What is the Price of Tea in China? Towards the Relative Cost of Living in Chinese and U.S. Cities[*]

Robert C. Feenstra
University of California, Davis

Mingzhi Xu
University of California, Davis

Alexis Antoniades
University of Georgetown

November 28, 2016, PRELIMINARY DRAFT

Abstract

We examine the price and variety of products at the barcode level in cities within China and the United States. In both countries, there is a greater variety of products in larger cities. But in China, unlike the United States, the prices of products tend to be lower in larger cities. We attribute the lower prices to a pro-competitive effect, whereby large cities attract more firms which leads to lower markups and prices. Combining the effect of greater variety and lower prices, it follows that the cost-of-living for grocery-store products in China is lower in larger cities. We further compare the cost-of-living indexes for particular products *between* China and the United States. In products with a significant presence of U.S. brands in the Chinese market, the availability of additional Chinese brands leads to greater variety than in the United States, and therefore lower Chinese price indexes for that reason. In products with much less presence of U.S. brands in the Chinese market, however, the observed prices differences between the countries (usually lower prices in China) are partially or fully offset by the variety differences (less variety in China), so that the cost of living in China is not as low as the price differences suggest, especially in smaller cities.

1

1 Introduction

In 2005, real GDP for the Chinese economy fell by 40%. Not in reality, of course, but in the estimates reported by the World Bank. As explained by Deaton and Heston (2010), the revised 2005 estimates made use of new price data collected for China under the International Comparisons Program (ICP):

> . . . the 2007 version of the World Development Indicators (WDI), World Bank (2007), lists 2005 per capita GDP for China as $6,757 . . . in current international dollars. The 2008 version, World Bank (2008), which includes the new [2005] ICP data, gives, for the same year, and the same concept $4,088 for China ...[1]

The logical explanation for this fall in estimated real GDP for China is that the prices collected for China were higher than expected. In fact, prices had never been collected for China before the 2005 round of the ICP, so prior estimates of real GDP used imputed prices. Because the actual Chinese prices were higher than imputed, then the quantity of goods consumed by the representative Chinese individual were lower. It follows that real GDP was also lower – 40% lower for China!

Two explanations have been provided for the unexpectedly high prices collected for China by the ICP 2005. The first explanation is that urban regions were over-sampled in China and that rural prices would be lower. That claim is evaluated by Feenstra, Ma, Peter Neary, and Prasada Rao (2013), who find that Chinese prices from the 2005 ICP round were indeed somewhat higher than for other developing countries at similar levels of GDP per capita. Besides the possible urban bias, another explanation has to do with a special feature of the 2005 ICP, whereby prices were collected in each region of the world and then "linked" to other regions using a different price survey conducted only in certain cities. Deaton and Aten (2014) and Inklaar and Rao (2015) argue that this "linking" procedure led to a systematic upward bias in the prices for Asia relative to the United States. That linking procedure was not followed in the 2011 round of the ICP, and as a result, relative prices in Asia and in other developing countries were lower than in 2005.

These facts, combined with anecdotal evidence of high prices in China,[2] are enough convince us that a source of price information for China independent of the ICP is very important; crucial, in fact, to obtain reliable estimates of its real GDP. The goal of this paper is to compare the cost of living for cities in China and in the United States using two sources of barcode data: scanner data from grocery stores; and prices for grocery-store products scraped from the web. In scraping prices from the web we are following the lead of Cavallo and Rigobon (2016) who collect prices from internet sources to make time-series and cross-country comparisons. Such scraped data alone it not well-suited to compute cost-of-living indexes, however, because there is no quantity or expenditure data available with the scraped prices. While expenditure data is available from scanner data, it is prohibitively expensive to obtain for China in more than a handful of products. Accordingly, in this paper we rely on Chinese scanner data for four products in 22 cities, purchased from Nielsen (China), and have supplemented these data with scraped prices for these same four products and a further 15 products over 60 cities, obtained from a mobile phone application that allows consumers to check

---

[1]Deaton and Heston (2010), p. 3. The references given in this quote are to the World Development Indicators.

[2]Anecdotal evidence that prices in China are high comes from Chinese students in the United States who return home and, finding a local price for an item that is higher than the U.S. price at the official exchange rate, will photograph that item and post it on their Facebook page!

for the various prices at supermarkets in each city. These Chinese price data are complemented by Nielsen (U.S.) barcode data at the city level. All data are for 2011 or more recent years.

In our results, we find that in both countries there is a greater variety of products in larger cities. That result has already been found by Handbury and Weinstein (2015) for the United States, and this paper is the first confirmation of the same result for China. For the U.S., the greater variety in larger cities offsets the higher prices found there, so that Handbury and Weinstein (2015) argue that the cost of living – incorporating both price and variety – is lower in larger cities. But in China, unlike the United States, the prices of products themselves tend to be *lower* in larger cities. We attribute those lower prices to a pro-competitive effect, whereby large cities attract more firms which leads to lower markups and prices. Combining the effect of greater variety and lower prices, it follows that the cost of living for grocery-store products in China is also lower in larger cities.

We further compare the cost-of-living indexes for particular products *between* China and the United States. In the four products for which we purchased Nielsen (China) scanner data, and supplemented it with scraped prices, we expected a significant presence of U.S. brands in the Chinese market and that turned out to be the case. Then the availability of additional Chinese brands leads to greater variety than in the United States, and therefore lower Chinese price indexes for that reason. In the other 15 products, however, there is much less presence of U.S. brands in the Chinese market. In these cases, the observed prices differences between the countries (usually lower prices in China) are partially or fully offset by the variety differences (less variety in China), so that the cost of living in China is not as low as the price differences suggest. This latter case applies to Tea, for example, because our data do not include the many Chinese varieties of tea that are not sold in supermarkets.

In section 2 we describe the nested CES framework that we shall use to measure the consumer gains from variety. Preliminary evidence from the four products with both Nielsen (China) barcode data and scanned prices is presented in section 3, where we describe the data sources in more detail. A model of multi-product firms in described in section 4, from which we derive the solutions for firm pricing and product scope. These equations are estimated as regressions described in section 5. Our empirical results on the prices of 15 other products is presented in section 6, and section 7 concludes.

2 Consumer Utility and Variety

We study an economy consisting of $c = 1, ..., D$ cities or destinations, which differ in population L_c and labor income w_c. In each city, labor is the only factor of production and it is not mobile across regions. A fraction ρ of total labor income is spent on the differentiated goods, and we denote that spending by $Y_c = \rho w_c L_c$.

The preferences of the representative consumer in each city are nested CES. Denoting the set of product varieties sold by firm f in city c by $i \in I_{fc}$, the sub-utility from the products of firm f are given by

$$X_{fc} = \left(\sum_{i \in I_{fc}} (b_{fic} x_{fic})^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}, \ \sigma > 1, \tag{1}$$

where σ is the elasticity of substitution across products sold by a firm, and b_{fic} are the taste parameters for each variety, which we will allow to differ across firms, products and to a certain extent

across cities as explained below. Aggregating across firms $f \in F_c$ that sell in city c, utility of the representative consumer is,

$$U_c = \left(\sum_{f \in F_c} X_{fc}^{(\eta-1)/\eta} \right)^{\eta/(\eta-1)}, \quad \eta > 1. \tag{2}$$

As in Hottman, Redding, and Weinstein (2016), we expect that the elasticity of substitution across products is larger within firms than across firms, so we assume that $\sigma \geq \eta > 1$. When the two elasticities are equal, then the nested CES system will collapse to a standard CES utility function as used in Feenstra (1994).

Let $\mathbf{p}_{fc}$ denote the vector of prices p_{fic} for firm f with the vector of taste parameters $\mathbf{b}_{fc}$, and let $P_{fc} = e(\mathbf{p}_{fc}, \mathbf{b}_{fc}, I_{fc})$ denote the minimum expenditure needed to obtain one unit of sub-utility, $X_{fc} = 1$. Then $e(\mathbf{p}_{fc}, \mathbf{b}_{fc}, I_{fc})$ takes on the CES form,

$$P_{fc} = e(\mathbf{p}_{fc}, \mathbf{b}_{fc}, I_{fc}) = \left(\sum_{i \in I_{fc}} (p_{fic}/b_{fic})^{1-\sigma} \right)^{1/(1-\sigma)}. \tag{3}$$

Feenstra (1994) shows how to measure the effect of new varieties on the exact price index. Specifically, consider firm f selling to destinations c and d. Suppose that there is a non-empty subset of "common" products $I_f \subseteq I_{fc} \cap I_{fd}$ sold by firm f in these two cities for which the taste parameters are equal, $b_{fic} = b_{fid}$, $i \in I$. Then the exact price index between the two cities can be expressed as

$$\frac{e(\mathbf{p}_{fc}, \mathbf{b}_{fc}, I_{fc})}{e(\mathbf{p}_{fd}, \mathbf{b}_{fd}, I_{fd})} = \left[\prod_{i \in I_f} \left(\frac{p_{fic}}{p_{fid}} \right)^{w_{fi}(I_f)} \right] \left(\frac{\lambda_{fc}}{\lambda_{fd}} \right)^{\frac{1}{\sigma-1}}. \tag{4}$$

The first term in brackets on the right of (4) is the Sato (1976)-Vartia (1976) price index, where $w_{fi}(I_f)$ is the weight defined by:

$$w_{fi}(I_f) \equiv \frac{\frac{s_{fic}(I_f) - s_{fid}(I_f)}{\ln s_{fic}(I_f) - \ln s_{fid}(I_f)}}{\sum_{j \in I_f} \left(\frac{s_{fjc}(I_f) - s_{fjd}(I_f)}{\ln s_{fjc}(I_f) - \ln s_{fjd}(I_f)} \right)}, \quad s_{fic}(I_f) \equiv \frac{p_{fic} x_{fic}}{\sum_{j \in I_f} p_{fjc} x_{fjc}}, \tag{5}$$

and likewise for the shares $s_{fid}(I_f)$ in city d, also defined over the common set of products. The second term on the right of (4) is the adjustment needed to take into account differing sets of goods available in the two cities, and is defined by:

$$\lambda_{fc} \equiv \frac{\sum_{i \in I_f} p_{fic} x_{fic}}{\sum_{i \in I_{fc}} p_{fic} x_{fic}} = 1 - \frac{\sum_{i \in I_{fc} \setminus I_f} p_{fic} x_{fic}}{\sum_{i \in I_{fc}} p_{fic} x_{fic}}. \tag{6}$$

To interpret these formulas, λ_{fc} in (6) denotes the spending in city c on the common products of firm f, sold in both cities, relative to total spending in city c on firm f's products. Equivalently, it equals one minus the share of expenditure on the unique products sold by firm f only in city c. Having access to more unique varieties in city c implies a smaller expenditure share on common products, λ_{fc}, and a lower cost-of-living index in (4).

To extend the exact price index to the nested CES case, let $\mathbf{P}_c$ denote the vector of CES price indexes P_{fc} shown in (3) for all firms $f \in F_c$ in city c, and let $P_c = E(\mathbf{P}_c, F_c) = \left(\sum_{f \in F_c} P_{fc}^{(\eta-1)} \right)^{1/(1-\eta)}$ denote the expenditure needed to obtain utility of one in city . Let $F \equiv F_c \cap F_d$ denote the non-empty

set of "common" firms selling to both cities c and d. Then again from Feenstra (1994), the cost of living between the two cities can be written as,

$$
\begin{aligned}
\frac{E(\mathbf{P}_c, F_c)}{E(\mathbf{P}_d, F_d)} &= \left[\prod_{f \in F} \left(\frac{P_{fc}}{P_{fd}} \right)^{W_f(F)} \right] \left(\frac{\lambda_c}{\lambda_d} \right)^{\frac{1}{\eta-1}} \\
&= \left[\prod_{f \in F} \prod_{i \in I_f} \left(\frac{p_{fic}}{p_{fid}} \right)^{W_f(F) w_i(I_f)} \right] \left[\prod_{f \in F} \left(\frac{\lambda_{fc}}{\lambda_{fd}} \right)^{W_f(F)} \right]^{\frac{1}{\sigma-1}} \left(\frac{\lambda_c}{\lambda_d} \right)^{\frac{1}{\eta-1}}
\end{aligned}
\tag{7}
$$

where the Sato-Vartia weights across firms are,

$$
W_f(F) \equiv \frac{\frac{S_{fc}(F) - S_{fd}(F)}{\ln S_{fc}(F) - \ln S_{fd}(F)}}{\sum_{g \in F} \left(\frac{S_{gc}(F) - S_{gd}(F)}{\ln S_{gc}(F) - \ln S_{gd}(F)} \right)}, \quad S_{fc}(F) \equiv \frac{\sum_{i \in I_{fc}} p_{fic} x_{fic}}{\sum_{g \in F} \sum_{i \in I_{gc}} p_{gic} x_{gic}},
\tag{8}
$$

and likewise for the shares $S_{fd}(F)$ for city d, also defined over the common set of firms. The final term on the right of (7) is defined by:

$$
\lambda_c \equiv \frac{\sum_{g \in F} \sum_{i \in I_{gc}} p_{gic} x_{gic}}{\sum_{g \in F_c} \sum_{i \in I_{gc}} p_{gic} x_{gic}} = 1 - \frac{\sum_{g \in F_c \setminus F} \sum_{i \in I_{gc}} p_{gic} x_{gic}}{\sum_{g \in F_c} \sum_{i \in I_{gc}} p_{gic} x_{gic}},
\tag{9}
$$

That is, λ_c denotes the spending on the common set of firms F relative to total spending in city c, or one minus the share of spending on firms selling only in city c. The greater the share of spending on unique firms selling only that city, the lower is the exact price index in (7).

3 Evidence on Prices for Four Products

3.1 Data

Our calculation of the cost-of-living relies on the data extracted from three sources. The first source is the Nielsen (China) Sales database which enables us to observe the annual sales and average price information of each product with a bar-code.[3] Significantly, the barcode systems used in China is the European Article Number (EAN-13), whereas that used in the United States in the Universal Product Code (UPC). While it is not difficult to identify similar product categories, it is impossible to exactly match the products across countries by their barcodes since the product descriptions themselves differ. That difference in the barcode systems means that the consumer theory outlined above due to Feenstra (1994) cannot be applied *between* countries, since it assumed that there is a non-empty subset of "common" products $I_f \subseteq I_{fc} \cap I_{fd}$ sold by firm f in two cities c and d for which the taste parameters are equal, $b_{fic} = b_{fid}$, $i \in I$. We certainly cannot assume that the tastes for products is the same when we cannot even match products across barcode systems. Accordingly, for the next several sections we focus on comparisons of price across cities *within* China and the U.S. It will not be until section 6 that we attempt to compare across countries, using an alternative formulation of exact price indexes for a nested CES utility function due to Redding and Weinstein (2016).

---

[3]We convert RMB data to US$ using the annual average exchange rate.

