International Price Comparison using Scanner Data*

David Argente†
FRB Minneapolis

Chang-Tai Hsieh‡
University of Chicago

Munseob Lee§
UC San Diego

February 2019
very preliminary and incomplete
please do NOT cite without permission

Abstract

Cross-country price indices are crucial to compare living standards between countries. An accurate measurement of these price indices has proven to be an extremely difficult task because the consumption patterns of different countries do not overlap. We construct a unique data on prices and quantities at the barcode-level across two countries with different income level, United States and Mexico. At the aggregate level, we identify heterogeneity in quality and variety as important sources of bias in international price comparisons. In addition, we compute income- and country-specific price indices using a non-homothetic preference structure. We find that the price indices crucially depend on both the income distribution of each country and on the local producers' choice of quality. Our results indicate that previous studies importantly understate real income inequality between countries.

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

*The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. Empirical results are calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by 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.

†Email: david.argenteamaya@mpls.frb.org. Address: 90 Hennepin Ave, Minneapolis, MN 55401.

‡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

How large is real income inequality across countries? For decades, economists have tried to study the patterns of poverty and inequality by focusing only on the disparities in nominal income or nominal wages. The measures of real income inequality that account for differences in the cost of living are arguably better measures of the differences in the standard of living.[1] For this reason, the International Comparison Program (ICP) collects prices in countries around the world and, together with the national accounts for each country, calculates purchasing power parities (PPP), which are used to compare the price of consumption in a country relative to another. As with price indices within a country, PPPs can be thought of as statistical averages of prices, or a cost of living. They can be used to not only to study the relative size of different economies but also to measure global inequality or to convert the global poverty line to its local equivalent (Deaton and Dupriez, 2011).

There are, nonetheless, several issues in the collection of prices and in the methodology of PPPs. The main issue in the collection is the absence of expenditure weights below aggregated categories (basic headings) for most countries. As a result, in order to compute PPPs, detailed prices have to be aggregated up to basic categories. This produces important biases in the estimation of PPPs because lower quality items in poor countries are often matched to higher quality items in rich countries and because it is harder to accurately observe when the consumption patterns of different countries do not overlap. These issues are aggravated the further apart countries are in their level of development.

The second issue is methodological. The estimation of PPPs assumes identical homothetic tastes across countries. These two assumptions could importantly bias PPP estimates. First, the assumption that all countries have identical homothetic tastes is contradicted by much empirical evidence.[2] And, second, the assumption that tastes are identical across countries

[1] In 2016, for example, the US nominal GDP was 18.5 trillion dollars and China's was 11.3 trillion. That year, at common prices, the Chinese economy was larger that the US economy with a real GDP of about 21.5 trillion.

[2] See for example figure 2 for the case of Mexico. The figure shows the relationship of the relative prices paid with household income, also known as quality Engel curve. The curves are computed under the premise

implies that all differences in the patterns of demand must be attributable to international differences in the structure of relative prices. Given that relaxing these two assumptions has proven to be difficult, many economists assign PPPs a statistical rather than a welfare interpretation (Deaton and Heston, 2010).

In this paper, we tackle both sampling and methodological issues of the PPPs. To do so, we first construct a unique data set on prices and quantities at the barcode-level across two countries with different income levels: the United States and Mexico. We rely primarily on the Nielsen Homescan scanner dataset for the US and Mexico. The Nielsen Homescan for the US tracks the shopping behavior of 40,000 to 60,000 households every year in 54 Metropolitan Statistical Areas (MSA). 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 is grouped into 1,235 product modules 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. The Nielsen Homescan for Mexico is a representative sample of 5,000 households available for 2012-2013 that is collected using in-home scanners as well. It covers 60 cities across the country and, although it contains less product categories than the US data, it also covers around 35-40 percent of all expenditures on goods in the CPI.

The combination of these two datasets offers considerable advantages over the sample collected to compute PPPs. First, the dataset contains expenditures at the finest level of disaggregation so that the prices do not need to be aggregated to broader categories or across regions to compute price indices. Second, using the structure of the barcode of each item, we are able to identify more than 5,000 products that overlap across the two countries. This is relevant because it will allow us to estimate the bias that arises from matching goods of different qualities across countries. Third, since we use representative sample of households in each of the countries, our data is able to solve other issues in the collection of prices faced by the ICP such as the fact that in large countries some regions are not represented if they

that across households, at a point in time, those paying higher unit prices are buying higher quality goods (typically richer households).

do not sell the type of goods that meet the ICP specifications. And, lastly, given that our dataset contains demographic information of the consumer, we will be able to estimate the importance of the homothetic preference assumption.

