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

John Romalis
University of Sydney

May 18, 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 also compare the cost-of-living indexes for particular products between China and the United States. In most cases, the observed prices differences between the countries (lower prices in China) are partially offset by the variety differences (reduced variety in China), so that the cost of living in China is not as low as the price differences suggest, especially in smaller cities.

*We thank Steve Redding and David Weinstein for providing the elasticities of substitution used here. Financial support from the National Science Foundation is gratefully acknowledged, Please direct correspondence to: Robert Feenstra, University of California, Davis, email: rcfeenstra@ucdavis.edu.

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 reduction in measured 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 expected, then the quantity of goods consumed by the representative Chinese individual were lower (since quantity is obtained by dividing expenditure by price). 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 prices from the Chinese prices from 2005 round of the ICP tend to be 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 following in the 2011 round the ICP, and relative prices in Asia and in other developing countries went lower.

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 cannot be used 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, but is prohibitively

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

[2] Addition 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!

expensive to obtain for China for more than a handful of products. Accordingly, in this paper we rely on Chinese scanner data for four products in 22 cities during, purchased from Nielsen (China), and have supplemented these data with scraped prices for the same four products, obtained from a mobile phone application that allows consumers to check for the various prices at supermarkets in each city. In this way the sample of 22 cities is extended to 60 cities of varying size, and in later work we plan to greatly extend the sample of products, too.[3] 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 most cases, the observed prices differences between the countries (lower prices in China) are partially 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.

In section 2 we describe the nested CES framework that we shall use to measure the consumer gains from variety. A model of multi-product firms in presented in section 3, from which we derive the solutions for firm pricing and product scope. These equations are estimated as regressions described in section 4, where we describe the data sources in more detail. Our empirical results are presented in section 5.

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

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

[3] We have collected prices for 5,000 barcode items using the mobile phone app.

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

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

As in Hottman, Redding, and Weinstein (2014), 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 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 varities 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:

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

and likewise for the shares $s_{fid}(I_f)$, also defined over the common set of products in city d. 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 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,

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

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

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

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 (7).

3 Firm Pricing and Choice of Variety

3.1 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,

$$p_{fic} x_{fic} = P_{fc} X_{fc} \left(\frac{p_{fic}/b_{fic}}{P_{fc}} \right)^{1-\sigma} = Y_c \left(\frac{p_{fic}/b_{fic}}{P_{fc}} \right)^{1-\sigma} \left(\frac{P_{fc}}{P_c} \right)^{1-\eta}. \tag{10}$$

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:

$$\begin{aligned}\epsilon_{fic} &= -\frac{d\ln x_{fic}}{d\ln p_{pic}} = 1 - \frac{d\ln(p_{fic}x_{fic})}{d\ln p_{fic}} \\ &= 1-(1-\sigma) - \left[(\sigma-\eta)+(\eta-1)\frac{d\ln P_c}{d\ln P_{fc}}\right]\frac{d\ln P_{fc}}{d\ln p_{fic}} \\ &= \sigma - [(\sigma-\eta)+(\eta-1)S_{fc}]s_{fic},\end{aligned} \tag{11}$$

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. [4]

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 the within-firm variety share s_{fic}.

3.2 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

$$\max_{p_{fid}, i\in I_{fd}} \sum_{d=1}^{D} \left\{ \left[\sum_{i\in I_{fd}} (p_{fid} - g_{fi}(w_c) - T_{cd})x_{fid} - k_{fid} \right] - K_{fd} \right\} \mathbf{1}(I_{fd} \neq \emptyset), \tag{12}$$

where K_{fd} denotes the fixed costs to enter a city and $\mathbf{1}(I_{fd} \neq \emptyset)$ is an indicator variable that takes value of unity if $I_{fd} \neq \emptyset$ and zero otherwise. We let $g_{fi}(w_c)$ denote the (constant) marginal costs of producing good i in city c, with factor prices w_c, and selling it in city d with transport costs of T_{cd}. The term k_{fid} are the fixed costs needed to sell each variety in city d. 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 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,

$$p_{fid} = \left[1 + \frac{1}{(\eta-1)(1-S_{fd})}\right][g_{fi}(w_c) + T_{cd}]. \tag{13}$$

[4] 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 $S_{fc} = S_{fc}(F_c)$.

When the firm sells multiple products, 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. As shown in Appendix A.1, the same pricing formula as in (13) is obtained. 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 (11) but with $s_{fid} = 1$, so that the markup depends only on the total market share S_{fd} of the firm in that city.[5]

3.3 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.[6] We allow marginal and fixed costs to differ across firms, however. Then the profit maximization problem (12) is simplified as:

$$\max_{p_{fid}, N_{fd} \geq 0} \sum_{d=1}^{D} \left\{ N_{fd} \left[(p_{fid} - g_f - T_{cd}) x_{fid} - k_{fd} \right] - K_{fd} \right\} \mathbf{1}(N_{fd} > 0). \tag{14}$$

The optimal price is still given by (13), though this price is the same across varieties i. 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 A.2 that the optimal variety is determined by,

$$N_{fd} = \frac{\eta - 1}{\sigma - 1} \left[\frac{S_{fd}(1 - S_{fd})}{\eta - (\eta - 1) S_{fd}} \right] \frac{Y_d}{k_{fd}}, \tag{15}$$

where $Y_d = \rho w_d L_d$ is the total expenditure in destination d. Substituting this equation into (14) and also using the optimal price from (13), it can be shown that the profits from entering a city, *before* deducting the fixed costs K_{fd}, are:[7]

$$\pi_{fd} = \frac{(\sigma - \eta) + (\eta - 1) S_{fd}}{\sigma - 1} \left[\frac{S_{fd}}{\eta - (\eta - 1) S_{fd}} \right] Y_d. \tag{16}$$

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 with a smaller city, 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 (13), revenue per product is given by (10), and the number of products is as in (15), 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.