We purchased Nielsen (China) scanner data for four product categories: Toothpaste, Laundry Detergent, Personal Wash items, and Shampoo, covering 22 cities in China, as shown in the red colored regions of Figure 1.[4] Besides the sales information for each product, we also observe manufacturer information such as brand and sub-brand of each product. In our theory and empirical analysis, we refer to the brand as the firm (e.g. Crest or Colgate for toothpaste).
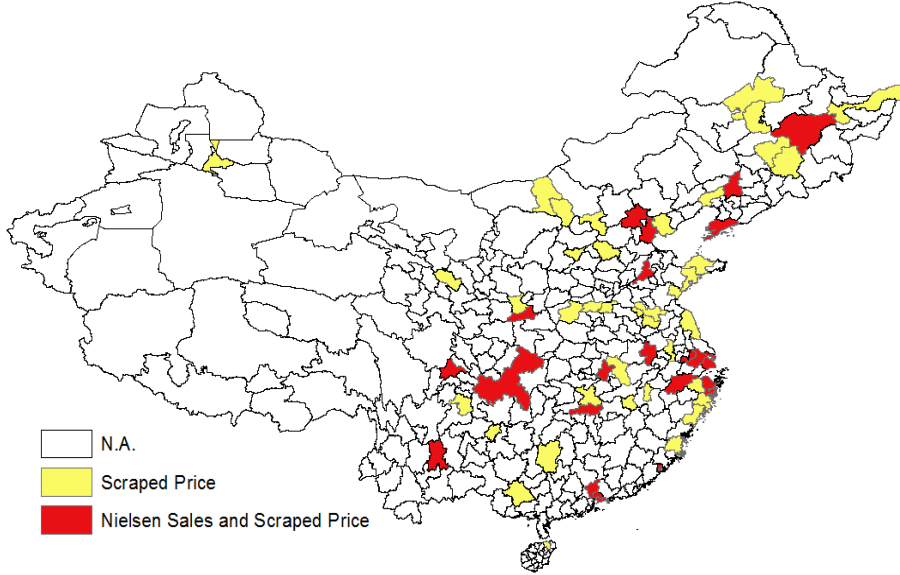

Figure 1: Regions included in Nielsen Sales Data and Scraped Price Data

The regions covered by Nielsen (China) database are large cities (most of them are capital cities). To address this issue, our analysis also relies on our second data source, which are scraped prices collected in 2015 from a mobile phone application that allows consumers to check for the various prices at supermarkets in each city. Details on this second source of data are provided in Appendix A, and it allows us also to include smaller cities in our sample. We end up expanding our sample to 60 cities, including the 22 cities provided by Nielsen (China) database, as shown in the yellow colored regions of Figure 1. Based on the first two data sources, we implement the formulas described above to calculate the cost of living in China for each of the four product categories. Our third source of data is the Nielsen Retailing Sales (U.S.) database, used to calculate the cost of living for 377 MSAs of the United States in 2013, for each of the four product categories.

We have prices for individual barcode products in four categories for 60 cities in China, but we have expenditure on these barcode products for only 22 cities in the Nielsen (China) database. We have implemented a Heckman procedure to estimate barcode-level expenditure in the remaining 38 cities, using a reduced-form equation for the barcode expenditure shares in the 22 cities for which we have those data. This Heckman procedure is described in Appendix B. The estimated expenditure shares are used to measure the terms λ_{fc} and λ_f appearing in the cost of living index (7).

Data are also needed on the elasticities σ between products within a firm, and η between firms. These elasticities of substitution differ across the four product categories. We rely on the estimates of

---

[4]We use 2011 and 2012 sales information for toothpaste, and 2014 for the other three product categories.

these elasticities from Hottman, Redding, and Weinstein (2016), as shown in Appendix B. These authors estimate the elasticities from the Nielsen Retailing Sales (U.S.) database, which is the same data that we use for the United States. Our assumption is that, for each product category, the elasticities are the same in the United States and China.

3.2 Prices Across Chinese and U.S. Cities

Panel (a) of Figures 2 to 5 exhibits how the average price of the common products vary with respect to the city population for Toothpaste, Laundry Detergent, Personal Wash Items and Shampoo. We measure the price index in terms of dollars per ounce and calculate the average price using the Sato-Vartia weights, which is the numerator of the Sato-Vartia index in (7), and we calculate the numerator of the variety adjustment in (7):

$$P_c^{SV} \equiv \prod_{f \in F} \prod_{i \in I_f} (p_{fic})^{W_f(F) w_i(I_f)}, \ \ \Lambda_c^{SV} \equiv \left[\prod_{f \in F} \lambda_{fc}^{W_f(F)} \right]^{\frac{1}{\sigma - 1}} \left(\lambda_c \right)^{\frac{1}{\eta - 1}}. \tag{10}$$

The terms P_c^{SV} and Λ_c^{SV} are shown in panels (a) and (b) of Figure 2 for Toothpaste, Figure 3 for Laundry Detergent, Figure 4 for Personal Wash items, and Figure 5 for Shampoo. In each panel and each Figure, we plot the variables for the United States (in blue dots) and China (in red x) against the log population of each city.

From panel (a), we observe that average prices P_c^{SV} within the United States there is no significant price difference between cities, whereas average prices in China have a clearly negative relationship with city size. We explore that finding more in the next section, where we argue that a pro-competitive effect in larger cities in China can account for that result.[5] In panel (b) we see that the term Λ_c^{SV}, which measures the inverse of variety, falls slightly with city size in the United States (though this is barely visible), but falls very noticably with city size in China.

Multiplying the terms $P_c^{SV} \times \Lambda_c^{SV}$, and choosing the normalization $P_{NY}^{SV} \Lambda_{NY}^{SV} = 1$ for New York, we obtain the exact price index for the United State shown in panel (c) of each Figure. These exact price indexes are slightly declining in city size for Toothpaste, Laundry Detergent, and Personal Wash items, but not for Shampoo. In the first three products, we therefore confirm the finding by Handbury and Weinstein (2015) for the United States that the greater variety in larger cities offsets the higher prices found there. But in China, unlike the United States, the prices of products themselves tend to be lower in larger cities, and that is reinforced by a much more greater increase in vartiety in larger cities, too. For both reasons, the cost of living indexes $P_c^{SV} \Lambda_c^{SV}$ for China decline much more rapidly with population than in the United States.

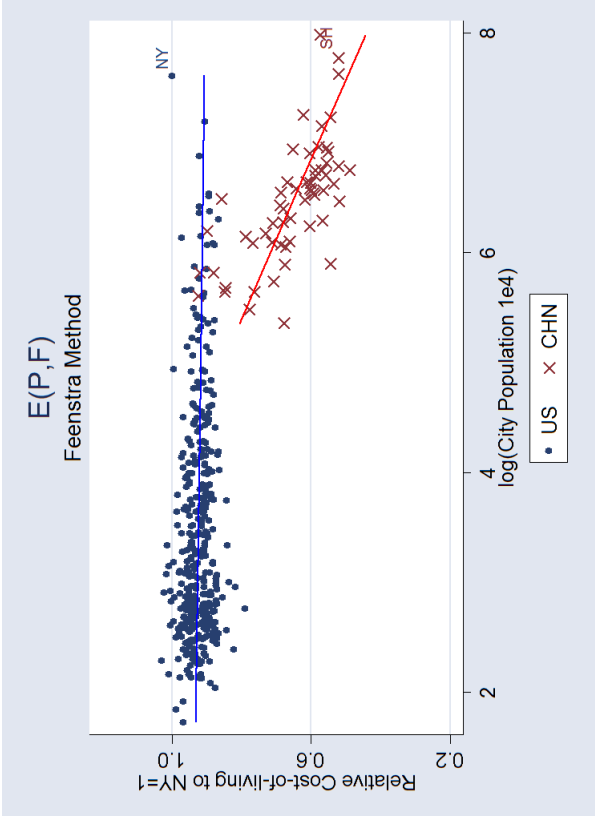[5]Campbell and Hopenhayn (2005) also find this result for thirteen retail trade industries across 225 U.S. cities.
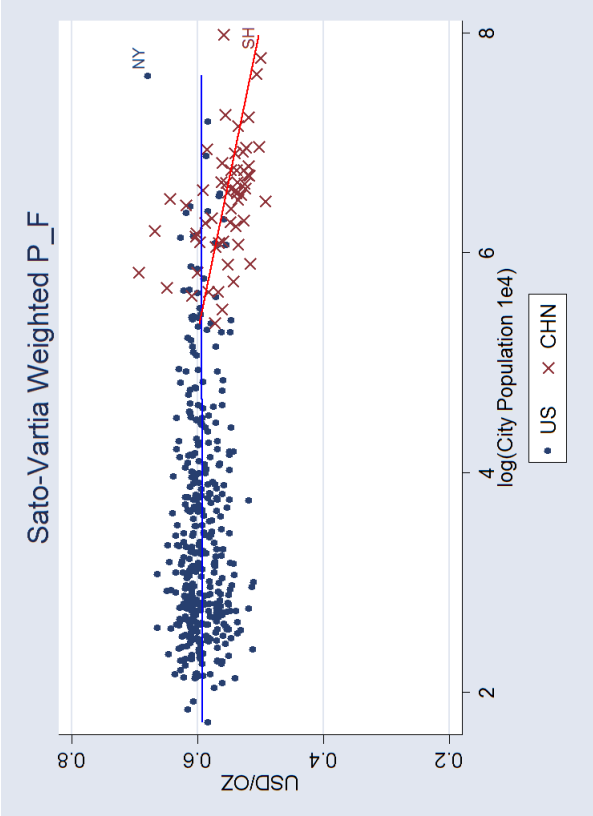
(a) Common Products Price Index
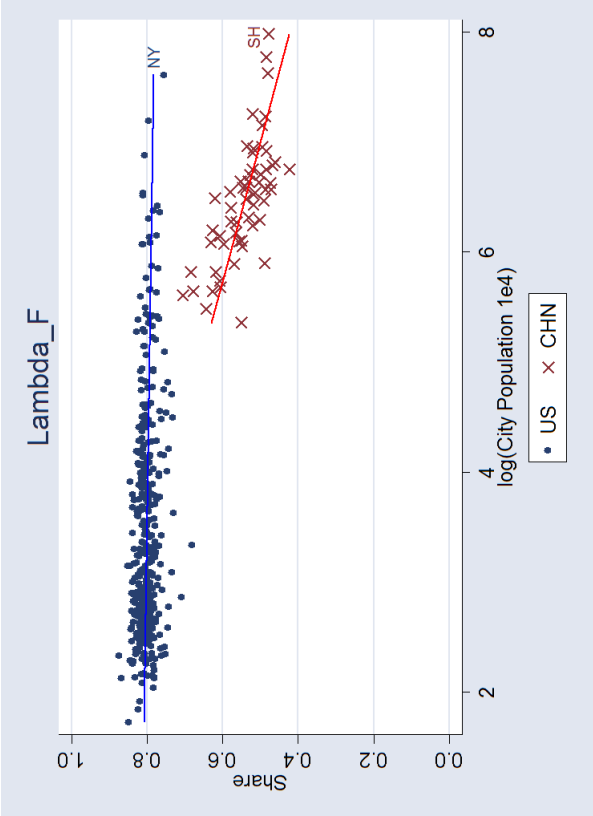
(b) Common Firm Sales Share

(c) Feenstra Cost-of-living Index

(d) Redding-Weinstein Cost-of-living Index

Data Source: City and MSA macro data are from China Statistical Yearbook and U.S. Bureau of Economic Analysis.

Figure 2: Cost-of-living Comparison: **Toothpaste**

(a) Common Products Price Index

(b) Common Firm Sales Share

(c) Feenstra Cost-of-living Index

(d) Redding-Weinstein Cost-of-living Index

Data Source: City and MSA macro data are from China Statistical Yearbook and U.S. Bureau of Economic Analysis.

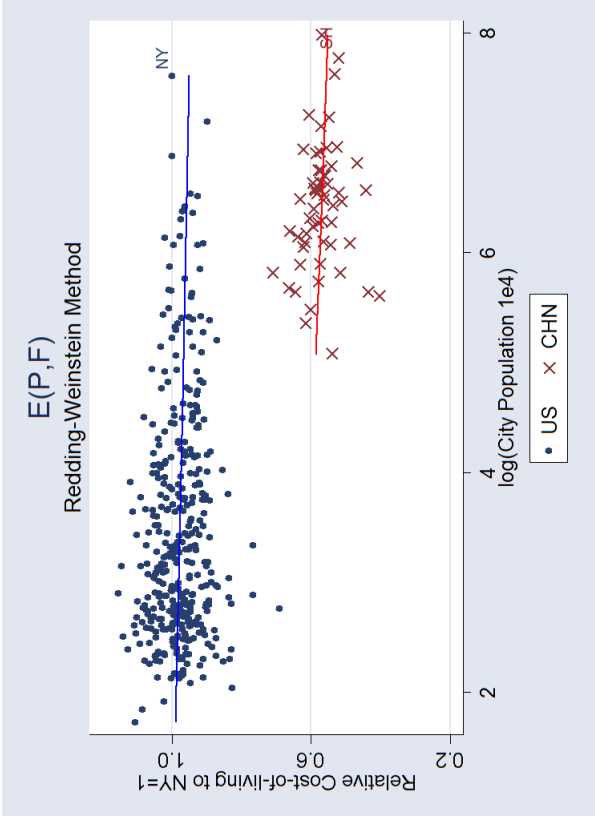Figure 3: Cost-of-living Comparison: **Laundry Detergent**

(a) Common Products Price Index

(b) Common Firm Sales Share

(c) Feenstra Cost-of-living Index

(d) Redding-Weinstein Cost-of-living Index

Data Source: City and MSA macro data are from China Statistical Yearbook and U.S. Bureau of Economic Analysis.

Figure 4: Cost-of-living Comparison: **Personal Wash Items**

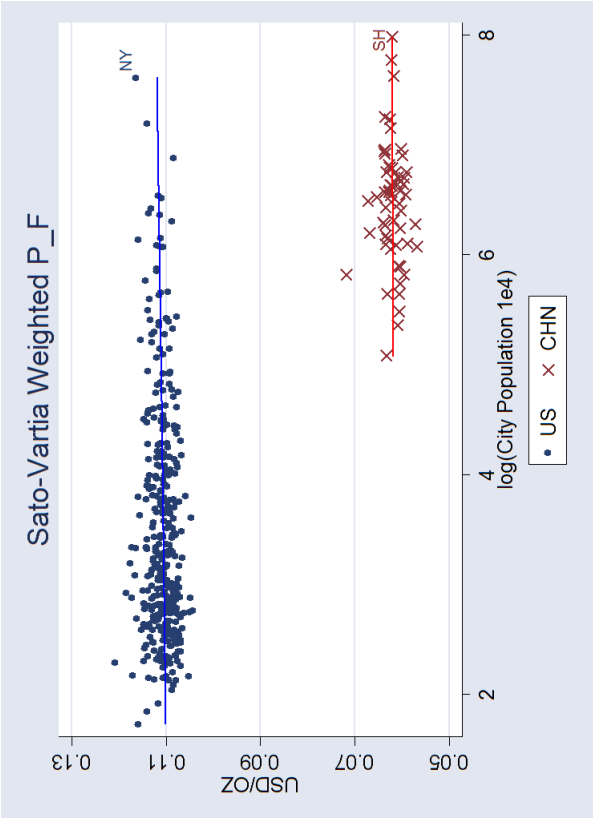(a) Common Products Price Index

(b) Common Firm Sales Share
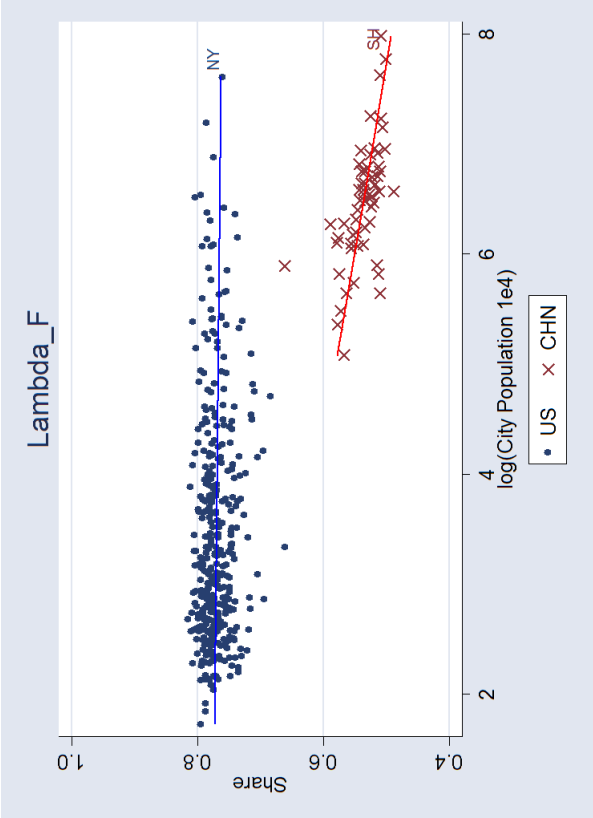
(c) Feenstra Cost-of-living Index