We begin by describing the distribution of the set of prices available in each of the countries. There is larger dispersion in prices in the US relative to Mexico. This dispersion is significantly reduced when we limit the sample to goods overlapping across the two countries. Nonetheless, we show that even in the set of overlapping goods, the law of one price does not hold. Given that the relative prices differ in the two countries, we then use the ICP methodology to compute PPP estimates and compare them with those obtained using the same methodology but computed with the Nielsen data. The correlation between the PPP estimates of the ICP and those computed taking advantage of the granularity of the Nielsen data is very weak. We find evidence that heterogeneity in quality and variety are important sources of bias in comparisons across the two countries. In particular, we find the closer comparability of prices tends to increase the PPP of Mexico relative to the one of the US.

We then propose an income- and country-specific price index that takes into account this differences and uses a non-homothetic preference structure. Our price index encompasses versions of random utility models as in Redding and Weinstein (2018) and makes explicit the complementarity between quality and quantity of consumption we found in the data. In our setup, the vector of prices and quality along with the distribution of nominal income pin down the distribution of the price index. This is, we estimate the distribution of qualities to provide the best match between the predicted and observed expenditure shares for all the income groups in each country. We show that, for the set of overlapping goods, we do not find large differences in tastes across the two countries. After obtaining the distribution of prices and qualities, we perform two exercises: 1) calculate the level of utility for each US income level assuming that US consumers face Mexican prices, 2) calculate utility for each Mexican income level assuming that Mexican consumers face US prices.

This work is related to the literature studying international price and variety differences

(Hummels and Klenow, 2005; Gopinath, Gourinchas, Hsieh and Li, 2011; Simonovska, 2015). It is also related to recent works on price inequality across regions or households in the United States. Handbury and Weinstein (2015) use scanner data in the U.S. and find that price level for food products falls with city size, after eliminating quality heterogeneity and variety biases. We introduce cross-country scanner data and find that the same biases arise in international price comparison.

2 Data

2.1 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 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.[3]

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

[3]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).

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 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.[4]

2.2 Nielsen Mexico

The Nielsen data for Mexico shares many of the same features as the US data. It 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 item. The data contain around 55,000 distinct products grouped into 100 product categories 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

[4]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.

size.

Households are classified in 7 socieconomic 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 socio-economic 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.[5] 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. In the Appendix A we provide a more detailed description of each of the levels.

2.3 Linking Products Across Countries

In order to link products across countries, we take advantage of the fact that both the US and Mexico adopted the UPC system as they 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 an 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 barcodes for businesses in

[5]The variables used to construct the index are also available in our data.

the US and Mexico beginning with 00-139 and 750 respectively. [6] In our data, more than that 5,000 barcodes are consumed by households both in the US and Mexico. Our matched sample includes approximately 260 categories in the US and 62 in Mexico.

2.4 International Comparison Program (ICP)

We use restricted micro-data for Mexico and the United States from the 2011 ICP. The ICP is a worldwide statistical initiative led by the World Bank that collects and compares price data and GDP expenditures to estimate purchasing power parities (PPPs). The data is collected through partnerships with international, regional, sub-regional and national agencies. It contains 26 analytical categories (e.g. food and non-alcoholic beverages), 155 basic headings (e.g. rice, bread), and more than one thousand individual products (e.g. long grain rice, jasmine rice, basmati rice) for which we obtained the national average price.

In addition, the data contains the national accounts expenditures for each analytical category and basic heading. For the individual products, given that expenditure weights are not available, 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 its PPPs. When we link the categories of the 2011 ICP with those available in the Nielsen Mexico data we are able to match over 100 categories of the Nielsen data and 34 basic headings of the ICP representing around 45% of all products available in that data set.

[6] 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.

3 Stylized Facts

3.1 Distribution of Prices

In this section we compare the distribution of prices and varieties across the two countries. 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 75%/25% ratio, 90%/10%, mean/50%, and the kelly skewness. The statistics are computed within a nivel, a category of goods as defined by Nielsen Mexico, which broader than the usual categories (modules) defined by Nielsen US. The table contains information of all the UPCs (barcodes) belonging to the categories that are matched across the two countries and also reports the statistics for the set of matched barcodes.

Table 2 documents the differences in varieties across the two countries splitting goods in their income elasticity approximated by the share of high income consumers that buy the good in each product category. For each product category and country, we compute the number of different UPCs each household consumes. In addition, for each product category and using the households' income information, we compute the share of expenditure for consumers above the median income and below the median income. We rank the product categories according to this measure and split them in 4 buckets each containing the same amount of product categories. Compared to the US, Mexican households buy more varieties in necessary product categories, while they buy less varieties in luxury product categories.