[5]This result is also shown by Hottman, Redding, and Weinstein (2014).

[6]In Appendix A3, 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 (15), but with an extra constant term.

[7]To derive (16) we use that fact that the revenue earned per product by firm f is $S_{fd} Y_d / N_{fd}$.

To establish this claim, start with the equilibrium conditions for city c and then reduce expenditure Y_c. From (15) we see that N_c falls in direct proportion, from which it follows that $S_{fc} = N_{fc} p_{fic} x_{fic} / Y_c$ is not affected.[8] Profits in (16) 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 (10)). 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 (13), 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 predition is borne out in the data for Chinese and U.S. cities.

4 Estimating Equations

4.1 Data

Our calculation of the cost of living rely 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.[9] The data used for analysis includes four product categories, namely Toothpaste, Laundry Detergent, Personal Wash items, and Shampoo, covering 22 cities in China, as shown in the red colored regions of Figure 1.[10] Besides the sales information for each product, we also observe manufacturer information such as brand and sub-brand of each product. In our empirical analysis, we refer to brand as the firm (e.g. Crest or Colgate for toothpaste).

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 B, 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 formula derived from theory to calculate the cost of living in China for each of the four product categories. Our third source of data is the Nielsen HomeScan (U.S.) database, used to calculate the cost of living for 377

[8] It can be confirmed from (10) 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.

[9] Note that the UPC systems used in China and in the United States differ, though it is not difficult to identify similar product categories.

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

MSAs of the United States in 2011, for each of the four product categories.[11]

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 these elasticities from Hottman, Redding, and Weinstein (2014), as shown in Appendix C. These authors estimate the elasticities from the Nielsen HomeScan (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. 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 to identify the factory locations in China, and therefore compute the distance between the factory and destination markets. But we have only completed this collection of factory distance for toothpaste, which we shall use as our baseline case.

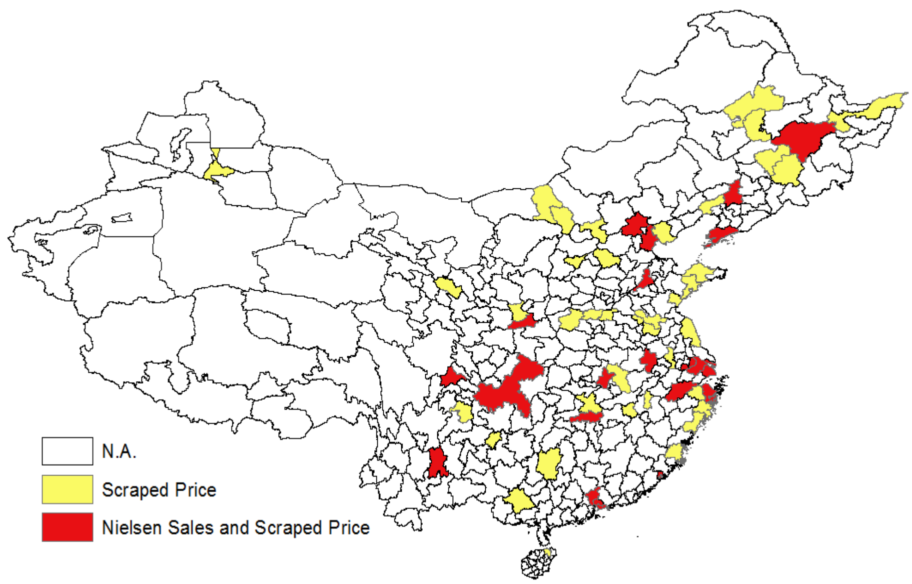


Figure 1: Regions included in Nielsen Sales Data and Scraped Price Data

4.2 Measuring Product Variety

As noted in the previous section, 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 D. The estimated expenditure shares are used to measure the terms λ_{fc} and λ_f appearing in the cost of living index (7).

[11]In the following analysis, we exchange RMB to USD using annual average exchange rate.

For the Sato-Vartia index appearing in that expression, however, we take a simpler approach and replace the Sato-Vartia weights with an *unweighted* geometric mean of the prices that are common across firms and cities. Using an unweighted geometric mean can be justified based on the new expression for a CES price index developed by Redding and Weinstein (2016). They allow for taste differences over time in their approach, which we will apply instead to taste differences across regions. 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. We shall measure common goods *separately* for the 60 cities in China and for the 377 MSAs of the United States. But when we want to measure the cost of living in China relative to the United States, then we will have to take a stand on how the taste parameters differ across these countries.

To briefly review the CES index derived by 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:

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

Because λ_{fc} defined in (6) equals $s_{fic}/s_{fic}(I_f)$, we can replace the share s_{fic} by $s_{fic} = s_{fic}(I_f)\lambda_{fc}$ in the above equation. Then we take the unweighted geometric mean across all products to obtain the formula in Redding and Weinstein (2016),

$$P_{fc} = \left[\prod_{i\in I_f} s_{fic}(I_f)^{\frac{1}{N_f(\sigma-1)}}\right]\left[\prod_{i\in I_f}\left(\frac{p_{fic}}{b_{fic}}\right)^{\frac{1}{N_f}}\right]\lambda_{fc}^{1/(\sigma-1)}. \tag{17}$$

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 (17), we obtain:

$$\begin{aligned} P_c = E(\mathbf{P}_c, F_c) = &\left[\prod_{f\in F} S_{fc}(F)^{\frac{1}{M(\eta-1)}}\prod_{i\in I_f} s_{fic}(I_f)^{\frac{1}{MN_f(\sigma-1)}}\right] \\ &\times\left[\prod_{f\in F}\prod_{i\in I_f}\left(\frac{p_{fic}}{b_{fic}}\right)^{\frac{1}{MN_f}}\right]\left[\prod_{f\in F}\lambda_{fc}^{1/M(\sigma-1)}\right]\lambda_c^{1/(\eta-1)}. \end{aligned} \tag{18}$$

This equation is the basis for our cost of living calculations. On the first line is an expression involving a geometric mean of firm and product shares. In our Chinese sample that expression will tend to be the same across cities, because we estimate the shares in some cities using the data for others, so the shares do not differ by much. Across countries, the shares in the first line will differ between China and the United States, and the first line reflects the product variety difference *within* the common goods for each country. We have not yet explored that term systematically, however, so it is not used in our calculations.