(d) Redding-Weinstein Cost-of-living Index

Data Source: City and MSA macro data are from China Statistical Yearbook and U.S. Bureau of Economic Analysis.
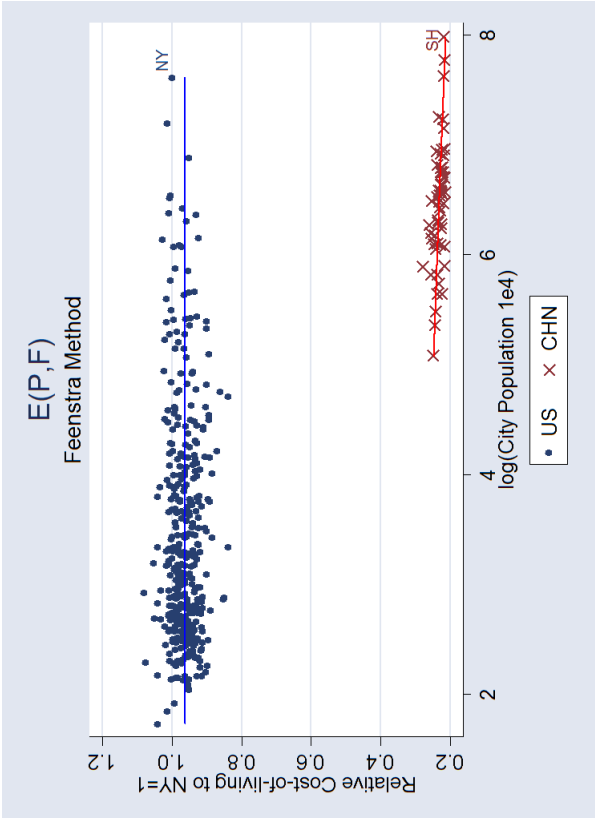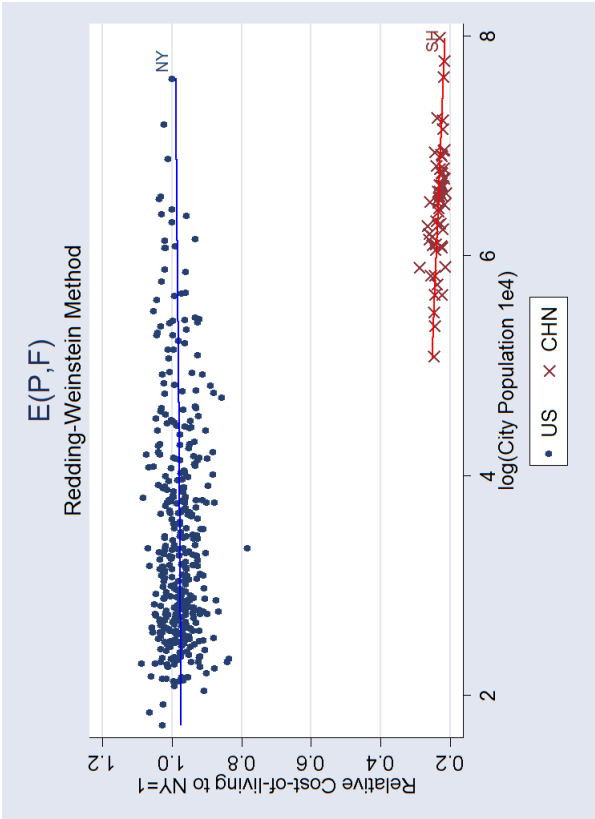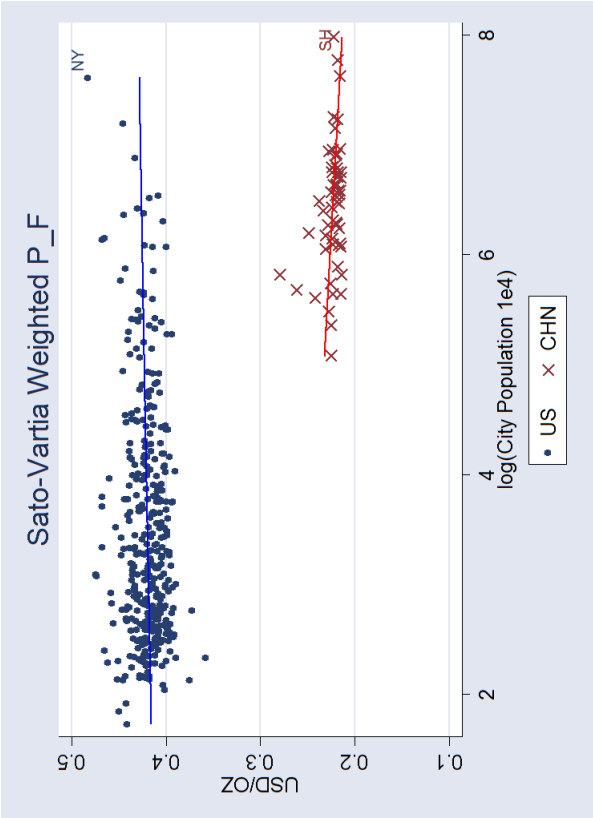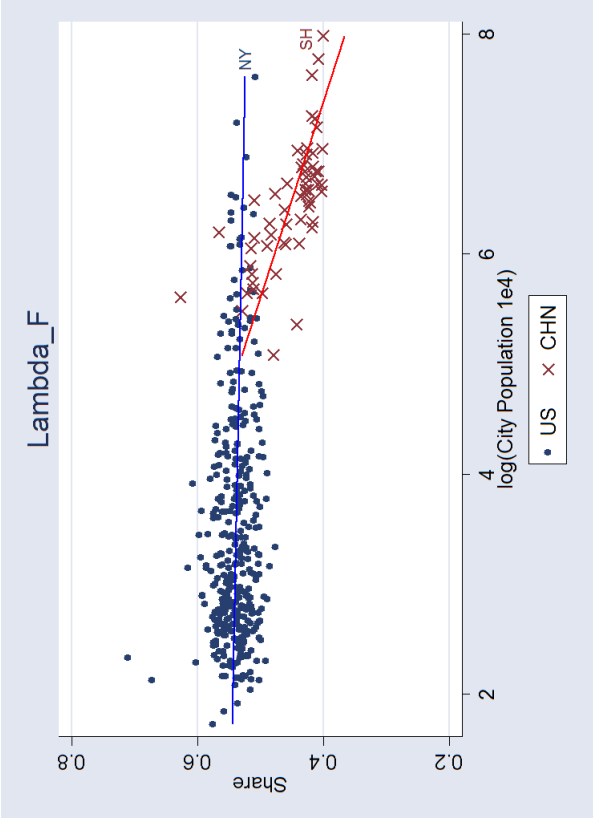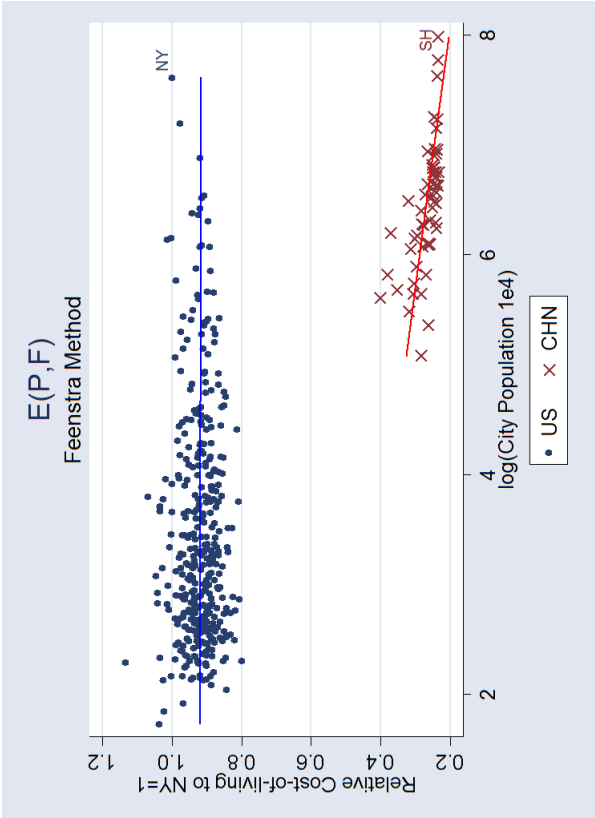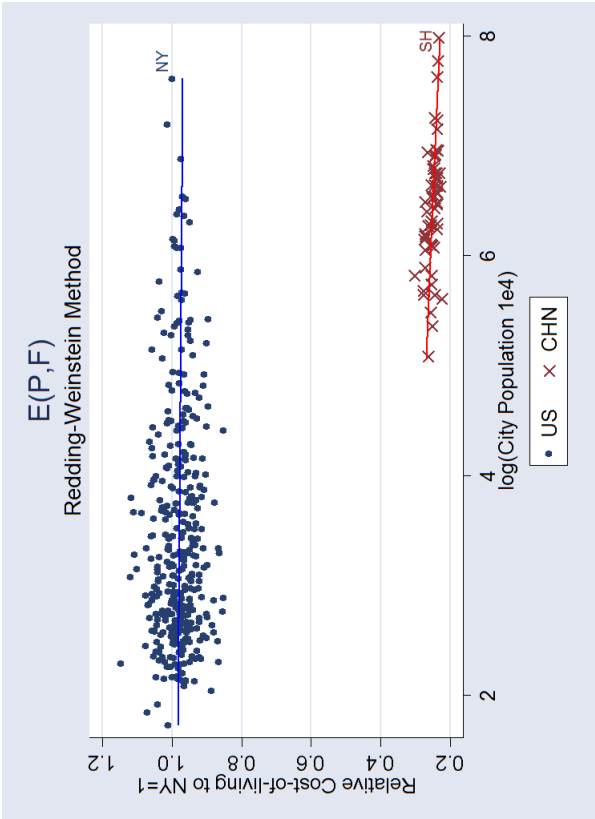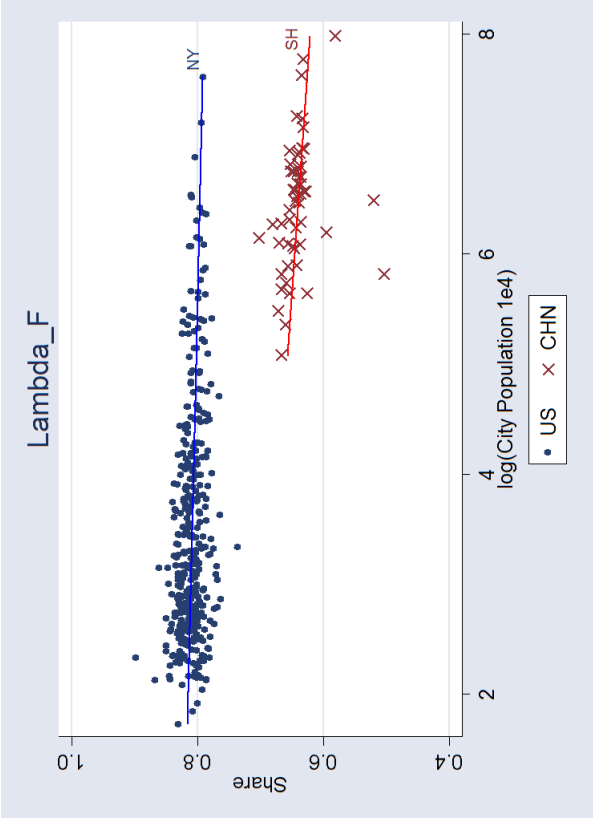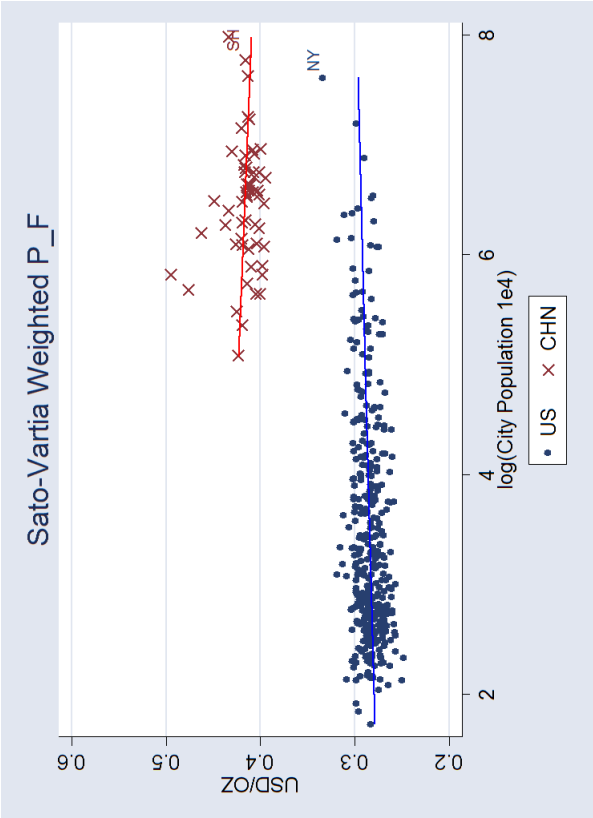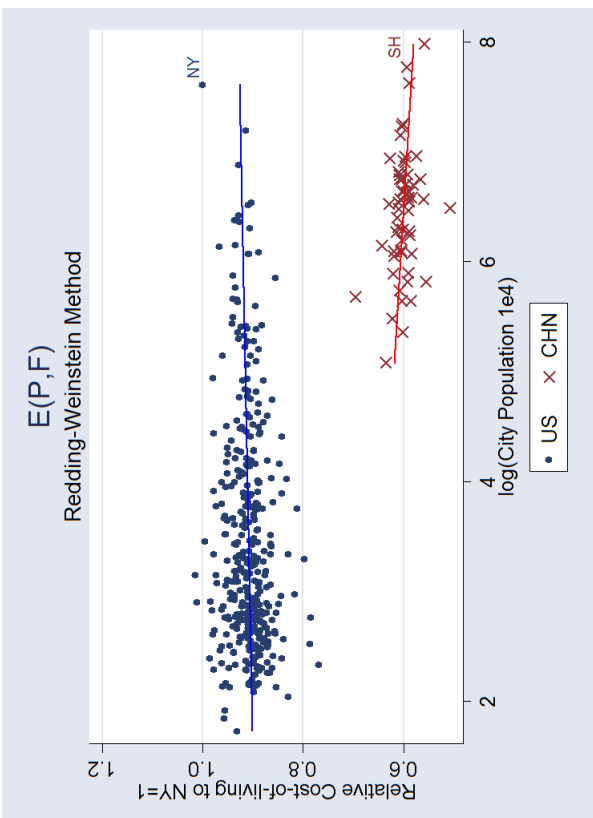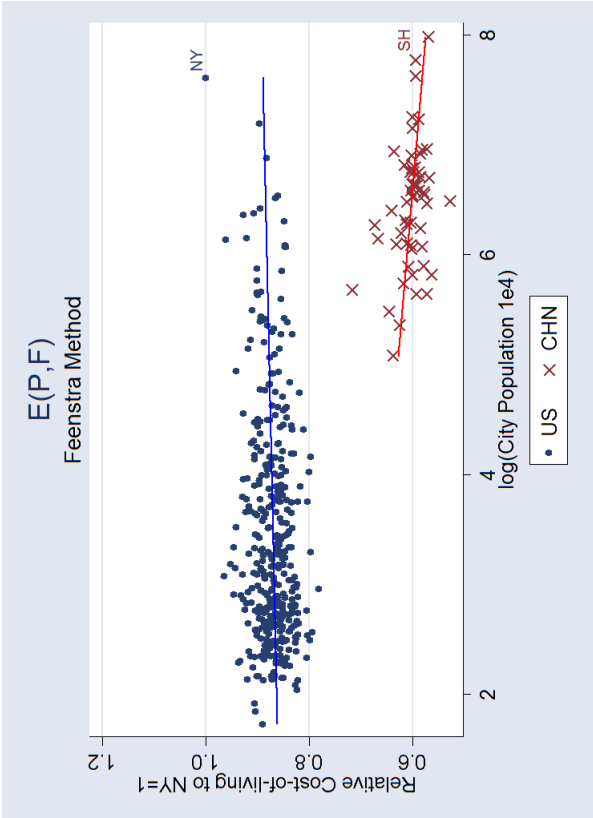
Figure 5: Cost-of-living Comparison: **Shampoo**