3.2 Law of One Price

Figure 1 is an example of how law of one price holds for different barcodes withing the category of shampoo. In Table 3, we follow the approach by Gopinath, Gourinchas, Hsieh and Li (2011) and measure the 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. In particular, we focus on the first 6 months of 2012.

For country, we define a store as a city×retailer combination. 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. This is, for all the common products in each store pair, we compute the difference in the log price between the two stores. Table 3 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).

Table 4 describes the results of running the log of the prices in the US at the barcode level on the log of the price of the same product in Mexico. The table reports the coefficients when no intercept is allowed. This is, if the law of one price holds, then the coefficient of this regression must be equal to 1. We conduct this analysis as follws. Using the sample of UPCs matched across the two countries for both years, we compute for each country $\log p^k_{umrct}$ where p is the mean price of UPC u, in module m, sold in retailer r, in city c, at time t, in country k. Then we run a regression of $\log p^{US}_{umrct}$ on $\log p^{MEX}_{umrct}$ without intercept.

3.3 Price Index Comparison

In this section we construct two price indices. The first one follows the same methodology as the ICP. In order to validate our data, we want to show that, without taking advantage of the dissagregation in our data we could match the index in the ICP. We then construct a price index using the detailed information of our data on prices and expenditure weights. This will provide further evidence of the advantages of our data and will portray the quality and variety biases present in the ICP.

Then, we estimate exact price index (Feenstra, 1994):

$$\mathrm{EPI}_{\mathrm{mex,us}} = \left(\frac{\lambda_{\mathrm{mex,us}}}{\lambda_{\mathrm{us,mex}}}\right)^{\frac{1}{\sigma-1}} \prod_{k\in\Omega_{\mathrm{mex,us}}} \left(\frac{P_{k,\mathrm{mex}}}{P_{k,\mathrm{us}}}\right)^{\omega_{k,\mathrm{mex,us}}} \tag{1}$$

$$\text{where } \omega_{k,\text{mex,us}} = \frac{\dfrac{s_{k,\text{mex}} - s_{k,\text{us}}}{\ln s_{k,\text{mex}} - \ln s_{k,\text{us}}}}{\sum_{k \in \Omega_{\text{mex,us}}} \dfrac{s_{k,\text{mex}} - s_{k,\text{us}}}{\ln s_{k,\text{mex}} - \ln s_{k,\text{us}}}}$$

where $\lambda_{\text{mex,us}}$ adjusts a variety bias as a function of consumption baskets, while second term captures price difference of common goods with ideal weights.

Cross-country scanner data allows us to construct the exact price index, by comparing prices of common goods and by adjusting variety differences. [THIS IS WORK-IN-PROGRESS]

3.4 Non-Homotheticity

Lastly, in this section we provide evidence of the complementarity between quality and quantity present in the data. In Figure 2 we plot quality Engel curves for Mexico. Within a product category, high socioeconomic status households pay 10 percent higher price than low socioeconomic status households. Substantial amount of difference comes from the fact that high socioeconomic status households buy high quality products. With UPC fixed effects, high socioeconomic households pay 2 percent higher price than low socioeconomic status households to buy the same good.

4 Model

There are $n \in \{1, ..., N\}$ different product modules (milk, bread), and within a product module, there are $k \in 1, ..., K$ vertically differentiated goods (organic milk, normal milk). A household allocates its expenditure to n product modules, picks the best good per each product module, and spend all the money to purchase that good. Households are heterogeneous in income and idiosyncratic preference on differentiated goods. We partition consumers into different income groups indexed by $r \in \{1, ..., R\}$.

Given empirical evidences on the Quality Engel Curve (Bils and Klenow, 2001; Argente and Lee, 2016), we work with non-homothetic utility function. The price index with homothetic demand is shown in Appendix C.

4.1 Non-homothetic Demand with Multiplicative Utility Function

We introduce complementatity between quality and quantity of consumption as:

$$U_{ik}^r = q_k \log C_{ik}^r \epsilon_{ik}^r \tag{2}$$

where q_k captures the quality of product k and ϵ_{ik}^r represents idiosyncratic consumer tastes for each product. Each consumer i of type r, therefore, chooses C_{ik}^r units of good k to maximize utility. Since the consumer only consumes their preferred good, their budget constraint implies that $C_{ik}^r = E_i^r / P_k$, where E_i^r is the consumer's expenditure, and P_k is the price of the good. Recall that all consumers of the same type are assumed to have the same expenditure: $E_i^r = E^r$. We assume law of one price within a country where consumers pay the same prices to purchase the product.