On the second line, the first expression is an unweighted geometric mean of prices measured relative to 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. It follows that the taste parameters cancel out when taking the ratio of (18) over two cities in the same country. But what about when we compare China with the United States? In that case we make use of the assumption also made by Redding and Weinstein (2016),

$$\prod_{f \in F} \prod_{i \in I_f} \left(b_{fic}\right)^{\frac{1}{MN_f}} = 1, \quad \forall c \in = 1, ..., D. \tag{19}$$

Within a country, this assumption is obviously weaker than assuming that the taste parameters for the common products of each firm are identical across cities. This assumption still guarantees that the taste parameters cancel out when taking the ratio of (18) over two cities in the same country. But that result holds equally well when comparing a city in China with a city in the United States: by applying (19) to the common products in *each* country, and then taking the ratio of (18), we make we are able to make a consistent comparison of the cost of living across countries. Furthermore, because the second line in (18) is decomposed into three multiplicative terms, we will be able to explore the extent to which each of terms accounts for the cost of living differences across cities and countries, as described in the next section.

4.3 Regression Equations

We follow the approach of the IO literature by using estimates of the demand elasticity ϵ along with the firms' shares to infer the markup; then the remaining marginal costs plus transport costs on the right will be estimated. We will estimate (6) by moving the markup term to the left:

$$\begin{aligned} p_{fid} \left[1 + \frac{1}{(\eta - 1)(1 - S_{fd})}\right]^{-1} &= g_{fi}(w_c) + T_{cd} \\ &= \alpha_i + \alpha_{cap} + \beta \ln Dist_{cd} + \epsilon_{fid}. \end{aligned} \tag{20}$$

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:

$$\lambda_{fd} = \frac{N_{min} p_{fd} x_{fd}}{N_{fd} p_{fd} x_{fd}} = \frac{N_{min}}{N_{fd}} = N_{min} \left(\frac{\sigma - 1}{\eta - 1} \right) \left[\frac{\eta - (\eta - 1) S_{fd}}{S_{fd}(1 - S_{fd})} \right] \frac{k_{fd}}{Y_d}, \tag{21}$$

using (15). The second regression equation comes from substituting $Y_d = \rho w_d L_d$ and taking logs,

$$\ln \lambda_{fd} = \gamma_{f,prov} + \delta_1 \ln \left[\frac{\eta - (\eta - 1) S_{fd}}{S_{fd}(1 - S_{fd})} \right] + \delta_2 \ln w_d + \delta_3 \ln L_d + \epsilon_{fd}. \tag{22}$$

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 (15). 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 (22) 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 (22) 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 shown 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 (20), so that we actually estimate,

$$p_{fid} \left[1 + \frac{1}{(\eta - 1)(1 - S_{fd})} \right]^{-1} = \alpha_i + \alpha_{cap} + \beta_1 \ln Dist_{cd} + \beta_2 \ln w_d + \beta_3 \ln L_d + \epsilon_{fid}. \tag{23}$$

A final set of regressions are used is to decompose the cost of living indexes into the three terms on the second line of (18), and study how each of these terms is related to the country size variables w_d and L_d. Specifically, we run the regressions:

$$\ln Z_d = \mu_1 + \mu_{cap} + \mu_2 \ln w_d + \mu_3 \ln L_d + \epsilon_d, \tag{24}$$

where Z_d denotes any of the three terms on the second line of (18).

[12] In fact, product scope N_{fd}in (15) is maximized for $\mathrm{S}_{fd} = \sqrt{\eta}/(1 + \sqrt{\eta})$.

5 Empirical Results

5.1 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 (23). 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.

Table 1: Price Regression of Toothpaste

	(1)		(2)		(3)		(4)	
Dep Var.	Price	P/mkup	Price	P/mkup	Unit-P	UP/mkup	Unit-P	UP/mkup
ln *Population*	-4.89***	-3.31***	-4.74***	-3.08***	-1.40***	-0.90***	-1.34***	-0.84***
	(0.48)	(0.32)	(0.48)	(0.32)	(0.17)	(0.11)	(0.17)	(0.11)
ln *Average Income*	-9.78***	-7.13***	-9.75***	-6.90***	-2.92***	-1.93***	-2.88***	-1.86***
	(0.67)	(0.42)	(0.67)	(0.43)	(0.27)	(0.17)	(0.27)	(0.16)
Capital City	0.41	0.45*	0.64	0.370	-0.05	0.03	0.02	0.02
	(0.41)	(0.26)	(0.42)	(0.26)	(0.13)	(0.08)	(0.13)	(0.08)
ln *Distance*			0.41**	0.95***			0.16***	0.25***
			(0.20)	(0.12)			(0.06)	(0.04)
Observations	37,828	37,828	36,433	36,433	37,828	37,828	36,433	36,433
Number of group	1,607	1,607	1,529	1,529	1,607	1,607	1,529	1,529
R-squared	0.03	0.03	0.03	0.04	0.02	0.02	0.02	0.03

Notes: All regressions include UPC fixed effects. *Average Income* is measured as GDP per capita in units of $1e4$ U.S. dollar and *Population* is in units of $1e4$. Prices (p and p_adj) are in units of U.S. cent and unit-prices (up and up_adj) are measured in U.S. cent per ounce. *Capital City* is dummy variable. *Distance* is in unit of *km*. Robust standard errors are clusters at group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

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 -4.89 to -3.31 when we switch from using the price as the dependent variable to using the price devided by markup. This finding implies a higher markup in smaller cities, as the markup-adjusted price (price devided by markup) rises less in smaller cities than the per-item price. We repeat the regressions but adding ln *Distance* in specification (2). As expected, the firm will charge a higher price if the destination is farther from the location of factory. The results in specification (2) continue to support the presence of a pro-competitive effect.