Because the terms in (10) use the common products within each country, these variables are only meaningful when comparing prices and variety for each country. In section 6 we develop a method for comparing these indexes across countries, drawing on Redding and Weinstein (2016). Continuing to normalize $P_{NY}^{SV} \Lambda_{NY}^{SV} = 1$ for New York, we apply this cross-country method to obtain the exact prices index for Shanghai, abbreviated as the point SH in panel (c). The the indexes $P_c^{SV} \Lambda_c^{SV}$ for all other Chinese cities are measured relative to Shanghai, and by construction, these can be compared to those in the United States. We see that the cost of living for each product in panel (c) is pulled down in China as compared to the Sato-Vartia geometric means in panel (a), at least in the larger Chinese cities. In other words, for these four products, the additional product variety due to both U.S. and local brands in the larger cities in China lowers the cost of living for consumers there as compared to the United States.

We further illustrate the relative cost-of-living in each China city relative to the average level of cost-of-living in the United States,[6] by plotting these in maps shown in Figure 6 to 9. We use the cold color (blue) to indicate a Chinese price level lower than that of the U.S. average, and the hot color (red) to represent a higher Chinese price level.[7] We see from Figure 6 to 9 that different product categories exhibit different relative price patterns. Toothpaste, Laundry Detergent and Personal Wash items are cheaper in China (the relative price varies from 34 to 81%, with U.S. average as unity), whereas prices of shampoo are 4% to 42% higher than than in the United States. Secondly, exact price levels are higher in inland cities than in the eastern coast cities that are usually of larger size. For example, the most expensive cities for Toothpaste, Laundry Detergent, Personal Wash items, and Shampoo are Urumqi (0.61), Lanzhou (0.71), Urumqi (0.82), and Hohhot (1.42) respectively. So the benefits of greater variety in China applies only to the larger cities, often found on the eastern coast and in certain inland locations such as Chongching, highlighted in the maps near the center of China and having relatively low prices.

---

[6]The U.S. average cost of living is calculated as the geometric mean of all MSAs

[7]The price follows an ascending order for colors ranking as dark blue, light blue, light red, and dark red.
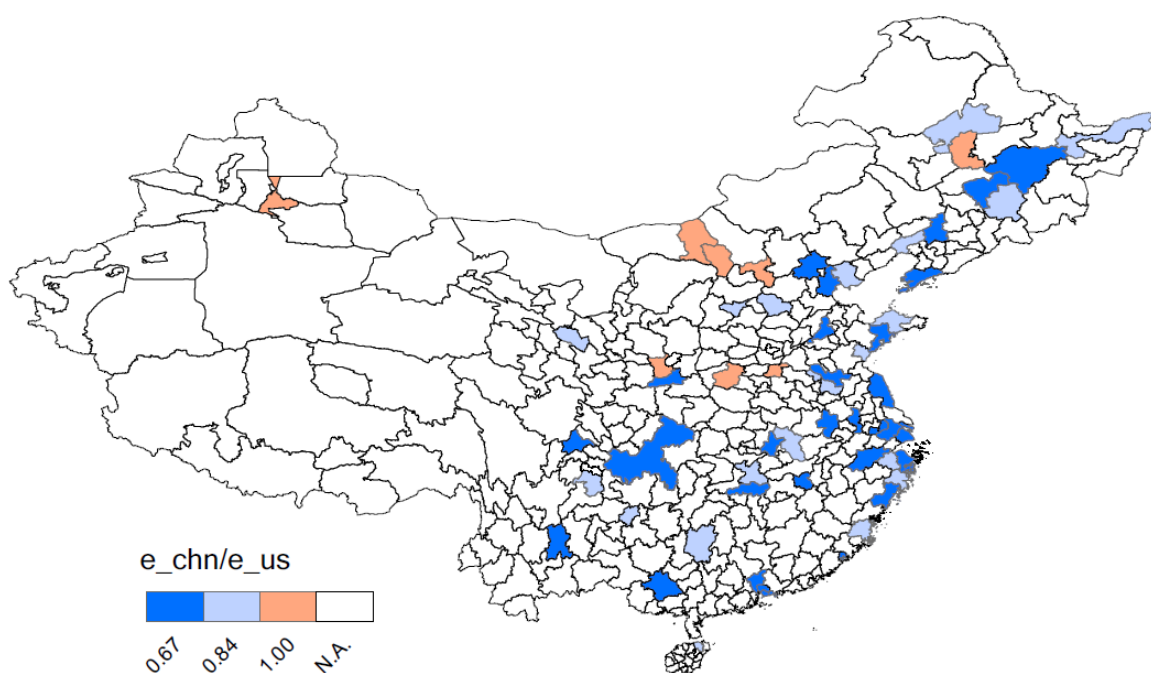
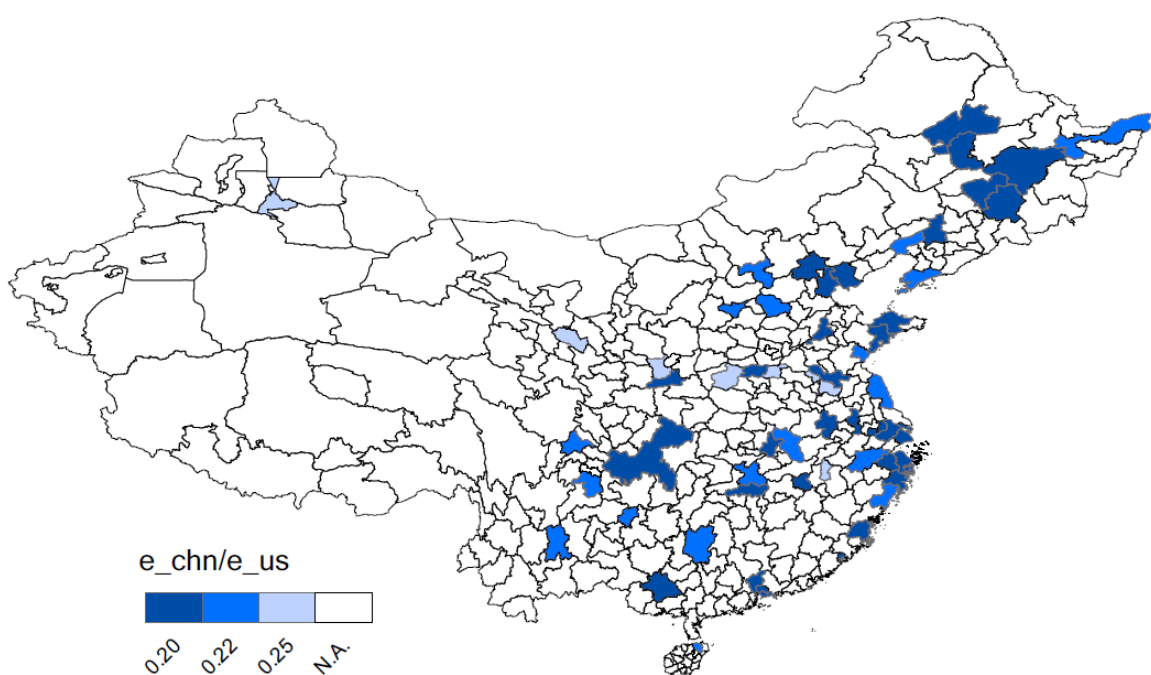Figure 6: Relative Cost-of-living in China: **Toothpaste**

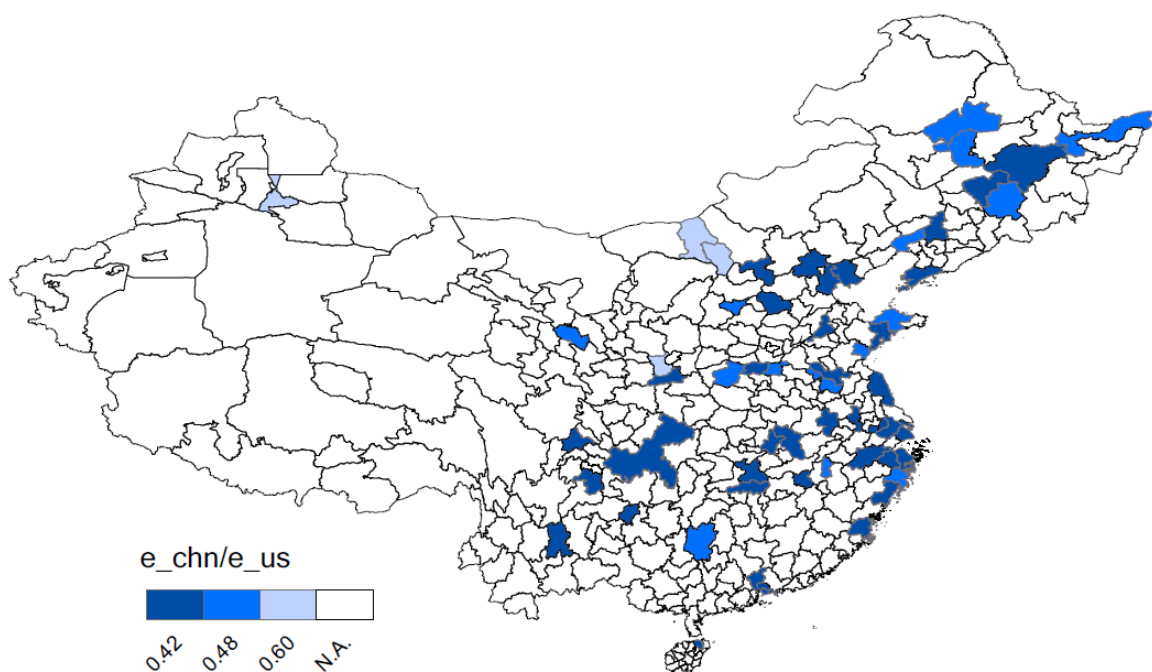Figure 7: Relative Cost-of-living in China: **Laundry Detergent**

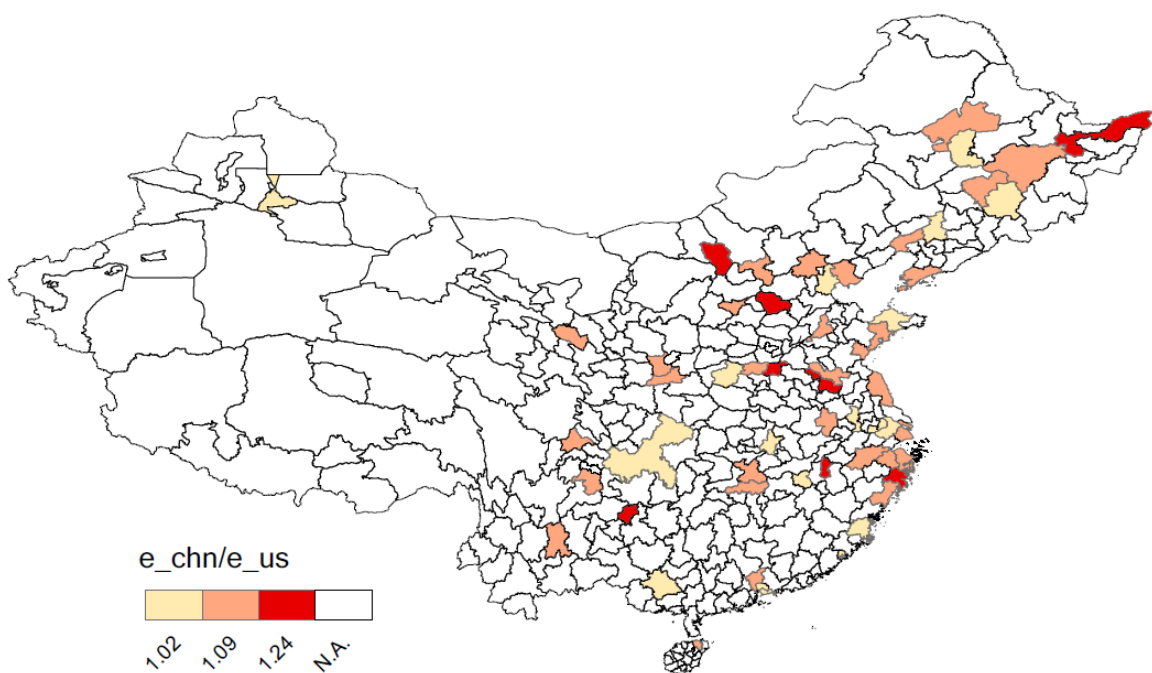Figure 8: Relative Cost-of-living in China: **Personal Wash Items**

Figure 9: Relative Cost-of-living in China: **Shampoo**

4 Firm Pricing and Choice of Variety

4.1 Nested CES Demand

To obtain the demand for each differentiated export consumed at home, let us start at the firm level. Demand for the aggregate of firm f products is $X_{fc} = (Y_c/P_c)(P_{fc}/P_c)^{-\eta}$, where Y_c denotes total expenditure, P_c is the overall CES price index, and P_{fc} is the CES price index (or unit-expenditure function) shown in (3). Then demand for each variety equals $b_{fic}x_{fic} = [(p_{fic}/b_{fic})/P_{fc}]^{-\sigma}X_{fc}$. Multiplying by (p_{fic}/b_{fic}) and using the equation for X_{fc}, we find that spending on each variety is,

$$p_{fic}x_{fic} = P_{fc}X_{fc}\left(\frac{p_{fic}/b_{fic}}{P_{fc}}\right)^{1-\sigma} = Y_c\left(\frac{p_{fic}/b_{fic}}{P_{fc}}\right)^{1-\sigma}\left(\frac{P_{fc}}{P_c}\right)^{1-\eta}, \tag{11}$$

using the above equation for X_{fc}.