Idiosyncratic tastes are assumed to have a Frechet (Type-II Extreme Value) distribution:

$$G(z) = e^{-z^{-\theta}} \tag{3}$$

where $\theta > 2$ determines the dispersion of idiosyncratic tastes.[7]

Then, the probability that an individual i of type r chooses product k represents the Quality Engel Curve:

$$S_{ik}^r = S_k^r = \frac{[q_k(\log E^r - \log p_k)]^\theta}{\sum_{j=1}^{M}[q_j(\log E^r - \log p_j)]^\theta} \tag{4}$$

[7] In order to capture Variety Engel Curve where high income households spread their expenditure across more varieties (Li, 2012), we could extend the model with Assumption 1 of Fajgelbaum, Grossman and Helpman (2011): θ is income-group specific and increasing in income group.

The expected utility of consumer i of type r is

$$E[U^r] = \Gamma(\frac{\theta-1}{\theta})[\sum_{j=1}^{M}\{q_j(\log E^r - \log p_j)\}^{\theta}]^{\frac{1}{\theta}} \tag{5}$$

where $\Gamma(.)$ is the Gamma function. This expected utility can be re-written as:

$$E[U^r] = \frac{E^r}{P^r} \tag{6}$$

where P^r is the unit expenditure function for consumers of type r:

$$P^r = E^r\frac{1}{E[U^r]} = E^r\Gamma(\frac{\theta-1}{\theta})^{-1}[\sum_{j=1}^{M}\{q_j(\log E^r - \log p_j)\}^{\theta}]^{-\frac{1}{\theta}} \tag{7}$$

which is income-group specific.[8]

5 Estimation

We begin by estimating the preference parameter, θ. Then, combining with data, we quantify distribution of product quality.

5.1 Preference Parameter Estimation: Homothetic Prefernce

We begin by estimation of θ under homothetic preference (equation 23 in the Appendix C), following Faber and Fally (2017).

$$\Delta \log s_{nkct} = -\theta\Delta \log p_{nkct} + \nu_{nct} + \epsilon_{nkct} \tag{8}$$

where as before n and k denote product modules and vertically differentiated goods (organic milk, normal milk). c and t indicate cities and four half-years (three changes), and s_{nkct} are budget shares within product module n. ν_{nct} are product module-by-city-by-half-year fixed

[8]Refer to Appendix B for detailed derivation.

effects. To address concerns about autocorrelation in the error term ϵ_{nkct} for the same city over time or within the city across modules, we cluster standard errors at the city level.

To address the standard simultaneity concern that taste shocks in the error term are correlated with observed price changes, we follow the empirical literature in industrial organization (e.g. Nevo (2000) and Hausman and Leibtag (2007)) and make the identifying assumption that consumer taste shocks are idiosyncratic across cities whereas supply-side cost shocks are correlated across space. For the supply-side variation needed to identify θ, we exploit the fact that store chains frequently price nationally or regionally without taking into consideration changes in local demand conditions. In particular, we instrument for local consumer price changes across brands $\Delta \log p_{nkct}$ with either national or state-level leave-out mean price changes.[9]

Table 5 and 6 shows estimation results. Table 5 trims top and bottom 1% of observations from all variables, and table 6 winsorize top and bottom 1% of observations from all variables. In all specifications, OLS estimates are upward biased. From our benchmark instrumental variable, state level leave-out mean price changes, θ is estimated at -1.263 from trimming and -0.353 from winsorizing.

5.2 Preference Parameter Estimation: Non-homothetic Preference

We then move to estimation of θ under non-homothetic preference. Equation 4 provides the following estimation equation:

$$\Delta \log s^r_{nkct} = \theta \Delta \log[\log(E^r_{nct}/p_{nkct})] + \nu_{nct} + \epsilon_{nkct} \tag{9}$$

[9] This shift-share style instrument has been widely used to isolate causal effect at the local level, tracing back to Bartik (1991) and then Beraja, Hurst and Ospina (2018); Faber and Fally (2017); Stroebel and Vavra (2016) with scanner data among others.

where as before n and k denote product modules and vertically differentiated goods (organic milk, normal milk). c and t indicate cities and four half-years (three changes), and s^r_{nkct} are budget shares of type r consumer within product module n. E^r_{nct} is the expenditure by type r for product module n in city c at time t. ν_{nct} are product module-by-city-by-half-year fixed effects. To address concerns about autocorrelation in the error term ϵ_{nkct} for the same city over time or within the city across modules, we cluster standard errors at the city level.

We begin by two nse groups: high nse (nse 1-4) and low nse (5-7) households. Refer to Appendix A for explanation on socioeconomic groups. [THIS IS WORK-IN-PROGRESS]

5.3 Calibration of Product Quality

In the case of homothetic demand, equation 23 exactly identifies product quality, given estimated θ. However, in the case of non-homothetic demand, the system from equation 4 is over-identified give estimated θ because product quality is not income-group specific. We calibrate θ by minimizing squared distance from predicted and actual budget shares for R (the number of consumer types) equations for each products.