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 coefficents 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.

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

	(1) Laundry Detergent		(2) Personal Wash		(3) Shampoo	
Dep Var.	Price	P/mkup	Price	P/mkup	Price	P/mkup
ln *Population*	-2.94***	-2.35***	-1.87***	-1.62***	-3.51***	-2.60***
	(0.51)	(0.38)	(0.42)	(0.30)	(0.71)	(0.55)
ln *Average Income*	-3.69***	-2.50***	-1.66***	-1.25***	-3.60***	-3.23***
	(0.58)	(0.43)	(0.62)	(0.43)	(0.88)	(0.69)
Capital City	2.63***	1.59***	3.65***	2.50***	5.84***	4.37***
	(0.70)	(0.50)	(0.55)	(0.38)	(0.88)	(0.67)
Observations	46,860	46,860	68,752	68,752	57,591	57,591
Number of group	1,913	1,913	2,717	2,717	2,008	2,008
R-squared	0.00	0.00	0.00	0.00	0.00	0.00

Notes: All regressions include UPC fixed effects. *Average Income* is measured as GDP per capita in units of $1e4$ U.S. dollar and *Population* is in units of $1e4$. Prices (p and p_adj) are in units of U.S. cent and unit-prices (up and up_adj) are measured in U.S. cent per ounce. *Capital City* is dummy variable. Robust standard errors are clusters at group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

In Tables 2 and 3, we explore the pro-competitive effect for the other product categories (Laundry Detergent, Personal Wash items and Shampoo). Table 2 shows the price regressions with per-item price and price/markup as the dependent variables, and Table 3 uses unit-price. From these tables, we still observe a lower price in richer and bigger cities, and firms charge a lower markup for richer and bigger cities.

Table 4 exhibits the result of our second structural identification as shown in equation 22. 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 $\ln \lambda_{fc}$, where λ_{fc} 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.

Table 3: Price Regression for the Other Product Categories (Unit-price)

	(1) Laundry Detergent		(2) Personal Wash		(3) Shampoo	
Dep Var.	Unit-P	UP/mkup	Unit-P	UP/mkup	Unit-P	UP/mkup
ln *Population*	-0.07***	-0.05***	-0.14***	-0.11***	-0.25***	-0.18***
	(0.01)	(0.01)	(0.04)	(0.03)	(0.05)	(0.04)
ln *Average Income*	-0.06***	-0.04***	-0.08	-0.06	-0.09	-0.10
	(0.01)	(0.01)	(0.07)	(0.05)	(0.13)	(0.10)
Capital City	0.08***	0.05***	0.31***	0.21***	0.54***	0.41***
	(0.02)	(0.01)	(0.05)	(0.04)	(0.08)	(0.06)
Observations	46,860	46,860	68,752	68,752	57,591	57,591
Number of group	1,913	1,913	2,717	2,717	2,008	2,008
R-squared	0.00	0.00	0.00	0.00	0.00	0.00

Notes: All regressions include UPC fixed effects. *Average Income* is measured as GDP per capita in units of 1*e*4 U.S. dollar and *Population* is in units of 1*e*4. Prices (p and p_adj) are in units of U.S. cent and unit-prices (up and up_adj) are measured in U.S. cent per ounce. *Capital City* is dummy variable. Robust standard errors are clusters at group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Firm Share Regression

	(1)	(2)	(3)	(4)
Dep Var: ln λ_{fd}	Toothpaste	Laundry Detergent	Personal Wash	Shampoo
ln *Population*	-0.150*	-0.236***	-0.295***	-0.156***
	(0.088)	(0.030)	(0.034)	(0.021)
ln *Average Income*	0.242***	-0.244***	-0.159***	-0.083***
	(0.072)	(0.030)	(0.030)	(0.012)
Capital City	-0.054	-0.179***	-0.136***	-0.069***
	(0.059)	(0.029)	(0.028)	(0.015)
$\ln \frac{\eta-(\eta-1)S_{fd}}{S_{fd}(1-S_{fd})}$	0.003	0.609***	0.396***	0.002
	(0.045)	(0.070)	(0.061)	(0.006)
Observations	660	420	600	840
Number of group	308	196	280	392
R-squared	0.046	0.662	0.586	0.348

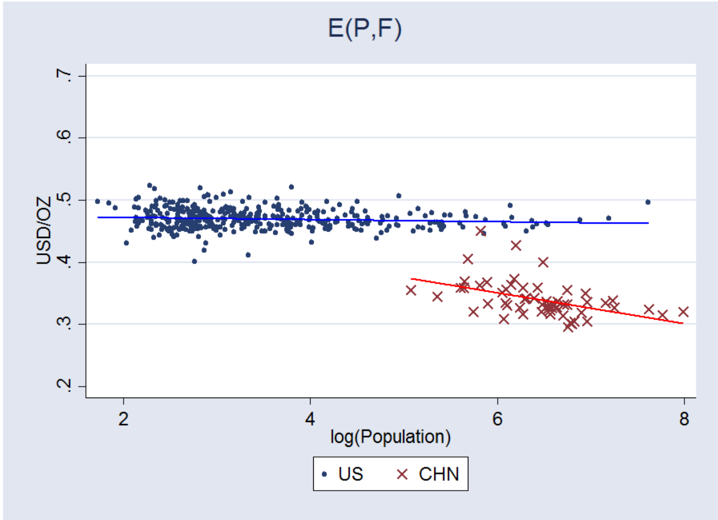
Notes: All regressions include firm-province fixed effects. *Average Income* is measured as GDP per capita in units of 1*e*4 U.S. dollar and *Population* is in units of 1*e*4. *Capital City* is dummy variable. Robust standard errors are clusters at group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