We are now in a position to compute the elasticity of demand for a individual variety. We use the log of spending and differentiate to obtain:

$$\begin{aligned}
\varepsilon_{fic} &= -\frac{d\ln x_{fic}}{d\ln p_{fic}} = 1 - \frac{d\ln(p_{fic}x_{fic})}{d\ln p_{fic}} \\
&= 1 - (1-\sigma) - \left[(\sigma-\eta) + (\eta-1)\frac{d\ln P_c}{d\ln P_{fc}}\right]\frac{d\ln P_{fc}}{d\ln p_{fic}} \\
&= \sigma - [(\sigma-\eta) + (\eta-1)S_{fc}]s_{fic},
\end{aligned} \tag{12}$$

where $s_{fic} = d\ln P_{fc}/d\ln p_{fic}$ is the share of expenditure on product i *within* the sales of firm f, and $S_{fc} = d\ln P_c/d\ln P_{fc}$ is the total share of sales of firm f in city c. [8]

Our assumption that $\sigma \geq \eta$ implies that as the variety share s_{fic} rises then the elasticity of demand will fall; likewise, as the firm share S_{fc} rises the elasticity also falls. If the firm ignored the cannibalization effect of each variety on other sales then it would charge a higher price whenever s_{fic} or S_{fc} rises. But we now show that when the firm jointly profit-maximizes over all goods, taking into account cannibalization effects, then the price charged for each good will depend only the firm share S_{fc} and not on the within-firm variety share s_{fic}.

4.2 Optimal Prices for the Multiproduct Firm

Consider a firm producing variety i in city c and delivering it to destination d. The firm chooses the range of products to sell in multiple destinations $d = 1, ...D$. The profit-maximization problem for this firm is

$$\max_{p_{fid}, i \in I_{fd}} \sum_{d=1}^{D}\left\{\left[\sum_{i \in I_{fd}}(p_{fid} - g_{fi}(w_c) - T_{cd})x_{fid} - k_{fid}\right] - K_{fd}\right\}\mathbf{1}(I_{fd} \neq \varnothing), \tag{13}$$

where K_{fd} denotes the fixed costs to enter a city and $\mathbf{1}(I_{fd} \neq \varnothing)$ is an indicator variable that takes value of unity if $I_{fd} \neq \varnothing$ and zero otherwise. That is, the firm chooses whether to enter into each city d, and pays the fixed cost of K_{fd} only if it does along with the fixed costs of k_{fid} to sell each variety i in city d. We let $g_{fi}(w_c)$ denote the (constant) marginal costs of producing good i in city c, with

---

[8]Using our notation of section 2, the within-firm share of expenditure is $s_{fic} = s_{fic}(I_{fc})$, and the firm total share of expenditure is $S_{fc} = S_{fc}(F_c)$.

factor prices w_c, and selling it in city d with transport costs of T_{cd}. We assume that firms treat the prices of other firms as given under Bertrand competition, and that demand in the various cities is independent.

Focus initially on the choice of optimal prices. If the firm sold only a single product i in destination d, so that $s_{fid} = 1$, then the elasticity of demand from (12) is $\epsilon_{fid} = \eta - (\eta - 1)S_{fd}$ so that $\epsilon_{fid} - 1 = (\eta - 1)(1 - S_{fd})$. It follows from the usual markup formula that the optimal price is,

$$p_{fid} = \left[1 + \frac{1}{(\eta - 1)(1 - S_{fd})}\right][g_{fi}(w_c) + T_{cd}]. \tag{14}$$

When the firm sells multiple products, however, then it must take into account how a reduction in the price of one will decrease demand for its other products: this is the cannibalization effect. Nevetheless, we show in Appendix C that the same pricing formula as in (14) is obtained.[9] In other words, when the firm jointly maximizes over all its prices, the markup obtained is "as if" it was using the elasticity of demand in (12) but with $s_{fid} = 1$, so that the markup depends only on the total market share S_{fd} of the firm in that city.

4.3 Optimal Variety for the Multiproduct Firm

For simplicity, suppose that each good sold by firm f has the same demand and marginal cost regardless of the product and which city it is sold in, and that the firm sells N_{fd} of these varieties in city d.[10] We allow marginal and fixed costs to differ across firms, however. Then the profit maximization problem (13) is simplified as:

$$\max_{p_{fid}, N_{fd} \geq 0} \sum_{d=1}^{D} \left\{ N_{fd} \left[(p_{fid} - g_f - T_{cd})x_{fid} - k_{fd}\right] - K_{fd} \right\} \mathbf{1}(N_{fd} > 0). \tag{15}$$

The optimal price is still given by (14), though now with symmmetry this price is the same across varieties i for each firm. As the firm expands the number of varieties sold, it draws demand away from existing varieties. Taking this cannibalization effect into account, it is shown in Appendix C.2 that the optimal variety is determined by,

$$N_{fd} = \frac{\eta - 1}{\sigma - 1} \left[\frac{S_{fd}(1 - S_{fd})}{\eta - (\eta - 1)S_{fd}}\right] \frac{Y_d}{k_{fd}}, \tag{16}$$

where $Y_d = \rho w_d L_d$ is the total expenditure in destination d. Substituting this equation into (15) and also using the optimal price from (14), it can be shown that the profits from entering a city, *before* deducting the fixed costs K_{fd}, are:[11]

$$\pi_{fd} = \frac{(\sigma - \eta) + (\eta - 1)S_{fd}}{\sigma - 1} \left[\frac{S_{fd}}{\eta - (\eta - 1)S_{fd}}\right] Y_d. \tag{17}$$

---

[9]This result is also shown by Feenstra (2016), chapter 9, and Hottman, Redding, and Weinstein (2016).

[10]In Appendix C3, we generalize the analysis to allow the rising marginal cost of products that are farther from the core-competency of the firm. If we restrict the analysis to iceberg rather than specific trade costs, then we find that the equilibrium condition for the scope of a firm is essentially the same as that shown by (16), but with an extra constant term.

[11]To derive (17) we use that fact that the revenue earned per product by firm f is $S_{fd} Y_d / N_{fd}$.

4.4 Firm Entry

In order for the firm to serve city d, we must have $\pi_{fd} \geq K_{fd}$. To see how city size affects the entry of firms into cities, consider comparing a large city c with a smaller city d, with $Y_d < Y_c$. An equilibrium consists of a set of firms F_c and F_d selling in each city such that prices for each firm are given by (14), revenue per product is given by (11), and the number of products is as in (16), so that $S_{fc} = N_{fc} p_{fic} x_{fic} / Y_c$ with $\pi_{fc} \geq K_{fc}$ and $\sum_{f \in F_c} S_{fc} = 1$, and likewise in city d. There is ambiguity in the equilibrium set of firms that enter each city, because once a firm enters then demand for other firms is reduced. But regardless of this ambiguity, we assert that with $Y_d < Y_c$ then there will be a weakly smaller set of entrants $F_d \subseteq F_c$ in city d, and that the entrants are definitely reduced if city d is small enough.

To establish this claim, start with the equilibrium conditions for city c and then reduce expenditure Y_c. From (16) we see that N_c falls in direct proportion to Y_c provided that $S_{fc} = N_{fc} p_{fic} x_{fic} / Y_c$ is not affected. In fact, that is the equilibrium result because it can be confirmed from (11) that with all firms selling fewer products in direct proportion to the fall in expenditure Y_c, then the change in the CES price indexes P_{fc} and P_c ensure that revenue per product is not affected. Profits in (17) will fall in direct proportion to the fall in Y_c, however. It follows that for a sufficiently large reduction in expenditure, the firm with the smallest initial ratio $\pi_{fc}/K_{fc} \geq 1$ will be the first to have $\pi_{fc}/K_{fc} < 1$, and will therefore exit the market. It follows that $F_d \subset F_c$, as we asserted.

Denoting the exiting firm by g, the equilibrium condition $\sum_{f \in F_c} S_{fc} = 1$ will become $\sum_{f \in F_c \setminus g} S_{fc} = 1$, so the market shares of all remaining firms will rise as that firm drops out (because expenditure $p_{fic} x_{fic}$ on all remaining products increases as the price index P_c rises in (11)). By this argument, we see that smaller cities will have fewer firms in our model, with higher market shares. That will lead to higher prices in (14), since we have assumed that marginal costs for each firm are the same across cities. In other words, there is an anti-competitive effect in smaller cities, or a pro-competitive effect in larger cities. We now turn to the empirical analysis to find whether this prediction is borne out in the data for Chinese and U.S. cities.

5 Estimating the Firm Model

We will estimate (14) by moving the markup term to the left, so that the remaining marginal costs plus transport costs on the right will be estimated:

$$
p_{fid} \left[1 + \frac{1}{(\eta - 1)(1 - S_{fd})} \right]^{-1} = g_{fi}(w_c) + T_{cd}
$$
$$
= \alpha_i + \alpha_{cap} + \beta \ln Dist_{cd} + \epsilon_{fid}. \tag{18}
$$

The first term on the right is an indicator variable for variety i, which together with the eror term ϵ_{fid} reflects the marginal cost of production. We include an indictor variable α_{cap} for the capital city of each province in China, and the log of distance $lnDist_{cd}$ between the production in city c and the destination city d.

Let $N_{min} = min_{c,d}\{N_{cd}\}$ denote the "common" varieties sold in all cities in each country. Then the common-goods share of firm f in destination d is:

$$\lambda_{fd} = \frac{N_{min}p_{fd}x_{fd}}{N_{fd}p_{fd}x_{fd}} = \frac{N_{min}}{N_{fd}} = N_{min}\left(\frac{\sigma-1}{\eta-1}\right)\left[\frac{\eta-(\eta-1)S_{fd}}{S_{fd}(1-S_{fd})}\right]\frac{k_{fd}}{Y_d}, \tag{19}$$

using (16). The second regression equation comes from substituting $Y_d = \rho w_d L_d$ and taking logs,

$$\ln \lambda_{fd} = \gamma_{f,prov} + \delta_1 \ln\left[\frac{\eta-(\eta-1)S_{fd}}{S_{fd}(1-S_{fd})}\right] + \delta_2 \ln w_d + \delta_3 \ln L_d + \epsilon_{fd}. \tag{20}$$

The first term on the right is an indicator variable for the firm-province, which together with the eror term ϵ_{fd} reflects the fixed costs k_{fd} of providing each product in city d. The second term reflects an inverted-U-shaped relationship between the product scope of a firm and its market share S_{fd}, as shown by (16). Small firms have low shares and low scope, but large firms with high shares will hold back on product scope to avoid cannibalizing their own sales; product scope is maximized for the intermediate value of the firm share.[12] The common-goods share of firms in each city, λ_{fd}, is *inversely* related to their product scope, so the second term on the right of (20) is a U-shaped function of the firm's market share, and should have a coefficient of $\delta_1 = 1$ in theory.

Notice that the presence of w_d and L_d in (20) reflects how higher expenditure in larger cities *reduces* the common-goods share of firms in that city, λ_{fd}, thereby increasing product variety from each firm. In theory, the coefficients of these variables are both $\delta_2 = \delta_3 = -1$. A negative sign on these estimated coefficients will show how larger cities (measured by population or average income) have a lower common-goods share for firms, and therefore more product variety. We will likewise want to evaluate how prices charged by firms differ across cities of different sizes, and for this purpose we include include the variables w_d and L_d in (18), so that we actually estimate,

$$p_{fid}\left[1 + \frac{1}{(\eta-1)(1-S_{fd})}\right]^{-1} = \alpha_i + \alpha_{cap} + \beta_1 \ln Dist_{cd} + \beta_2 \ln w_d + \beta_3 \ln L_d + \epsilon_{fid}. \tag{21}$$

A final set of regressions are used is to decompose the cost of living indexes into the three terms on the right of (25), and study how each of these terms is related to the country size variables w_d and L_d. Specifically, we run the regressions:

$$\ln Z_d = \mu_1 + \mu_{cap} + \mu_2 \ln w_d + \mu_3 \ln L_d + \epsilon_d, \tag{22}$$

where Z_d denotes any of the three terms on the right of (25).

5.1 Estimating the Pro-Competitive Effect

Our first testable hypothesis is the existence of the pro-competitive effect, which is evaluated with the price equation, as shown in (21). For each of the four products, we use both the price (p) and the price divided by the markup ($p/markup$) as the dependent variables. We use the scraped prices at the barcode level and rely on Nielsen (China) data to calculate firm-destination shares and the markups. Information on city populations and average incomes, used to measure $Y_c = \rho w_c L_c$, are obtained for China and for MSAs in the United States, from standard sources. Finally, we used company reports

---

[12]In fact, product scope N_{fd} in (16) is maximized for $S_{fd} = \sqrt{\eta}/(1+\sqrt{\eta})$.

to identify the factory locations in China, and therefore compute the distance between the factory and destination markets.

Table 1 shows the result for toothpaste. We use price-per-item (U.S. cents/item) as the dependent variable for specification (1) and (2), and unit-price (U.S. cents/oz) for specification (3) and (4). The coefficients of ln *Population* and ln *Average Income* are all significantly negative, which implies larger and richer cities will benefit from lower prices for a given product at the barcode level. Notice that in specification (1), the coefficient on ln *Population* changes from -2.87 to -1.92 when we switch from using the price as the dependent variable to using the price divided by markup. This finding implies a higher markup in smaller cities, as the markup-adjusted price (price divided by markup) rises less in smaller cities than the per-item price.

We repeat the regressions but controlling for retailer (or supermarket) fixed effects in specification (2). The finding of higher markup in smaller and poorer cities remains robust. As expected, the firm will charge a higher price if the destination is farther from the location of factory. Interestingly, the regression result also indicates that access to supermarkets help reduce the price charged by firms all else equal, the coefficient changes from -2.87 to -4.13 after controlling supermarket fixed effects in specifications using the price as the dependent variable. The results in specification (2) continue to support the presence of a pro-competitive effect, even after controlling the supermarket.

Specifications (3) and (4) use the unit-price (U.S. cents/oz) as the dependent variable. We observe similar results as in specifications (1) and (2), and the pro-competitive effect remains significant: the coefficients of ln *Population* and ln *Average Income* are uniformly negative, and the coefficients are reduced when using unit-price/markup as a dependent variable rather than just unit-price.

In Tables 2, we explore the pro-competitive effect for the other product categories (Laundry Detergent, Personal Wash items and Shampoo). We focus on the price regressions with per-item price and price/markup as the dependent variables, but the results using unit-price are similar. From these tables, we still observe a lower price in richer and bigger cities, and firms charge a lower markup within richer and bigger cities.