From homothetic preference, product quality is deviation of item's budget share from its relative price. Suppose that budget shares are the same across items. Then, relative price represents item's quality, which is an assumption widely used in trade literature, using relative unit value as a proxy for the quality. [THIS IS WORK-IN-PROGRESS]

References

Argente, David and Munseob Lee, "Cost of Living Inequality during the Great Recession," 2016.

Bartik, Timothy J, *Who Benefits from State and Local Economic Development Policies?*, W.E. Upjohn Institute for Employment Research, 1991.

Beraja, Martin, Erik Hurst, and Juan Ospina, "The aggregate implications of regional business cycles," *NBER Working Paper No. 21956*, 2018.

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

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

_ and Olivier Dupriez, "Purchasing power parity exchange rates for the global poor," *American Economic Journal: Applied Economics*, 2011, *3* (2), 137–66.

Faber, Benjamin and Thibault Fally, "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data," *NBER Working Paper No. 23101*, 2017.

Fajgelbaum, Pablo, Gene M Grossman, and Elhanan Helpman, "Income Distribution, Product Quality, and International Trade," *Journal of Political Economy*, 2011, *119* (4), 721–765.

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

Gopinath, Gita, Pierre-Olivier Gourinchas, Chang-Tai Hsieh, and Nicholas Li, "International prices, costs, and markup differences," *American Economic Review*, 2011, *101* (6), 2450–86.

Handbury, Jessie and David E Weinstein, "Goods Prices and Availability in Cities," *Review of Economic Studies*, 2015, *82*, 258–296.

Hausman, Jerry and Ephraim Leibtag, "Consumer benefits from increased competition in shopping outlets: Measuring the effect of Wal-Mart," *Journal of Applied Econometrics*, 2007, *22* (7), 1157–1177.

Hummels, David and Peter J Klenow, "The variety and quality of a nation's exports," *American Economic Review*, 2005, *95* (3), 704–723.

Li, Nicholas, "An Engel Curve for Variety," 2012.

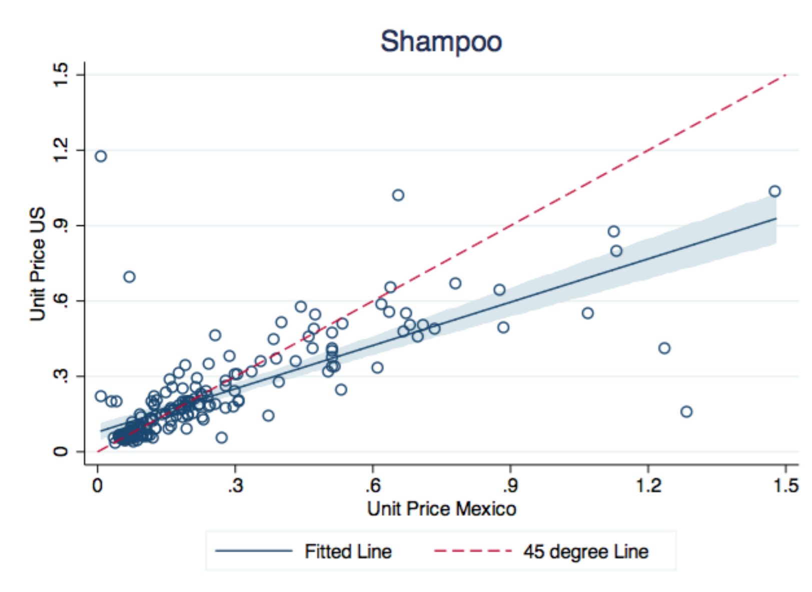
Nevo, Aviv, "Mergers with differentiated products: The case of the ready-to-eat cereal industry," *The RAND Journal of Economics*, 2000, pp. 395–421.

Redding, Stephen J and David E Weinstein, "Measuring Aggregate Price Indexes with Demand Shocks: Theory and Evidence for CES Preferences," *NBER Working Paper No. 22479*, 2018.