5.2 Cost of Living Indexes

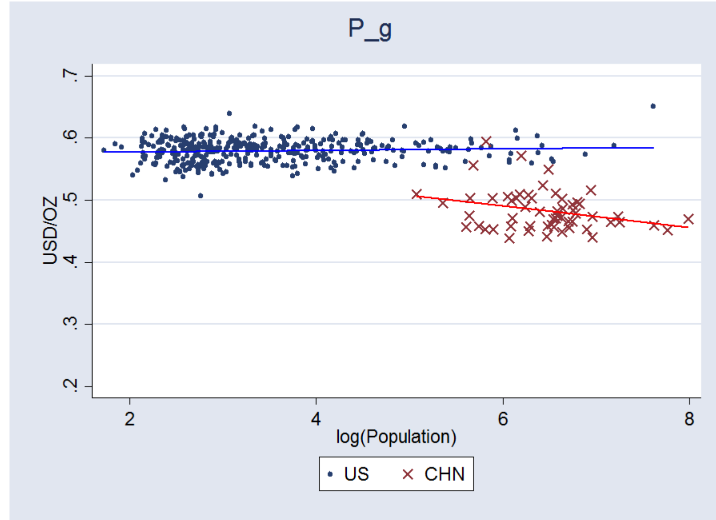
Given the predicted product shares, we calculate the cost-of-living index according to the second line of (18), in terms of unit-price (U.S. cents/oz). We also make use of assumption (19), so that the taste parameters in (18) do not appear. The results are plotted in Figure 2 for Toothpaste, Figure 3 for Laundry Detergent, Figure 4 for Personal Wash items, and Figure 5 for Shampoo. In each panel and Figure, we compare the variables between China and the United States. Working counter-clockwise, the bottom-left panel exhibits the relationship between the expenditure share on common set of firms, λ_c as defined in (9), and city population. The bottom-right panel compares city's average firm sales share of common products, $\overline{\lambda}_{fc} \equiv \prod_{f \in F} (\lambda_{fc})^{W_f(F)}$, and the top-right panel is the geometric mean of the common-goods price index, $P_c^g \equiv \prod_{f \in F} \prod_{i \in I_f} (p_{fic})^{W_f(F) w_i(I_f)}$. Lastly, the variety adjusted cost-of-living index, which combines these three terms as in the second line of (18), is plotted in the top-left panel.

As demonstrated in Figure 2 for Toothpaste, smaller cities have a higher expenditure share on common products from firms, $\overline{\lambda}_{fc}$, and purchase more from the common set of firms, λ_c. That result holds in both China (shown in red) and the United States (shown in blue). Because variety is inversely related to the common-good or common-firm shares, these results indicate that larger cities have more variety. That tendency is more pronounced in China than that in the U.S., as the common-firm share, λ_c, decreases more rapidly with city population in China.

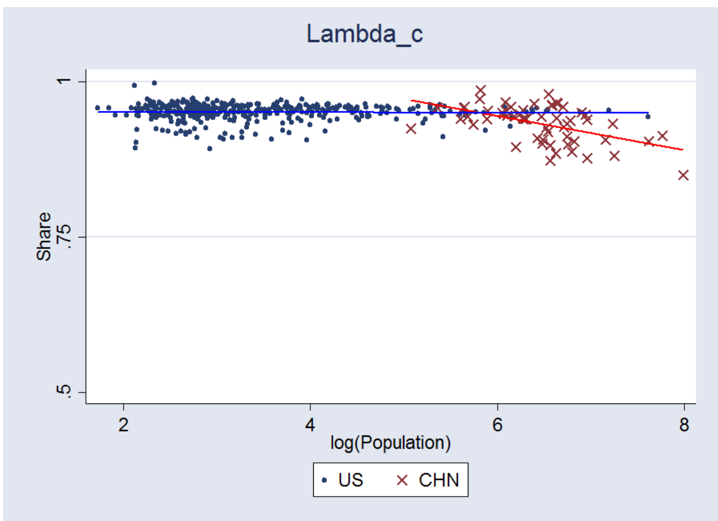
For the products that are sold in all cities, there is no significant price difference between cities in the U.S., as shown by the geometric mean P_c^g, whereas in China there is a clear downward sloping pattern of prices, i.e., the unit price is higher in smaller cities in China. This confirms the pro-competitive effect in large cities. The greater entry of firms into larger cities decreases the expenditure share on common firms and, in turn, leads to a fall in cost of living. After taking everything (price and variety) into consideration, the cost of living is higher in small cities than that in large ones. Similar patterns are also found in Laundry Detergent, Personal Wash items and Shampoo.



