economic classification in the Nielsen data has seven categories, and it is widely used in other surveys in Mexico. The 7 levels are: A/B, C+, C, C-, D+, E, and E. The level A/B consists mainly of households (82%) whose head has at least and undergraduate degree, they have internet available at home, and spend approximately 13% of their income on education and less than 25% on food. On the other hand, level E consists of households (95%) whose household head finished at most elementary school, do not have access to internet, and most of its expenditures are destined for food (52%) and only 5% to education.

Figure C1 plots the quality Engel curves for Mexico using the Socio-economic classification in the Nielsen data. In the figure, the relative prices are measured in a regression of the log unit price paid against socioeconomic status dummies and category, region, chain, and quarter fixed effects. Each dot represents how much more households in each income category pay per unit for products within a category than households at the lowest socioeconomic status. Within a product category, high socioeconomic status households pay 10 percent higher price than low socioeconomic status households. A substantial amount of difference comes from the fact that high socioeconomic status households buy higher quality products. With barcode fixed effects, high socioeconomic households pay 2 percent higher price than low socioeconomic status households to buy the same good.

Figure C1: Quality Engel Curve in Mexico Using Socio-economic Status



The figure plots the cross-sectional relationship between the relative prices paid and household income. The relative prices are measured in a regression of the log unit price paid against socioeconomic status dummies and category, region, chain, quarter and household fixed effects. Each dot represents how much more households in each socioeconomic bin pays per unit for products with respect to households belonging to the lowest socioeconomic status.

D Derivation of the Model-based Engel Curve

In this section, we derive the model-based Engel curve as equation 27. We start with the demand:

$$s_{kt}^{h} = (\varphi)^{\sigma-1} \left(\frac{p_{kt}}{P_t^h}\right)^{1-\sigma} \left(C_t^h\right)^{\epsilon_k - 1} \tag{30}$$

Taking logs

$$\ln s_{kt}^{h} = (\sigma-1)\ln P_t^h - (\sigma-1)\ln p_{kt} + (\sigma-1)\ln \varphi_{kt} + (\epsilon_k - 1)\ln C_t^h \tag{31}$$

Subtract the demand for a baseline product

$$\ln \frac{s_{kt}^{h}}{s_{bt}^{h}} = (1-\sigma)\ln \frac{p_{kt}}{p_{bt}} + (\sigma-1)\ln \frac{\varphi_{kt}}{\varphi_{bt}} + (\epsilon_k - \epsilon_b)\ln C_t^h \tag{32}$$

We want to find an expression to substitute for C_t^h. We begin by using 31 for the baseline product.

$$\ln s_{bt}^{h} = (\sigma-1)\ln P_t^h - (\sigma-1)\ln p_{bt} + (\sigma-1)\ln \varphi_{bt} + (\epsilon_b - 1)\ln C_t^h$$

Using that $P_t^h C_t^h = E_t^h$ and dividing by $(1-\sigma)$ we get

$$\frac{1}{1-\sigma}\ln s_{bt}^{h} = -\ln E_t^h + \ln p_{bt} - \ln \varphi_{bt} + \epsilon_b \ln C_t^h$$

which implies

$$\ln C_t^h = \frac{1}{\epsilon_b(1-\sigma)}\ln s_{bt}^{h} + \frac{1}{\epsilon_b}\ln \frac{E_t^h}{p_{bt}} + \frac{1}{\epsilon_b}\ln \varphi_{bt} \tag{33}$$

we now substitute 33 into 32

$$\begin{aligned}
\ln \frac{s_{kt}^{h}}{s_{bt}^{h}} &= (1-\sigma)\ln \frac{p_{kt}}{p_{bt}} + (\sigma-1)\ln \frac{\varphi_{kt}}{\varphi_{bt}} + (\epsilon_k - \epsilon_b)\left(\frac{1}{\epsilon_b(1-\sigma)}\ln s_{bt}^{h} + \frac{1}{\epsilon_b}\ln \frac{E_t^h}{p_{bt}} + \frac{1}{\epsilon_b}\ln \varphi_{bt}\right) \\
&= (1-\sigma)\ln \frac{p_{kt}}{p_{bt}} + (\sigma-1)\ln \frac{\varphi_{kt}}{\varphi_{bt}} + \frac{(\epsilon_k - \epsilon_b)}{\epsilon_b(1-\sigma)}\ln s_{bt}^{h} + \frac{(\epsilon_k - \epsilon_b)}{\epsilon_b}\ln \frac{E_t^h}{p_{bt}} + \frac{(\epsilon_k - \epsilon_b)}{\epsilon_b}\ln \varphi_{bt}
\end{aligned}$$

using that $\varepsilon_k = \epsilon_k/\epsilon_b$ and collecting terms

$$\ln \frac{s_{kt}^{h}}{s_{bt}^{h}} = (1-\sigma)\ln \frac{p_{kt}}{p_{bt}} + (\varepsilon_k - 1)\ln \frac{E_t^h}{p_{bt}} + \frac{(\epsilon_k - 1)}{(1-\sigma)}\ln s_{bt}^{h} + \ln \frac{\varphi_{kt}^{\sigma-1}}{\varphi_{bt}^{\sigma-\epsilon_k}}$$