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

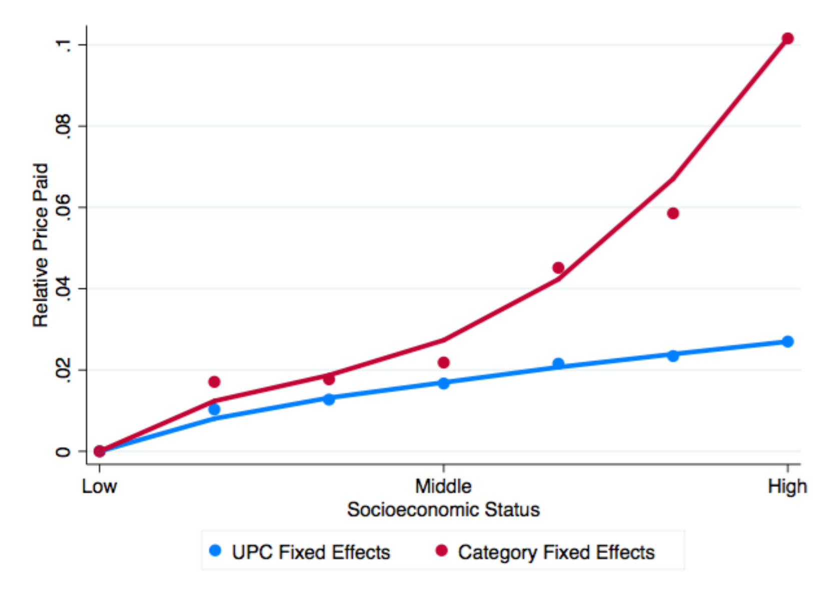
Stroebel, Johannes and Joseph Vavra, "House prices, local demand, and retail prices," *NBER Working Paper No. 20710*, 2016.

Figure 1: Law of One Price (Shampoo)



The figure plots the law of one price for the category shampoo.

Figure 2: Quality Engel Curve in Mexico



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 income category dummies and product category, region, chain, quarter and household fixed effects.

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, 90%/10% ratio, mean/50% ratio, and the kelly skewness.

Country	Sample	75%/25%	90%/10%	Avg/50%	Kelley Skew
US	All	-.26	-.65	.92	-1.01
Mexico	All	-.44	-.61	1.74	-.28
US	Matched	-.3	-.69	.96	-.47
Mexico	Matched	-.4	-.67	1.04	.61

Table 2: Number of Varieties - Luxury vs Inferior Goods

The table reports the differences in varieties across Mexico and the US splitting goods in their income elasticity approximated by the share of high income consumers that buy the good in each product category. Q4 represents luxury product categories, and Q1 represents inferior product categories.

	US	Mexico
Q1	7.34	16.37
Q2	5.52	8.96
Q3	8.59	8.49
Q4	6.94	4.05
Average	7.12	9.33

Table 3: Deviations from the 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.

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	4.644	4.635	4.845
Median	4.589	4.582	4.688
Std. Dev.	.77	.764	.974

Table 4: Law of One Price (No Intercept)

The table reports the results of running the log of the prices in the US at the barcode level on the log of the price of the same product in Mexico.

	(1) log p_{umt}^{US}	(2) log p_{umt}^{US}	(3) log p_{umt}^{US}	(4) log p_{umt}^{US}	(5) log p_{umt}^{US}
log p_{umt}^{MEX}	-0.534	0.481	0.597	0.483	0.491
	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)
Observations	7382	7382	7382	7382	7382
Adjusted R^2	0.362	0.826	0.468	0.826	0.826
Nivel FE	No	Yes	No	Yes	No
Time FE	No	No	Yes	Yes	No
Nivel x Time FE	No	No	No	No	Yes

Standard errors in parentheses

Table 5: Preference Parameter Estimation: Homothetic Demand (Trimming top and bottom 1% of obs.)

The table reports the coefficients of OLS and IV regressions. The dependent variable is change in budget shares of products within product module. The main independent variable is change in average prices of products. Controls include module-by-city-by-half-year fixed-effects. Standard errors are clustered at city level and presented in parentheses. ***, **, and * represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var.: $\Delta \log s_{nkct}$	(1) OLS	(2) IV: Nat'l change	(3) IV: Leave-out
$\Delta \log p_{nkct}$	0.657***	0.187***	-1.263***
	(0.023)	(0.028)	(0.349)
Observations	438,489	434,491	423,576
R-squared	0.044	0.037	-0.062
module-by-city-by-half-year FEs	Yes	Yes	Yes

Table 6: Preference Parameter Estimation: Homothetic Demand (Winsorizing top and bottom 1% of obs.)

The table reports the coefficients of OLS and IV regressions. The dependent variable is change in budget shares of products within product module. The main independent variable is change in average prices of products. Controls include module-by-city-by-half-year fixed-effects. Standard errors are clustered at city level and presented in parentheses. ***, **, and * represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var.: $\Delta \log s_{nkct}$	(1) OLS	(2) IV: Nat'l change	(3) IV: Leave-out
$\Delta \log p_{nkct}$	0.842***	0.554***	-0.353**
	(0.015)	(0.027)	(0.145)
Observations	456,099	455,760	446,903
R-squared	0.053	0.050	0.008
module-by-city-by-half-year FEs	Yes	Yes	Yes

APPENDIX – FOR ONLINE PUBLICATION ONLY

A Socieconomic Classification According to AMAI

Socieconomic classicification in the Nielsen data has seven categories, and it is widely used in other surveys in Mexico. It is known that the highest socio-economic level is made up mainly (82%) of households in which the head of the family has professional studies. 98% of these households have a fixed Internet in the home. It is the level that invests the most in education (13% of its expenditure) and also the one that spends the smallest proportion in food (25%). The vast majority of households in the lowest level (95%) have a head of family with no higher education than primary education. The possession of fixed internet in the house is minimal (0.1%).