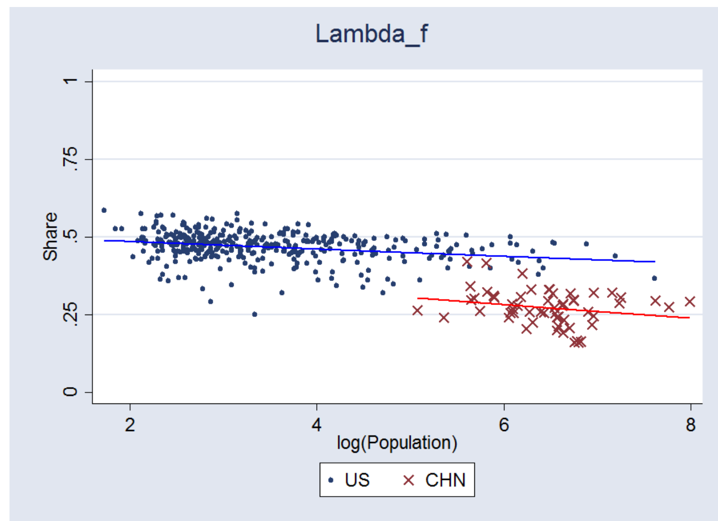
(a) Cost-of-living Index



(b) Common Products Price Index



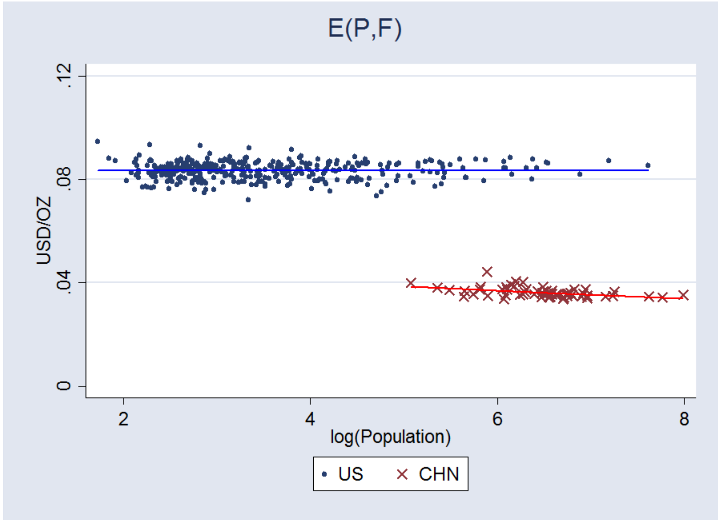
(c) Common Firm Sales Share



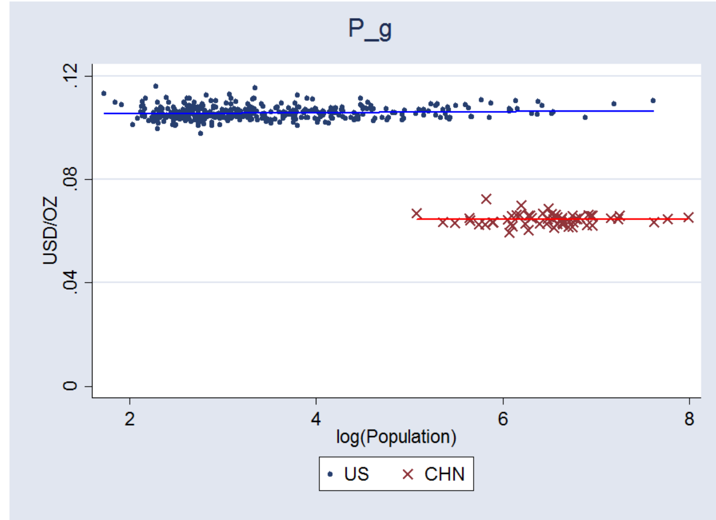
(d) Within-firm Common Products Sales Share

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

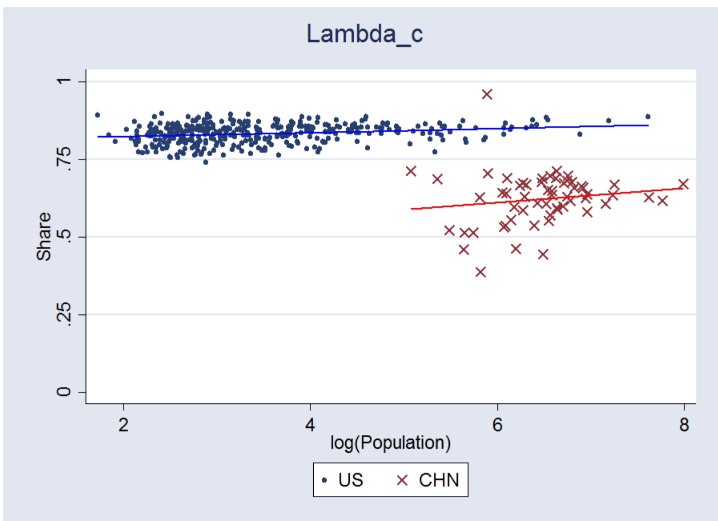
Figure 2: Cost-of-living Comparision: **Toothpaste**