Recall that in our setup $\varphi_{kt} = \varphi_k$ for every k. Then our estimation equation is

$$\ln \frac{s_{kt}^h}{s_{bt}^h} = (1-\sigma)\ln \frac{p_{kt}}{p_{bt}} + (\varepsilon_k - 1)\ln \frac{E_t^h}{p_{bt}} + \frac{(\epsilon_k - 1)}{(1-\sigma)}\ln s_{bt}^h + \delta_k + \nu_{kt}^h$$

where we add an extra term ν_{kt}^h that account for measurement error and can be further separated using time effects or household $\times$ time effects:

$$\ln \frac{s_{kt}^h}{s_{bt}^h} = (1-\sigma)\ln \frac{p_{kt}}{p_{bt}} + (\varepsilon_k - 1)\ln \frac{E_t^h}{p_{bt}} + \frac{(\epsilon_k - 1)}{(1-\sigma)}\ln s_{bt}^h + \delta_k + \xi_t^h + \eta_{kt}^h$$

which is consistent to equation 27 we use for elasticity of consumption estimation for each barcode.

E Decomposition Results with Common Elasticity of Substitution

In this section, we report decomposition results with common elasticity of substitution across items, $\sigma_{ib} = 9$, which is chosen from a mean of σ_{ib} in Table 2. Table E2, and Table E3 report decomposition results for the non-homothetic CES price index and homothetic exact price index, respectively. Table E4 reports multiple price indexes. The homothetic price index (0.69) and the non-homothetic CES price index (0.72) are similar to the case with estimated σ_{ib} (0.71 and 0.72).

Table E2: Decomposition Results for the Non-Homothetic CES Price Index with $\sigma_{ib} = 9$

	Price Indexes		Biases			Agg. Bias
	$\mathbb{ICP}^M$	$\mathbb{NH}^M$	$\mathbb{S}^M$	$\mathbb{Q}^M$	$\mathbb{V}^M$	Θ^M
CES	0.80	0.69	0.82	1.45	0.73	0.87
Fisher	0.82	0.72	0.84	1.45	0.72	0.88

Note: The table reports ICP-style price index ($\mathbb{ICP}^M$) and the non-homothetic CES price index ($\mathbb{NH}^M$). By equation 18, the gap between two price indexes can be decomposed into sampling bias ($\mathbb{S}^M$), quality bias ($\mathbb{Q}^M$), and Engel-curve variety bias ($\mathbb{V}^M$). Θ^M is the aggregate bias defined as $\mathbb{S}^M \times \mathbb{Q}^M \times \mathbb{V}^M$. We report two different way of aggregations at the top level of aggregation: CES aggregation consistent to the model and Fisher aggregation as a robustness check.

Table E3: Decomposition Results for the Homothetic Exact Price Index with $\sigma_{ib} = 9$

	Price Indexes		Biases			Agg. Bias
	$\mathbb{ICP}^M$	$\mathbb{EPI}^M$	$\mathbb{S}^M$	$\mathbb{Q}^M$	$\tilde{\mathbb{V}}^M$	$\tilde{\Theta}^M$
CES	0.80	0.69	0.82	1.45	0.72	0.86
Fisher	0.82	0.72	0.84	1.45	0.72	0.88

Note: The table reports ICP-style price index ($\mathbb{ICP}^M$) and the homothetic Exact Price Index ($\mathbb{EPI}^M$). By equation 18, the gap between two price indexes can be decomposed into sampling bias ($\mathbb{S}^M$), quality bias ($\mathbb{Q}^M$), and homothetic variety bias ($\tilde{\mathbb{V}}^M$). $\tilde{\Theta}^M$ is the aggregate bias defined as $\mathbb{S}^M \times \mathbb{Q}^M \times \tilde{\mathbb{V}}^M$. We report two different way of aggregations at the top level of aggregation: CES aggregation consistent to the model and Fisher aggregation as a robustness check.

Table E4: Multiple Price Indexes with $\sigma_{ib} = 9$

	Price Indexes			
	$\mathbb{ICP}^M$	$\mathbb{SV}^M$	$\mathbb{EPI}^M$	$\mathbb{NH}^M$
CES	0.80	0.95	0.69	0.69
Fisher	0.82	1.00	0.72	0.72

Note: The table reports ICP-style price index ($\mathbb{ICP}^M$), Sato-Vartia price index ($\mathbb{SV}^M$), homothetic Exact Price Index ($\mathbb{EPI}^M$), and the non-homothetic CES price index ($\mathbb{NH}^M$). We report two different way of aggregations at the top level of aggregation: CES aggregation consistent to the model and Fisher aggregation as a robustness check.