B Proof of Non-homothetic Price Index

We want to find the probability that an individual of type i in income group r chooses product k. Given that the idiosyncratic consumer taste ϵ_{ik}^r is distributed according to a Frechet distribution, the distribution of utilities obtained by an individual of type i in income group r from product k is dercribed in the following equation:

$$G_{ik}^r = e^{-[q_k(\log E^r - \log P_k)]^\theta (U^r)^{-\theta}}, \tag{10}$$

where $\log E^r - \log P_k$ comes from the definition of utility and the fact that $p_k C_{ik}^r = E_i^r$.

The probability that an individual of type i in income group r chooses product k is then equal to the probability that product k delivers the highest utility in the set, given that the individual consumes only their prefered good. Therefore, we are looking for the following expression:

$$
\begin{aligned}
S^r_{ik} &= \Pr\Big[U^r_{ik} \geqq \max\{U^r_{ij}\}, \forall j \neq k\Big] \\
&= \int_0^\infty \Big[\prod_{j\neq k} G^r_{ij}(U^r)\Big] g^r_{ik}(U^r)\, dU^r
\end{aligned} \tag{11}
$$

In this case, $g^r_{ik}(U^r)$ is the defivative of G^r_{ik} and is equal to the following expression:

$$
g^r_{ik}(U^r) = [q_k(\log E^r - \log P_k)]^\theta \theta (U^r)^{-(1+\theta)} e^{-[q_k(\log E^r - \log P_k)]^\theta (U^r)^{-\theta}} \tag{12}
$$

Substiting 12 and 10 into 11, we get the following expression for S^r_{ik}:

$$
S^r_{ik} = \int_0^\infty \Big[\prod_{j\neq k} e^{-[q_j(\log E^r - \log P_j)]^\theta (U^r)^{-\theta}}\Big][q_k(\log E^r - \log P_k)]^\theta \theta (U^r)^{-(1+\theta)} e^{-[q_k(\log E^r - \log P_k)]^\theta (U^r)^{-\theta}}\, dU^r
$$

Notice that we can collect the exponent term from $g^r_{ik}(U^r)$ into the product to get the simplified expression:

$$
S^r_{ik} = \int_0^\infty \Big[\prod_{j} e^{-[q_j(\log E^r - \log P_j)]^\theta (U^r)^{-\theta}}\Big][q_k(\log E^r - \log P_k)]^\theta \theta (U^r)^{-(1+\theta)}\, dU^r \tag{13}
$$

where the product is now defined over all j in the set of products.

Using the laws of indices and noting that U^r is independent of j, we can simplify the first term of hte integral:

$$
\prod_j e^{-[q_j(\log E^r - \log P_j)]^\theta (U^r)^{-\theta}} = e^{-(U^r)^\theta \sum_j [q_j(\log E^r - \log P_j)]^\theta}
$$

For notation simplicity define Y^r:

$$
Y^r \equiv \sum_j [q_j(\log E^r - \log P_j)]^\theta \tag{14}
$$

Using 14, we get:

$$S^r_{ik} = \int_0^\infty e^{-Y^r(U^r)^\theta}[q_k(\log E^r - \log P_k)]^\theta \theta (U^r)^{-(1+\theta)}\, dU^r$$
$$= [q_k(\log E^r - \log P_k)]^\theta \int_0^\infty e^{-Y^r(U^r)^\theta} \theta (U^r)^{-(1+\theta)}\, dU^r \tag{15}$$

as $[q_k(\log E^r - \log P_k)]^\theta$ is constant in U^r.