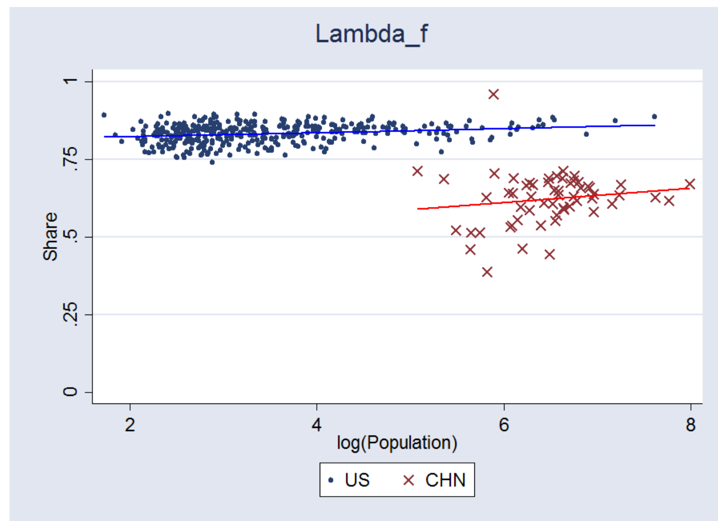
(a) Cost-of-living Index



(b) Common Products Price Index



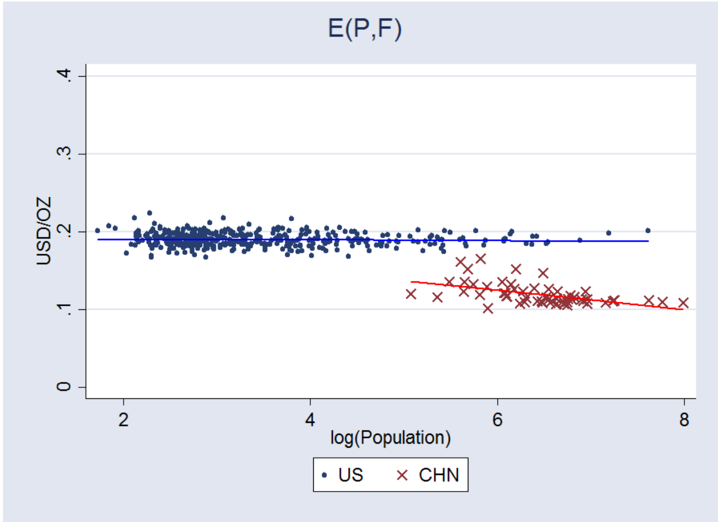
(c) Common Firm Sales Share



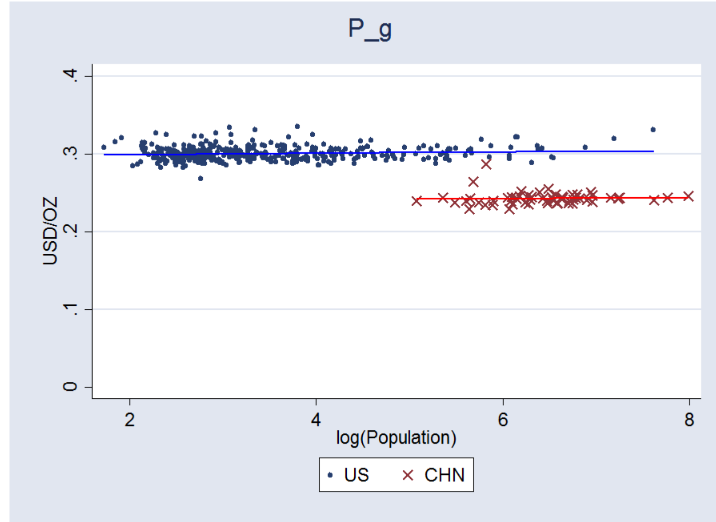
(d) Within-firm Common Products Sales Share

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

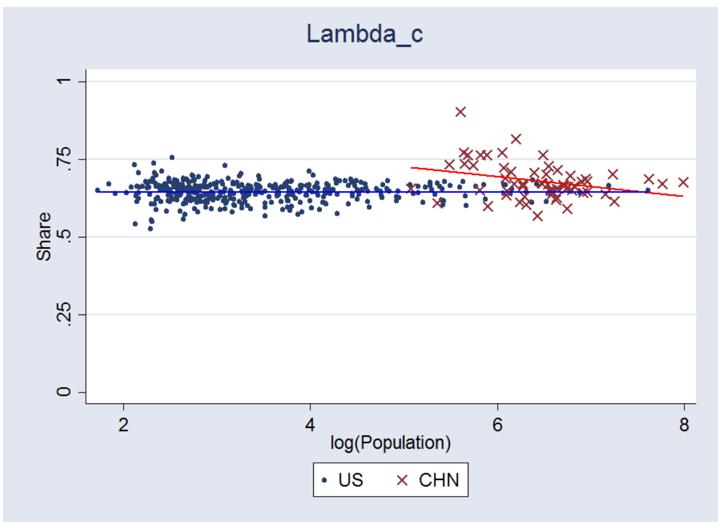
Figure 3: Cost-of-living Comparision: **Laundry Detergent**