Table 1: Price Regression of Toothpaste

Dep Var.	(1)		(2)		(3)		(4)	
	Price	P/mkup	Price	P/mkup	Unit-P	UP/mkup	Unit-P	UP/mkup
ln *Population*	-2.87***	-1.92***	-4.13***	-2.68***	-0.61***	-0.40***	-0.94***	-0.60***
	(0.25)	(0.18)	(0.24)	(0.15)	(0.07)	(0.05)	(0.06)	(0.04)
ln *Average Income*	-3.99***	-3.32***	-5.95***	-4.54***	-0.90***	-0.69***	-1.35***	-0.98***
	(0.32)	(0.24)	(0.31)	(0.20)	(0.08)	(0.06)	(0.08)	(0.06)
Capital City	3.59***	2.15***	1.16***	0.65***	0.85***	0.52***	0.27***	0.16***
	(0.26)	(0.17)	(0.22)	(0.14)	(0.07)	(0.04)	(0.06)	(0.04)
ln *Distance*	0.75***	0.49***	1.07***	0.67***	0.18***	0.12***	0.27***	0.17***
	(0.18)	(0.10)	(0.17)	(0.11)	(0.05)	(0.03)	(0.05)	(0.03)
Barcode FE	YES	YES	YES	YES	YES	YES	YES	YES
Retailer FE			YES	YES			YES	YES
Observations	89,399	89,399	89,399	89,399	89,399	89,399	89,399	89,399
Number of group	1,758	1,758	21,614	21,614	1,758	1,758	21,614	21,614
R-squared	0.006	0.009	0.015	0.019	0.005	0.006	0.010	0.012

Notes: *Average Income* is measured as GDP per capita in each city of the given *Population*. Prices (Price and P/mkup) are in units of U.S. cent and unit-prices (Unit-P and UP/mkup) are measured in U.S. cent per ounce. *Capital City* is dummy variable. *Distance* is in unit of *km*. Robust standard errors are clustered at group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Price Regression for the Other Product Categories (Price)

Dep Var.	(1) Laundry Detergent		(2) Personal Wash		(3) Shampoo	
	Price	P/mkup	Price	P/mkup	Price	P/mkup
ln *Population*	-3.69***	-2.72***	-1.47***	-1.37***	-4.88***	-3.44***
	(0.41)	(0.30)	(0.33)	(0.23)	(0.48)	(0.37)
ln *Average Income*	-4.24***	-3.12***	-4.86***	-3.48***	-6.50***	-5.68***
	(0.45)	(0.33)	(0.50)	(0.35)	(0.63)	(0.49)
Capital City	0.86**	0.65**	0.45	0.29	0.67	0.18
	(0.42)	(0.31)	(0.37)	(0.25)	(0.53)	(0.41)
ln *Distance*	1.14***	0.84***	1.66***	1.09***	3.65***	2.85***
	(0.30)	(0.22)	(0.33)	(0.23)	(0.39)	(0.30)
Observations	109,650	109,650	159,673	159,673	142,826	142,826
Number of group	28,001	28,001	35,965	35,965	31,711	31,711
R-squared	0.003	0.003	0.002	0.002	0.005	0.005

Notes: All regressions include Barcode-Retailer fixed effects. *Average Income* is measured as GDP per capita in each city of the given *Population*. Prices (Price and P/mkup) are in units of U.S. cent. *Capital City* is dummy variable. Robust standard errors are clustered at group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

5.2 Estimating the Variety Effect

Table 3 exhibits the result of our second structural equation as shown in 20. Motivated by the model, we add brand-province fixed effects to control for the fixed costs needed to sell each variety in city, k_{fd}. The dependent variable is $\ln \lambda_{fc}$, which denotes the expenditure in city c on the common products of firm f relative to firm f's total sales in that city. Columns (1) to (4) corresponds Toothpaste, Laundry Detergent, Personal Wash items, and Shampoo. The regressions results show that larger and richer cities are likely to have smaller expenditure share of common products for a given firm. This result implies that larger and richer cities get access to more varieties than the smaller ones. Consistent with model, the variable $\ln \frac{\eta-(\eta-1)S_{fd}}{S_{fd}(1-S_{fd})}$ contributes to the common product share positively, but with a coefficient less than its theoretical value of unity .

Table 3: Firm Share Regression

Dep Var: $\ln \lambda_{fd}$	(1) Toothpaste	(2) Laundry Detergent	(3) Personal Wash	(4) Shampoo
$\ln$ *Population*	-0.209***	-0.236***	-0.295***	-0.156***
	(0.037)	(0.030)	(0.034)	(0.021)
$\ln$ *Average Income*	-0.059**	-0.244***	-0.159***	-0.083***
	(0.029)	(0.030)	(0.030)	(0.012)
Capital City	-0.125***	-0.179***	-0.136***	-0.069***
	(0.023)	(0.029)	(0.028)	(0.015)
$\ln \frac{\eta-(\eta-1)S_{fd}}{S_{fd}(1-S_{fd})}$	0.147***	0.609***	0.396***	0.002
	(0.033)	(0.070)	(0.061)	(0.006)
Observations	660	420	600	840
Number of group	308	196	280	392
R-squared	0.376	0.662	0.586	0.348

Notes: All regressions include firm-province fixed effects. *Average Income* is measured as GDP per capita in each city of the given *Population*. *Capital City* is dummy variable. Robust standard errors are clustered at group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

6 Extension to other Products

6.1 Measuring Product Variety

We turn now to additional products for which we collected scraped data on their retail prices in 60 cities, but for which we do not have any scanner data purchased from Nielsen (China). This means that we do not have any information on their market shares, beyond know howing many supermarkets in each city sell each product (as is evident from the scraped data). In addition, as noted earlier, the United States and China use different barcode systems (UPC and EAN-13, respectively). For that reason, it is desirable to develop formulae for the exact price indexes that do not rely on having identical taste parameters b_{fic} for common goods that are available in every city and country. We

now show how the CES index derived by Redding and Weinstein (2016) allows for a much weaker assumption on taste parameters that allows the cost of living to be compared between countries.

To briefly review the results of Redding and Weinstein (2016), we start with the demand for each product variety, which equals $b_{fic}x_{fic} = [(p_{fic}/b_{fic})/P_{fc}]^{-\sigma}X_{fc}$. Multiplying by (p_{fic}/b_{fic}) and dividing by $P_{fc}X_{fc}$, we obtain an equation for the share of each variety within the total sales of firm f, which depends on the CES price index P_{fc}. Inverting that equation to solve for P_{fc}, we readily obtain:

$$P_{fc} = e(\mathbf{p}_{fc}, \mathbf{b}_{fc}, I_{fc}) = s_{fic}^{1/(\sigma-1)}\left(\frac{p_{fic}}{b_{fic}}\right). \tag{23}$$

The term λ_{fc} defined in (6) equals $s_{fic}/s_{fic}(I_f)$, so we can replace the share s_{fic} by $s_{fic} = s_{fic}(I_f)\lambda_{fc}$ in the above equation. Then because (23) holds *for every* product i, we take the unweighted geometric mean across all products to obtain the formula in Redding and Weinstein (2016):

$$P_{fc} = \left[\prod_{i\in I_f} s_{fic}(I_f)^{\frac{1}{N_f(\sigma-1)}}\right]\left[\prod_{i\in I_f}\left(\frac{p_{fic}}{b_{fic}}\right)^{\frac{1}{N_f}}\right]\lambda_{fc}^{1/(\sigma-1)}. \tag{24}$$

Significantly, this formula gives use the *level* of the CES price index P_{fc}, and not just its ratio as shown in section 2 by equation (7).

We can aggregate over firms in a city using a similar approach. The aggregate demand for each firm's products are $X_{fc} = (P_{fc}/P_c)^{-\eta}(Y_c/P_c)$. Multiplying by P_{fc} and dividing by Y_c, we obtain the share of each firm in city c, which depends on the CES price index P_c. Inverting that equation to solve for P_c, we readily obtain:

$$P_c = E(\mathbf{P}_c, F_c) = S_{fc}^{1/(\eta-1)}P_{fc}.$$

We again replace the share S_{fc} by $S_{fc} = S_{fc}(F)\lambda_c$, and take the unweighted geometric mean over the number of common firms M selling to *all* cities in each country. Then using that geometric mean along with (24), we obtain:

$$P_c = E(\mathbf{P}_c, F_c) = \left[\prod_{f\in F}\prod_{i\in I_f}\left(\frac{p_{fic}}{b_{fic}}\right)^{\frac{1}{MN_f}}\right]S_c(F)\Lambda_c(F), \tag{25}$$

where,

$$S_c(F) \equiv \left[\prod_{f\in F}S_{fc}(F)^{\frac{1}{M(\eta-1)}}\prod_{i\in I_f}s_{fic}(I_f)^{\frac{1}{MN_f(\sigma-1)}}\right], \quad \Lambda_c(F) \equiv \left[\prod_{f\in F}\lambda_{fc}^{1/M(\sigma-1)}\right]\lambda_c^{1/(\eta-1)}. \tag{26}$$

The first expression on the right of (25) is a simple geometric mean of the prices of common products, but these prices are adjusted for the taste parameters b_{fic}. As noted above, we have already assumed in our discussion of (7) that the taste parameters b_{fic} are identical for common goods that are available in every city within each country. We can now follow the assumption that Redding and Weinstein (2016) make in a time-series context, and use the *weaker* assumption that the geometric mean of taste parameters are equal in every city within each country:

$$\prod_{f\in F}\prod_{i\in I_f}\left(b_{fic}\right)^{\frac{1}{MN_f}} = 1, \quad \forall c \in= 1,...,D. \tag{27}$$

With this assumption, the first term on the right of (25) is easily measured as the geometric mean of prices. The second term is a geometric mean of the common products and common firms shares, as shown in (26). We interpret that term as possibly adjusting for differences in tastes for specific products *within* the common set, but subject to the maintained assumption that average tastes are the same as in in (26). The third term is the correction for variety outside the common set as in Feenstra (1994).

We will see that (25) and (26) can be readily used to measure the CES price index P_c across all cities *within* each country. But what about when we compare China and the United States? In that case, we need some analogue to the assumption in (27) that we are willing to apply across countries. To develop this idea, let us denote the *global brands* by the name of the common firms that sell in both countries by the set $G \equiv F^{US} \cap F^{China}$, which we assume is not empty.[13] We then pick the two large and comparable cities in each country, New York (NY) and Shanghai (SH). Our assumption is that there has been enough "convergence" of tastes and availability of global brands across these two cities that the average taste parameters are the same,

$$\prod_{g\in G}\prod_{i\in I_g}\left(b_{gi,NY}\right)^{\frac{1}{MN_g}} = \prod_{g\in G}\prod_{i\in I_g}\left(b_{gi,SH}\right)^{\frac{1}{MN_g}}. \tag{28}$$

With this assumption, we can compare the CES price indexes P_c between *countries*, as follows. We first use (25) to measure the CES price index P_c for New York in U.S. and Shanghai in China, with the common international firms G, which are denoted as $E(\mathbf{P}_{NY}, G)$ and $E(\mathbf{P}_{SH}, G)$. Next, we apply the same calculations to the cost-of-living measures across all cities in China, with the common firms F^{China}, and all cities in the United States, with the common firms F^{US}. We then *renormalize* the Chinese and U.S. price indexes so that the cost-of-living indexes are the same when using the common set as the global brands (G) and national brands (F^{China} and F^{US}) for Shanghai and New York. This is achieved by computing the terms (25) for the global brands G only in New York and in Shanghai, and then renormalizing the same terms for the Chinese and U.S. sets of common product F^{China} and F^{US} as:

$$\widetilde{E}(\mathbf{P}_c, F_c^{China}) \equiv E(\mathbf{P}_c, F_c^{China}) \times \frac{E(\mathbf{P}_{SH}, G)}{E(\mathbf{P}_{SH}, F_{SH}^{China})}$$

$$\widetilde{E}(\mathbf{P}_c, F_c^{US}) \equiv E(\mathbf{P}_c, F^{US}) \times \frac{E(\mathbf{P}_{NY}, G)}{E(\mathbf{P}_{NY}, F_{NY}^{US})} \tag{29}$$

---

[13]The global brands by product categories are listed in Table 9 in Appendix A3. In practice, G is empty for Chips and we alternatively define the global firms as $G \equiv [F^{US} \cap (\cup_c F_c^{China})] \cup [F^{China} \cap (\cup_c F_c^{US})]$, i.e. as the set of firms that sell national common products in one country and also sell to some cities in the other.

6.2 Cost of Living Indexes for Four Main Products

We calculate the cost-of-living indexes in New York and Shanghai, making use of assumption (28) so that these indexes are comparable between countries. The renormalization in (29) has already been applied to the indexes in panel (c) in Figure 2 to 5, which are otherwise constructed according to the method of Feenstra (1994). For comparison, we also calculate the indexes for these four products according to the method of Redding and Weinstein (2016), which uses the unweighted geometric mean of prices in (25), denoted by P_c^G, together with the multiplied terms $S_c(F) \times \Lambda_c(F)$. Those indexes are illustrated in panel (d) in in Figures 2 to 5, which are again renormalized according to 29.

A glance at those earlier Figures will show that the negative relationship between the exact price index and city size is noticably weaker, especially for Toothpaste, when using the Redding-Weinstein (RW) index than when using the Feenstra index. To formalize that difference, in Tables 4 and 5 we report the coefficients of regressions of the exact price indexes and its components on the log of city size (measured now with city GDP). Table 4 uses the Feenstra exact index and its components, while Table 5 uses the RW index and its compnents.

For Toothpaste, we see that the coefficients of the Feenstra exact index and its components using Sato-Vartia weights are more than twice as sensitive to city size than the RW exact index and its geometric price and variety components. The percentage contribution of these regression coefficients shows that the negative relationship between the exact price index and city size is caused roughly *equally* by the negative relationship between price and variety with city size (see the percentages reported in the Tables 4 and 5 for Toothpaste). For two other products – Personal Wash items and Shampoo – we also see that the negative relationship between the exact price index and city GDP is about twice as strong using the Feenstra index in Table 4 as with the RW index in Table 5. In these two products, it is the Sato-Vartia weighted prices themselves that decline more rapidly with city size in the Feenstra index, as compared to the simple geometric mean of prices in the RW index.

The final variable appearing in Table 5 is $\ln Share_c$, which is the log of the share variable $S_c(F)$ in (25). It is either insignficantly related to city size or has a very small coefficient. Thus, it is not the presence of that variable in the RW index that accounts for its reduced sensitivity to city size. But this variable *will* have a very important impact on the *average level* of the exact price index in China as compared to the U.S., especially as we turn to the other 15 products that we shall consider.

Table 4: Decomposition of Cost-of-living with City Size (Feenstra Method)

Dep.↓\Indep.→	Toothpaste		Laundry Detergent		Personal Wash		Shampoo	
	$\ln Y$	R^2	$\ln Y$	R^2	$\ln Y$	R^2	$\ln Y$	R^2
$\ln E(P_{fc}, F_c)$	-0.122***	0.424	-0.044***	0.419	-0.081***	0.320	-0.029***	0.233
	(0.022)		(0.007)		(0.018)		(0.007)	
$\ln P_c^{sv}$	-0.055***	0.363	-0.003	0.051	-0.021**	0.122	-0.015*	0.095
	(0.010)		(0.006)		(0.008)		(0.007)	
	45.1%		6.8%		25.6%		51.7%	
$\ln \Lambda_c$	-0.067**	0.361	-0.041***	0.524	-0.061***	0.395	-0.014	0.079
	(0.015)		(0.006)		(0.013)		(0.009)	
	51.9%		93.2%		74.4%		48.3%	

Notes: All regressions also include *Capital* dummy variable and constant, which are omitted in above table. For definition of each margin see theoretical part of the paper. Percentages describe the contribution of each margin to the overall cost-of-living elasticity. P_c^{sv} denotes the Sato-Vartia weighted mean of unit price. Y is measured as city's GDP. Standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Decomposition of Cost-of-living with City Size (Redding-Weinstein Method)

Dep.↓\Indep.→	Toothpaste		Laundry Detergent		Personal Wash		Shampoo	
	$\ln Y$	R^2	$\ln Y$	R^2	$\ln Y$	R^2	$\ln Y$	R^2
$\ln E(P_{fc}, F_c)$	-0.045***	0.166	-0.051***	0.453	-0.042***	0.357	-0.014	0.069
	(0.016)		(0.008)		(0.008)		(0.007)	
$\ln P_c^{g}$	-0.025***	0.185	-0.006	0.056	-0.004	0.058	-0.006	0.149
	(0.009)		(0.006)		(0.005)		(0.005)	
	55.6%		11.8%		9.5%		42.9%	
$\ln \Lambda_c$	-0.027**	0.261	-0.037***	0.478	-0.059***	0.363	-0.010	0.054
	(0.009)		(0.006)		(0.014)		(0.009)	
	60.0%		72.5%		140.5%		71.4%	
$\ln Share_c$	0.007	0.090	-0.008*	0.094	0.021	0.121	0.002	0.027
	(0.017)		(0.003)		(0.013)		(0.003)	
	-15.6%		15.7%		-50%		-14.3%	

Notes: All regressions also include *Capital* dummy variable and constant, which are omitted in above table. For definition of each margin see theoretical part of the paper. Percentages describe the contribution of each margin to the overall cost-of-living elasticity. P_c^{g} denotes the geometric mean of unit price. Y is measured as city's GDP. Standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

6.3 Cost of Living Indexes for 15 Other Products

When we have only the scraped price data and not the Nielsen (China) barcode expenditures, it is challenging to the construct the share term $S_c(F)$ and also the variety term $\Lambda_c(F)$ that appears in in (25). We create a proxy for the expenditure on each barcode product by simply counting the number of retail stores that sell each barcode in a city (and dividing by the total number of retailers). That is, the sales share of each product is approximated by:

$$\hat{Share}_{ic} = \frac{\sum_{s \in S_c} \mathbf{1}(i \text{ is observed in retailer } s)}{N_c^{retailers}}, \tag{30}$$

where $\mathbf{1}(i \text{ is observed in retailer } s)$ equals unity if we observe price quotes of barcode product i in retailer s and zero otherwise; $s \in S_c$ denote the collection of retailer stores in city c; and $N_c^{retailers}$ is the total number of retailer stores in city c.