Finally, we are ready to solve the integral:

$$\begin{aligned} S^r_{ik} &= [q_k(\log E^r - \log P_k)]^\theta \Big[\frac{1}{Y^r} e^{-Y^r(U^r)^\theta}\Big]_0^\infty \\ &= \frac{[q_k(\log E^r - \log P_k)]^\theta}{Y^r} \\ &= \frac{[q_k(\log E^r - \log P_k)]^\theta}{\sum_j [q_j(\log E^r - \log P_j)]^\theta} \end{aligned} \tag{16}$$

We can also define the distribution of U_r, the final utility value that the individual in group r get from comsuming their preferred good from the set $1, ..., k$:

$$\begin{aligned} G(U^r) = \prod_j G^r_{ij}(U^r) &= \prod_j e^{-[q_k(\log E^r - \log P_j)]^\theta (U^r)^{-\theta}} \\ &= e^{-(U^r)^{-\theta} \sum_j [q_k(\log E^r - \log P_j)]^\theta} \\ &= e^{-Y^r(U^r)^{-\theta}} \end{aligned} \tag{17}$$

Then we are ready to define the expected utility of an individual from income group r who is choosing a good from the set $\{1, ..., k\}$:

$$\begin{aligned} E[U^r] &= \int_0^\infty U^r g(U^r)\, dU^r \\ &= \int_0^\infty \theta Y^r (U^r)^{-\theta} e^{-Y^r(U^r)^{-\theta}}\, dU^r \end{aligned} \tag{18}$$

In order to solve this integral, introudce a new variable $z \equiv Y^r(U^r)^{-\theta}$ with $dz = -\theta Y^r (U^r)^{-(1+\theta)}\, dU^r$.

Then:

$$
\begin{aligned}
E[U^r] &= \int_0^\infty \theta Y^r (U^r)^{-\theta} e^{-Y^r (U^r)^{-\theta}}\, dU^r \\
&= \int_0^\infty (Y^r)^{\frac{1}{\theta}} z^{-\frac{1}{\theta}} e^{-z}\, dz \\
&= (Y^r)^{\frac{1}{\theta}} \int_0^\infty z^{-\frac{1}{\theta}} e^{-z}\, dz
\end{aligned} \tag{19}
$$

Notice that the term inside the brackets can be represented as Gamme function with parameter $\frac{\theta-1}{\theta}$. Therefore, the expected utility is:

$$
\begin{aligned}
E[U^r] &= (Y^r)^{\frac{1}{\theta}} \int_0^\infty z^{-\frac{1}{\theta}} e^{-z}\, dz \\
&= (Y^r)^{\frac{1}{\theta}} \Gamma\Big(\frac{\theta-1}{\theta}\Big) \\
&= \Gamma\Big(\frac{\theta-1}{\theta}\Big) \Big[\sum_j [q_j(\log E^r - \log P_j)]^\theta\Big]^{\frac{1}{\theta}}
\end{aligned} \tag{20}
$$

C Homothetic Demand

This subsection replicates key results of Redding and Weinstein (2018) in the environment where products are vertically differentiated by intrinsic quality.[10]

The utility of an individual i of type r who consumes C_{ik}^r units of product k is:

$$
U_{ik}^r = q_k C_{ik}^r \epsilon_{ik}^r \tag{21}
$$

where q_k captures the quality of product k and ϵ_{ik}^r represents idiosyncratic consumer tastes for each product. Each consumer i of type r, therefore, chooses C_{ik}^r units of good k to maximize utility. Since the consumer only consumes their preferred good, their budget constraint implies that $C_{ik}^r = E_i^r / P_k$, where E_i^r is the consumer's expenditure, and P_k is the price of the good. Recall that all consumers of the same type are assumed to have the same expenditure:

[10] Redding and Weinstein (2018)'s model represents horizontal differentiation where type-r consumers have common tastes for product, which doesn't systematically vary across income groups.

$E_i^r = E^r$. We assume law of one price where consumers pay the same prices to purchase the product.

Idiosyncratic tastes are assumed to have a Frechet (Type-II Extreme Value) distribution:

$$G(z) = e^{-z^{-\theta}} \tag{22}$$

where $\theta > 2$ determines the dispersion of idiosyncratic tastes.

Using the monotonic relationship between idiosyncratic tastes and utility, the probability that an individual i of type r chooses product k is the same across all individuals of that type and equal to:

$$S_{ik}^r = S_k = \frac{(p_k/q_k)^{-\theta}}{\sum_{l=1}^{M}(p_l/q_l)^{-\theta}} \tag{23}$$

which is not income-group specific.

The expected utility of consumer i of type r is

$$E[U^r] = \Gamma(\frac{\theta-1}{\theta})[\sum_{k=1}^{M}(E^r)^\theta(\frac{p_k}{q_k})^{-\theta}]^{\frac{1}{\theta}} \tag{24}$$

where $\Gamma(.)$ is the Gamma function. This expected utility can be re-written as:

$$E[U^r] = \frac{E^r}{P^r} \tag{25}$$

where P^r is the unit expenditure function for consumers of type r:

$$P^r = \Gamma(\frac{\theta-1}{\theta})^{-1}[\sum_{k=1}^{M}(\frac{p_k}{q_k})^{-\theta}]^{-\frac{1}{\theta}} \tag{26}$$

Therefore, the price index is the same across different types of consumers.