In Table 6, we apply the RW index to our 4 main products and the other 15 product categories, while summarizing each component of the Chinese cost of living relative to the average value of its counterpart of the U.S. A value greater than one implies the corresponding variable is larger in China than in the United States. Beginning with the geometric mean of prices, as shown in the first column, these are lower in China for all products except for Shampoo, Baby Formula, and Milk Powder.

In the second column we show the ratios of terms $[S_{SH}(G)\Lambda_{SH}(G)/S_{NY}(G)\Lambda_{NY}(G)]$ for Shanghai relative to New York. This is the key ratio that adjusts the relative geometric mean of prices to obtain the relative cost of living in these two cities. This ratio is less than unity for our four main products, indicating that there is greater product variety in China than in the United States. We interpret that result as saying that the global brands G are sufficiently common in China for these four products that the presence of *additional* local brands in China gives it greater product variety.

For the 15 other products, this ratio is greater than unity for all products except for three: Coffee, Gum and Milk Powder. Coffee and Gum are similar to our four main products, where there are (inexpensive) local brands in China that are sufficient popular to result in greater product variety there. Along with Soft Drinks, which also has local brands in China, these three products in addition to the four main products have cost of living indexes in China relative to the U.S. that are *lower* than the relative geometric mean of prices shown in the first column.

Milk Powder is a unique case because there was a consumer scare about this product in China in 2008 and now the prices are higher than in the United States, as shown in the first column. Those high prices are slightly offset by a value of $[S_{SH}(G)\Lambda_{SH}(G)/S_{NY}(G)\Lambda_{NY}(G)]$ that is less than unity, but the exact price index shown in the final column is still much higher in China. That feature is shared by all the remaining ten products: the cost of living in China relative to the U.S., as shown in the final column, is higher than the relative geometric mean of prices. That higher cost of living in China for the remaining products is in most cases due to lower variety in China.[14] In some products (Biscuits, Chocolate, Soap, and Tea), that lower variety is strong enough that it *reverses* the lower relative price in China, as shown by the geometric mean of prices, to become a higher exact price index in China.

It is surprisingly that lower relative variety in China to a product like Tea, since China has spe-

---

[14]We further decompose $[S_{SH}(G)\Lambda_{SH}(G)/S_{NY}(G)\Lambda_{NY}(G)]$ into three margins as shown in Appendix D, Table 11.

cialized in the product from ancient times. The reason that we find lower relative variety in our data is that the many loose varieties of Chinese tea are not sold in supermarkets, which is where we obtain our scraped data. So we miss entirely the local varieties of Chinese tea and end up with only Western-style teas, with Lipton as the unique global brand that is sold in all cities in both countries. The finding of a higher relative price of tea in China illustrates a deficiency of our data, or more precisely, of the attempt to correct the prices of barcode products for variety. It remains to be seen whether other methods of obtaining prices and expenditure data across countries can overcome this limitation.

Table 6: Summary of Cost-of-Living by Product Categories (relative to the U.S.)

Product Category\Variable	P_c^G	$\frac{S_{SH}(G)\Lambda_{SH}(G)}{S_{NY}(G)\Lambda_{NY}(G)}$	Other Terms	$\tilde{E}(P_{fc}, F_c)$
Four Main Products				
Toothpaste	0.81	0.64	1.13	0.58
Laundry Detergent	0.59	0.52	0.76	0.23
Personal Wash Item	0.80	0.62	0.55	0.25
Shampoo	1.59	0.66	0.61	0.65
Other Products				
Baby Formula	1.30	1.12	1.02	1.48
Battery	0.24	1.01	1.27	0.31
Biscuit	0.64	2.18	3.05	4.23
Cereal	1.29	1.95	0.70	1.76
Chips	0.31	1.76	1.16	0.64
Chocolate	0.74	1.74	1.37	1.77
Coffee	0.70	0.97	1.00	0.68
Diaper	0.74	1.12	1.04	0.86
Dog Food	0.23	1.46	1.06	0.35
Gum	0.93	0.88	0.82	0.67
Milk Powder	5.29	0.74	1.60	6.20
Soap	0.37	4.28	0.89	1.43
Soft Drink	0.41	1.19	0.67	0.33
Tea	0.81	2.13	1.12	1.93
Toothbrush	0.53	1.57	1.13	0.93

Notes: The table summarizes population weighted average of each varaible in relative terms (greater than unity implies the corresponding variable is larger in China.). P_c^G is in unit of cents per ounce for the four main products and dollars per item for other products; $\frac{S_{SH}(G)\Lambda_{SH}(G)}{S_{NY}(G)\Lambda_{NY}(G)}$ denotes the composite share of varieties in Shanghai relative to New York; and Other Terms is the further adjustment needed to obtain $\tilde{E}(P_{fc}, F_c)$, which is the cost-of-living index in China relative to the U.S.

For each of the 15 product categories, we also explore to what extent the three components (denoted as P_c^G, Λ_c and $Share_c$) contribute to the overall cost of living. Table 7 presents the regression results of the three components on total city size (i.e. city GDP).[15] . The first column is the elasticity of cost-of-living with respect to the total sizet size. As expected, larger cities generally benefit from having a lower cost-of-living index. Cities with twice market size have a lower indexes with declines ranging from 1% (Chocolate) to 10.4% (Cereal). Pro-competitive effect are active in most cases, with a

---

[15]In each regression, we also include a dummy variable for the capital city which is not reported in the Table.

few exceptions of Battery, Biscuit, Chocolate, Soap and Toothbrush, as shown in the second column. Lastly, access to varieties become an important sources of potential gain of living in big cities, and it explains a sizable proportion of the decline in the cost of living on average.

Table 7: Decomposition of Cost-of-living with City Size (Other Product Categories)

Product	$\ln E(P_{fc}, F_c)$	$\ln P_c^G$	$\ln(\lambda_c \times Share_c)$
Baby Formula	-0.014	-0.008	-0.006
		59.0%	41.0 %
Battery	-0.017	0.005	-0.022
		-27.2%	127.2%
Biscuit	-0.028	0.001	-0.028
		-2.2 %	102.2%
Cereal	-0.104	-0.020	-0.084
		19.2%	80.8%
Chips	-0.019	-0.011	-0.008
		57.2 %	42.8%
Chocolate	-0.010	0.003	-0.014
		-32.9%	132.9 %
Coffee	-0.054	-0.012	-0.042
		22.6 %	77.4 %
Diaper	-0.020	-0.015	-0.005
		76.2%	23.8%
Dog Food	-0.069	-0.014	-0.055
		20.1%	79.9 %
Gum	-0.029	-0.012	-0.017
		41.0%	59.0%
Milk Powder	-0.063	-0.007	-0.057
		10.7 %	89.3%
Soap	-0.033	0.008	-0.041
		-22.8%	122.8%
Soft Drink	-0.072	-0.022	-0.050
		30.5 %	69.5%
Tea	-0.057	-0.006	-0.051
		10.3%	89.7%
Toothbrush	-0.080	0.005	-0.085
		-6.3 %	106.3%

Notes: The reported coefficients denote the elasticity with respect to the total market size. All regressions also include *Capital* dummy variable and constant, which are omitted in above table. For definitions of each margin see theoretical part of the paper. Percentages describe the contribution of each margin to the overall cost-of-living elasticity; P_c^G denotes the geometric mean of unit price; and Y is city GDP.

7   Conclusions

[to be added]

Appendix

A. Collection Method for Retail Price Data of China

The retail price data of China is collected by scraping data from a mobile phone application. The mobile phone app that we rely on is *Wochacha* (the English meaning is "I search"), a leading consumer-product information platform in China.

A1. Introduction to *Wochacha APP*

The mobile phone application *Wochacha* was developed by Wochacha Info Tech Co. Ltd in January 2010. Detailed information regarding *Wochacha* can be obtained from its web-page.[16] It is a widely used price comparison app, using a QR code scanner so that people could see where a particular product can be bought at the lowest price. Consumers could also search for price quotes based on the barcode of the product.[17] By February 2014, the number of users had exceeded 210 million and distributed across all the provinces in China.

According to its Public Service Rules, *Wochacha* uses two methods to collect the retail price data. First, the firm hires and trains data collectors who visit supermarkets to collect prices. The second source of retail price data relies on partnerships with supermarkets, from which *Wochacha* directly obtains the scanner data. In either method, the collected prices are not published until *Wochacha* data auditors have examined them.

A2. Data Collection Example

Figure 10 exhibits the results when one search for prices quotes of Pampers[18]in Shanghai. Based on the results, there are eight retail price quotes varying form $112RMB$ by Walmart to $165RMB$ by Lotus. Besides the price-by-supermarket information, we also observe the producer name, country of origin, as well as the unit count and producer ID. We developed software to mimic the mobile phone app on the computer to automatically export price data for the given set of EAN code in the selected city. Then we repeated this process across 60 sampled cities to construct our retailing price database. The collection took place from October to December in 2015.

A3. Cities for Analysis

Cities for which we scraped price data are listed in Table 8. We include 60 cities from 28 provinces, which are widely spread across China.

---

[16]The official web-page: http://www.wochacha.com/about/index.
[17]Countries like U.S. and Canada use Universal Product Code (UPC) as the unique identity for products, whereas China adopts coding system of European Article Number which technically refers to EAN-13.
[18]The EAN-13 is 6903148040737.

Table 8: Cities with Scraped Data

Province	City	Province	City	Province	City
Anhui	**Hefei**	Guizhou	Guiyang	Shanxi	Taiyuan
	Suzhou	Hainan	Haikou		Datong
Beijing	**Beijing**	Hebei	Shijiazhuang	Shan'xi	**Xi'an**
Fujian	**Xiamen**		Tangshan		Xianyang
	Fuzhou	Henan	Kaifeng	Shanghai	**Shanghai**
Gansu	Lanzhou		Luoyang	Sichuan	**Chengdu**
Guangdong	**Guangzhou**		Zhengzhou		Yibin
	Shenzhen	Heilongjiang	Qiqihar	Tianjin	**Tianjin**
	Dongguan		Daqing	Xinjiang	Urumqi
Guangxi	Nanning		Jiamusi	Yunnan	**Kunming**
	Guilin		**Harbin**	Chongqing	**Chongqing**
Hubei	**Wuhan**	Jiangxi	Nanchang	Zhejiang	**Hangzhou**
	Huanggang		Jingdezhen		**Ningbo**
Hunan	**Changsha**	Liaoning	**Shenyang**		Shaoxing
	Yueyang		Jinzhou		Taizhou
Jilin	Changchun		**Dalian**		Wenzhou
	Jilin	Inner Mongolia	Baotou		
Jiangsu	Nanjing		Hohhot		
	Suzhou	Shandong	**Jinan**		
	Wuxi		Qingdao		
	Xuzhou		Rizhao		
	Yancheng		Yantai		

Notes: Cities also included in Nielsen (China) sales database are marked in bold.

Figure 10: Interface of *Wochacha* Application

A3. Global Brands

The global brands for each product are listed in Table 9. For all products except for Chips, the global brands are computed as the intersection of the common brands in the U.S. and in China, $G \equiv F^{US} \cap F^{China}$. For Chips this set is empty, so we adopt the broader definition of global brands described in note 13.

Table 9: List of Global Brands by Product

Product Category	Global Brands	Product Category	Global Brands
Toothpaste	Colgate	**Chocolate**	M&M
	Crest	**Coffee**	Maxwell
	Sensodyne	**Diaper**	Huggies
Shampoo	Dove		Pampers
	L'Oreal	**Dog Food**	Pedigree
	Pantene	**Gum**	Wrigley
Personal Wash Items	Olay	**Milk Powder**	Nestle
Laundry Detergent	Tide	**Soap**	Johnson
Baby Formula	Enfamil		Olay
	Nestle		Safeguard
	Similac	**Soft Drink**	Minute Maid
Battery	Duracell		Red Bull
	GP	**Tea**	Lipton
Biscuit	Meiji	**Toothbrush**	Colgate
Cereal	Quaker		Crest
Chips	Doritos		
	Kellogg		
	Lay's		
	Oishi		

31

B. Other Data

B1. Elasticities of Substitution

We directly use the elasticities of substitution for our four main products (Toothpaste, Laundry Detergent, Personal Wash items and Shampoo) as well as for other product categories from Hottman, Redding, and Weinstein (2016), which are listed in Table 10.

Table 10: Elasticity of Substitution

Product Category	Firm Elasticity (η)	Product Elasticity (σ)
Toothpaste	2.75	5.23
Laundry Detergent	3.85	6.60
Personal Wash	3.42	5.11
Shampoo	4.43	6.54
Baby Formula	7.47	14.62
Battery	8.72	15.10
Biscuit	5.51	7.31
Cereal	2.25	4.66
Chips	4.37	7.32
Chocolate	5.13	7.44
Coffee	3.89	7.31
Diaper	7.44	24.30
Dog Food	5.51	8.15
Gum	3.14	4.86
Milk Powder	2.41	5.05
Sanitizer	2.96	5.15
Soap	3.42	5.11
Soft Drink	5.18	7.82
Tea	4.46	6.12
Toothbrush	2.75	5.23

B2. Prediction Method for Four Main Products

The cost-of-living calculation requires firm sales information, whereas the Nielsen (China) sales database only provide us with this information for only 22 large cities. To allow our analysis to also cover smaller cities, we estimate the the market share by using the product, retailer and city information for the four main products, as discussed in this part.

For the four main products, we have the detailed EAN-level market sales in 22 cities, in addition to the price quotes collected by scraping the mobile phone application. We implement a Heckman two-step approach to estimate the product share within the 22 cities using the Nielsen (China) sales database, as shown below:

$$Prob(Share_{ic} > 0) = \alpha_0 + \alpha_1 \ln Y_c + \alpha_2 \ln L_c + \alpha_3 \ln N_c + \alpha_4 Capital_c \qquad (31)$$
$$+ \delta_i Variety_i + \gamma_{1r} Region_r + \epsilon_{ic},$$

$$E(\ln Share_{ic}|Share_{ic} > 0) = \beta_0 + \beta_1 Capital_c + \phi_{ic} Brand_i \times \ln Y_c$$

$$+ \mu_{ic} Brand_i \times \ln L_c + \psi_{ic} Brand_i \times \ln N_c + \gamma_{2r} Region_r, \quad (32)$$

where (31) is the selection equation and 32) is the observation equation. Y_c denotes city c's GDP; L_c is city population; N_c denotes the total number of differentiated products (at EAN level) in city c; $Capital_c$ is a dummy variable which equals unity if city c is a capital city; $Region_r$ is a set of region dummies (East, Middle and West China). Finally, for each product i at EAN-13 level, $Brand_i$ denotes a collection of Brand dummies, e.g., Crest, Colgate etc. for Toothpaste.; $Variety_i$ is a collection of variety categories which is constructed by jointly combining several product characteristics[19]; $Share_{ic}$ is calculated as product i's sales share in city c. The two-step Heckman approach is motivated by the fact that we only observe price quotes in the mobile phone application conditional on that product is sold in a particular city. To reduce the endogeneity issues in price regression, we do not include the product price on the right of the share specification (32).

After obtaining the coefficients $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}, \hat{\phi}, \hat{\mu}, \hat{\psi})$, we predict the market share in the cities that do not have sales information, and we further rescale them so that the sum of all product shares equals unity within each city. For consistency, we use the predicted shares to calculate the cost-of-living index for all cities, including the 22 for which we have the expenditure data from Nielsen (China).

Though this prediction method works well in most cities, it results in extreme predictions in some cities, particularly when some product accounts for 80% or more of the total market. To avoid this situation, we exclude cities that have the extremely low city's average firm sales share of common products $(\prod_{f \in F}(\lambda_{fc})^{W_f(F)})$ from our analysis. Specifically, we firstly regress city's average firm sales share of common products, $\lambda_c^f \equiv \prod_{f \in F}(\lambda_{fc})^{W_f(F)}$, on city's population ($\ln L_c$), and derived the prediction error $\hat{u}_c = \lambda_c^f - \hat{\lambda}_c^f$. Then we drop the cities from our sample if its prediction error is twice larger than the standard deviation of the prediction errors, i.e., we exclude city c if $|\hat{u}_c| > 2 \times std(\hat{u})$[20]. After implementing the criteria, we finally have (out of 60 cities) 58 cities for Toothpaste, Laundry Detergent and Personal Wash Items, and 57 cities for Shampoo.

C. Solution to the Firm's Problem

C1. Choice of price by the firm

The solution to the problem (13) w.r.t. the price p_{fid} is

$$x_{fid} + \sum_{j \in I_{fd}} [p_{fid} - g_{fi}(w_c) - T_{cd}] \frac{dx_{fjd}}{dp_{fjd}} = 0, \quad \forall i \in I_{fd} \quad (33)$$

This expression is more complicated than for a single-product firm because the firm is selling all the products $i \in I_{fd}$, and therefore must take into account the effect of a change in each price p_{fic} on all these products. To simplify this expression, it can be confirmed that the CES demand derivatives are symmetric, $dx_{fjd}/dp_{fid} = dx_{fid}/dp_{fjd}$. Using (33) and dividing by demand x_{fid}, we can re-express the

---

[19]For example, for Toothpaste, we define a variety by brand, subbrand, special function, flavors, and form. Variety *Crest-Complete Plus-Deep Clean-Mint-Paste* is different from *Crest-Complete Plus-Mouth Wash-Mint-Paste*

[20]We also tried with one and three standard deviations, and the main results barely changed.

condition as

$$1 + \sum_{j \in I_{fd}} \left[1 - \frac{g_{fj}(w_c) + T_{cd}}{p_{fjd}} \right] \frac{d \ln x_{fid}}{d \ln p_{fjd}} = 0, \quad \forall i \in I_{fd} \tag{34}$$

Let us denote the ratio of price to marginal cost, inclusive of transport costs, by $\mu_{fjd} = p_{fjd}/(g_{fjd} + T_d) \geq 1$. We can see that the expression in brackets in (34) equals $(\mu_{fjd} - 1)/\mu_{fjd} \geq 0$, which is the difference between price and marginal cost measured relative to price. This is the 'Lerner index' of monopoly power for a single-product firm, and with price chosen optimally will equal the inverse of its elasticity of demand. To see how this Lerner pricing rule is modified with multiproduct firms, let us conjecture a solution where the price-cost ratios are constant across all products sold by the firm in question, $\mu_{fjd} = \mu_{fd}$. Then it is immediately clear that the solution to (34) is

$$\left(\frac{\mu_{fd} - 1}{\mu_{fd}} \right) = - \left(\frac{d \ln x_{fid}}{d \ln p_{fjd}} \right)^{-1} \tag{35}$$

Expression (35) says that the Lerner index for the firm equals the inverse of the sum of demand elasticities (rather than just the inverse of the own demand elasticity, as occurs for a single-product firm). In order for this solution to be valid, however, we need to have that the sum of elasticities on the right of (35) be independent of good i, because we have assumed that the markup is common across goods. This independence holds if an equi-proportional increase in all prices charged by a firm needs to lead to the same percentage drop in demand for any product sold by that firm. It turns out that this condition is satisfied for CES demands, in which case the sum of elasticities is[21]

$$- \sum_{j \in I_{fd}} \frac{d \ln x_{fid}}{d \ln p_{fjd}} = \eta - (\eta - 1)S_{fd} > 1 \tag{36}$$

Notice that the expression on the right-hand side of (36) is precisely what we get from (12) if we replace the share s_{fid} within the firm by unity, since the firm sells all of its own products. Then the pricing formula is derived as (14).

C2. Choice of variety by the firm

The first-order condition of (15) with respect to N_{fd} yields[22]

$$[p_{fd} - g_f(w_c) - T_{cd}]x_{fd} + N_{fd}[p_{fd} - g_{fi}(w_c) - T_{cd}]\frac{dx_{fd}}{dN_{fd}} = k_{fd} \tag{37}$$

Using demand function, the elasticity of demand w.r.t. product scope N_{fd} is given in (38):

$$\begin{aligned} \frac{d \ln x_{fd}}{d \ln N_{fd}} &= (\sigma - \eta)\frac{d \ln P_{fd}}{d \ln N_{fd}} + (\eta - 1)\frac{d \ln P_d}{d \ln P_{fd}}\frac{d \ln P_{fd}}{d \ln N_{fd}} \\ &= \frac{\sigma - \eta + (\eta - 1)S_{fd}}{1 - \sigma} \end{aligned} \tag{38}$$

The second inequality comes from $\frac{d \ln P_{fd}}{d \ln N_{fd}} = \frac{1}{1-\sigma}$ when marginal cost are the same. Condition (37) can be rewritten as(16),

$$N_{fd} = \frac{\eta - 1}{\sigma - 1} \left[\frac{S_{fd}(1 - S_{fd})}{\eta - (\eta - 1)S_{fd}} \right] \frac{Y_d}{k_{fd}}$$

---

[21]The cross-elasticity is $d \ln x_{fid}/d \ln p_{fjd} = -[(\sigma - \eta) + (\eta - 1)S_{fc}]s_{fjc}$, for $i \neq j$ and along with(12), (36) is obtained.

[22]As products have the same marginal cost, it is convenient to omit subscript i in the analysis.

C3. Model of core-competency and iceberg trade cost detailed

In this section, we resolve firm's problem with heterogeneous product marginal cost and iceberg trade cost. The production technology of firm is characterized by a core competence and flexible manufacturing. Firms are able to adjust the number of varieties with adaption cost in marginal cost. Specifically, the marginal cost of producing variety ω is assumed to be monotonically increasing function. There is also a bilateral iceberg transportation cost τ_{cd} when firm sells to other cities. With these assumptions, the profit maximization of firm f in region c is expressed as,

$$\max_{N_{fd}, \{p_{fd}(\omega)|\omega\in[0,N_{fd}]\}} \sum_{d=1}^{D} \left\{ \int_{0}^{N_{fd}} [p_{fd}(\omega) - \tau_{cd}g_f(\omega)]x_{fd}(\omega)d\omega - k_{fd}N_{fd} - K_{fd} \right\} \mathbf{1}(N_{fd} > 0). \tag{39}$$

Following the steps shown in Appendix C.1, the optimal price is:

$$p_{fd}(\omega) = \left[1 + \frac{1}{(\eta-1)(1-S_{fd})}\right] \tau_{cd}g_f(\omega), \quad \forall \omega \in [0, N_{fd}]. \tag{40}$$

The first order condition with respect to N_{fd} yields,

$$[p_{fd}(N_{fd}) - \tau_{cd}g_f(N_{fd})]x_{fd}(N_{fd}) + \int_{0}^{N_{fd}} [p_{fd}(\omega) - \tau_{cd}g_f(\omega)]\frac{d \ln x_{fd}(\omega)}{dN_{fd}}d\omega = k_{fd}. \tag{41}$$

Using demand function, the elasticity of demand with respect to product scope is given by:

$$\frac{d \ln x_{fd}}{d \ln N_{fd}} = (\sigma - \eta)\frac{d \ln P_{fd}}{d \ln N_{fd}} + (\eta - 1)\frac{d \ln P_d}{d \ln P_{fd}}\frac{d \ln P_{fd}}{d \ln N_{fd}}$$

$$= \frac{\sigma - \eta + (\eta - 1)S_{fd}}{1 - \sigma}s_{fd}(N_{fd})N_{fd}. \tag{42}$$

In this expression, $s_{fd}(N_{fd})$ denotes the within firm share of the very last product, indexed by $\omega = N_{fd}$, which is defined as,

$$s_{fd}(N_{fd}) = \frac{p_{fd}(N_{fd})x_{fd}(N_{fd})}{\int_{0}^{N_{fd}} p_{fd}(\omega)x_{fd}(\omega)d\omega} = \frac{g_f(N_{fd})^{1-\sigma}}{\int_{0}^{N_{fd}} g_f(\omega)^{1-\sigma}d\omega} \equiv H(N_{fd}),$$

where $H(.)$ increases with its argument given $g_f' > 0$. Jointly using (41) and (42), the optimal scope is derived as,

$$H(N_{fd}) = \frac{\eta - 1}{\sigma - 1}\left[\frac{S_{fd}(1 - S_{fd})}{\eta - (\eta - 1)S_{fd}}\right]\frac{Y_d}{k_{fd}}. \tag{43}$$

If we further assume $g_f(\omega) = w_c\omega^\theta$, with $\theta \geq 0$ and $\sigma < 1 + (1/\theta)$, then $H(N_{fd}) = (1 + \theta - \sigma\theta)/N_{fd}$. It follows that the equation for λ_{fd} is derived as,

$$\lambda_{fd} = \frac{N_{fd}^{(\sigma-1)\theta}N_{min}^{1+\theta-\theta\sigma}}{1+\theta-\theta\sigma}\left(\frac{\sigma-1}{\eta-1}\right)\left[\frac{\eta - (\eta-1)S_{fd}}{S_{fd}(1-S_{fd})}\right]\frac{k_{fd}}{Y_d}.$$

D. Summary of Decomposition of $S(G) \times \Lambda(G)$ for New York and Shanghai

Table 11: Summary of Decomposition of $S(G)\Lambda(G)$ for New York and Shanghai

Product Category\Variable	$\frac{\prod_{f \in G} S_{f,SH}(G)}{\prod_{f \in G} S_{f,NY}(G)}$	$\frac{\prod_{f \in G} \prod_{i \in I_f} s_{fi,SH}(I_f)}{\prod_{f \in G} \prod_{i \in I_f} s_{fi,NY}(I_f)}$	$\frac{S_{SH}(G)}{S_{NY}(G)}$	$\frac{\Lambda_{SH}(G)}{\Lambda_{NY}(G)}$	$\frac{S_{SH}(G)\Lambda_{SH}(G)}{S_{NY}(G)\Lambda_{NY}(G)}$
Four Main Products					
Toothpaste	1.01	1.53	1.55	0.41	0.64
Laundry Detergent	1.00	1.13	1.13	0.46	0.52
Personal Wash Item	1.00	0.60	0.60	1.03	0.62
Shampoo	1.06	1.05	1.11	0.60	0.66
Other Products					
Baby Formula	1.25	1.01	1.27	0.88	1.12
Battery	1.00	1.35	1.35	0.75	1.01
Biscuit	1.00	0.98	0.98	2.23	2.18
Cereal	1.00	2.15	2.15	0.91	1.95
Chips*	1.75	0.85	1.49	1.18	1.76
Chocolate	1.00	1.62	1.62	1.07	1.74
Coffee	1.00	1.20	1.20	0.80	0.97
Diaper	1.01	1.12	1.13	0.99	1.12
Dog Food	1.00	1.03	1.03	1.42	1.46
Gum	1.00	1.26	1.26	0.70	0.88
Milk Powder	1.00	0.60	0.60	1.24	0.74
Soap	1.46	1.18	1.72	2.49	4.28
Soft Drink	1.05	1.56	1.63	0.73	1.19
Tea	1.00	1.52	1.52	1.41	2.13
Toothbrush	1.99	1.33	2.65	0.60	1.57

Notes: The table summarizes population weighted average of each varaible in relative terms (greater than unity implies the corresponding variable is larger in China.). The definition of each variables can refer to the notation used in the main text. Specifically, $S_c(G) \equiv \left[\prod_{f \in G} S_{f,c}(G)^{1/M(\eta-1)} \right] \times \left[\prod_{f \in G} \prod_{i \in I_f} s_{fi,c}(I_f)^{1/MN_f(\sigma-1)} \right]$ and we omit the exponents in the column headings for simplicity. Population data is from China Statistical Yearbook and U.S. Bureau of Economic Analysis.

References

CAMPBELL, J. R., AND H. A. HOPENHAYN (2005): "Market Size Matters," *Journal of Industrial Economics*, 53(1), 1–25.

CAVALLO, A., AND R. RIGOBON (2016): "The Billion Prices Project: Using Online Prices for Measurement and Research," NBER Working Paper 22111.

DEATON, A., AND B. ATEN (2014): "Trying to Understand the PPPs in ICP2011: Why are the Results so Different?," NBER Working Paper 20244.

DEATON, A., AND A. HESTON (2010): "Understanding PPPs and PPP-based National Accounts," *American Economic Journal: Macroeconomics*, 2(4), 1–35.

FEENSTRA, R. C. (1994): "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1), 157–177.

——— (2016): *Advanced International Trade,*Princeton University Press, 2nd edition.

FEENSTRA, R. C., H. MA, J. PETER NEARY, AND D. PRASADA RAO (2013): "Who Shrunk China? Puzzles in the Measurement of Real GDP," *The Economic Journal*, 123(573), 1100–1129.

HANDBURY, J., AND D. E. WEINSTEIN (2015): "Goods Prices and Availability in Cities," *The Review of Economics and Statistics*, 82(1), 258–296.

HOTTMAN, C., S. J. REDDING, AND D. E. WEINSTEIN (2016): "Quantifying the Sources of Firm Heterogeneity," *Quarterly Journal of Economics*, forthcoming.

INKLAAR, R., AND D. S. P. RAO (2015): "Cross-Country Income Levels Over Time: Did the Developing World Suddenly Become much Richer?," *American Economic Journal: Macroeconomics, forthcoming*.

REDDING, S. J., AND D. E. WEINSTEIN (2016): "A Unified Approach to Estimating Demand and Welfare," Discussion paper, Princeton University and Columbia University.

SATO, K. (1976): "The Ideal log-Change Index Number," *Review of Economics and Statistics*, 58, 223–228.

VARTIA, Y. O. (1976): "Ideal log-Change Index Numbers," *Scandinavian Journal of Statistics*, 3, 121–126.