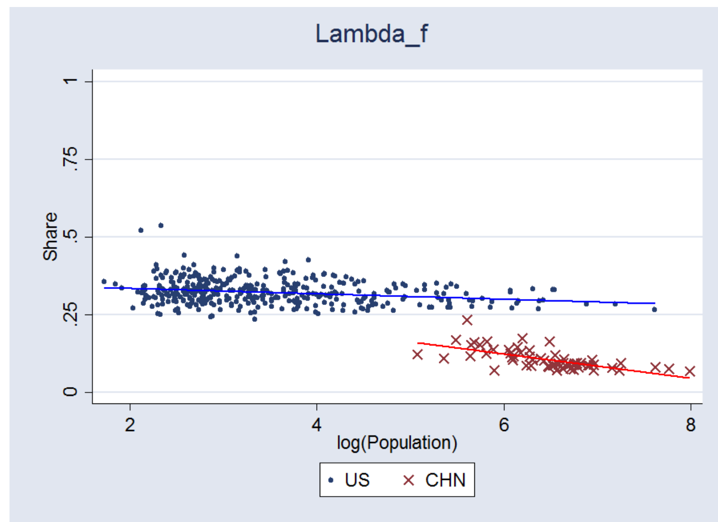
(a) Cost-of-living Index



(b) Common Products Price Index



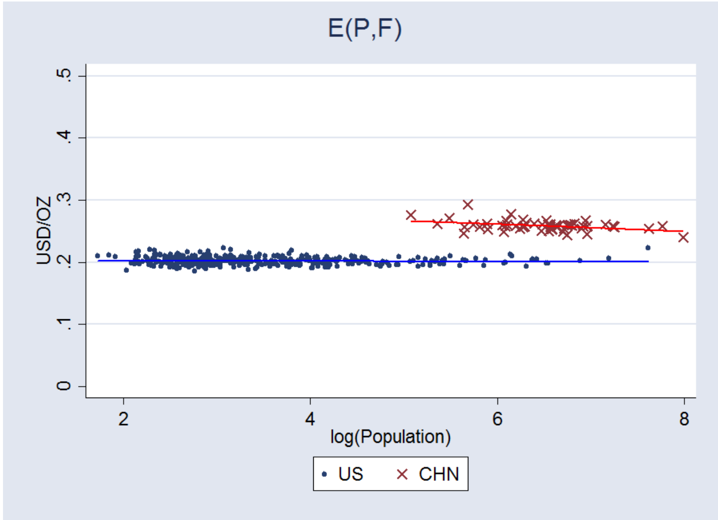
(c) Common Firm Sales Share



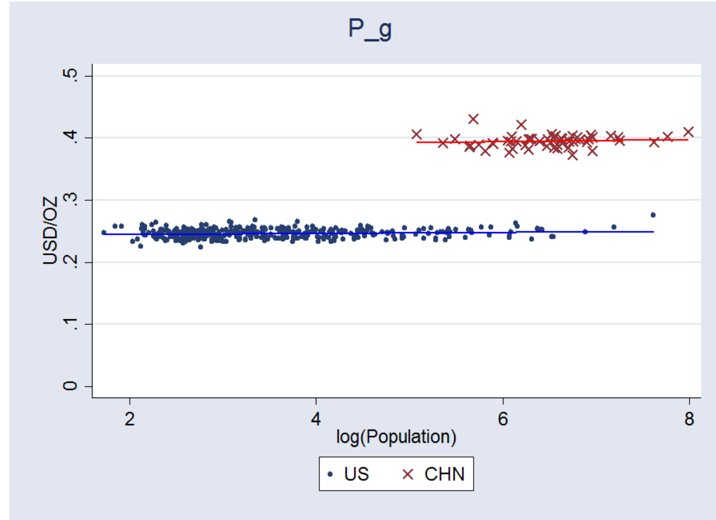
(d) Within-firm Common Products Sales Share

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

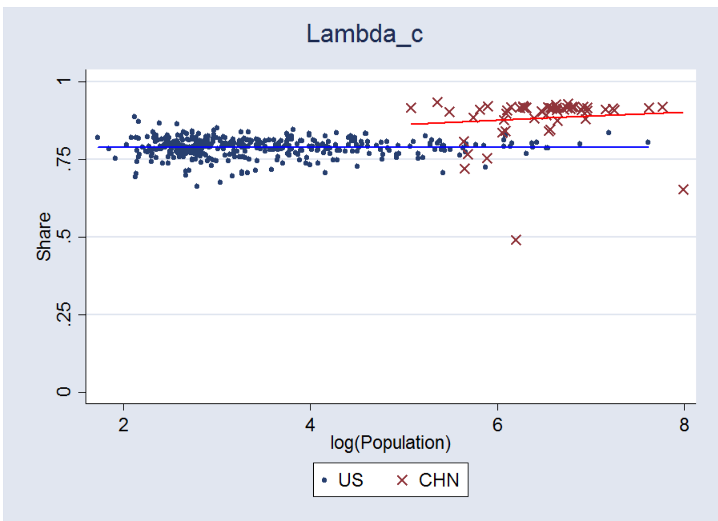
Figure 4: Cost-of-living Comparision: **Personal Wash Items**



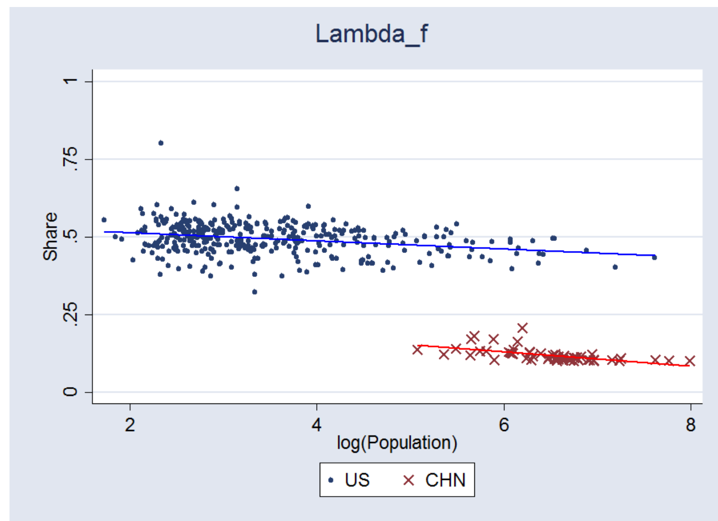
(a) Cost-of-living Index



(b) Common Products Price Index



(c) Common Firm Sales Share



(d) Within-firm Common Products Sales Share

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

Figure 5: Cost-of-living Comparision: **Shampoo**

We further calculate the relative cost of living in each China city relative to the average for the United States,[13] and plot them 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.[14] We see from Figure 6 to 9 that different product categories exhibit different relative price patterns. Laundry Detergent, Personal Wash items, and Toothpaste are cheaper in the China (the relative price varies from 42 to 96%, with U.S. average as unity), whereas prices of shampoo are 24% to 45% higher than than in the United States. Secondly, price levels are higher in inland cities than in the eastern coast cities that are usually larger and wealthier. For example, the most expensive cities for Toothpaste, Laundry Detergent, Personal Wash items, and Shampoo are Urumqi (0.96), Lanzhou (0.53), Urumqi (0.87), and Hohhot (1.45) respectively. The least expensive cities for the same order of product categories are Xuzhou (0.63), Jilin (0.40), Xiamen (0.54) and Chongqing (1.19). The geographic disparity reinforces our concern that the cost of living in China is *higher* than found by just examining prices in larger cities.

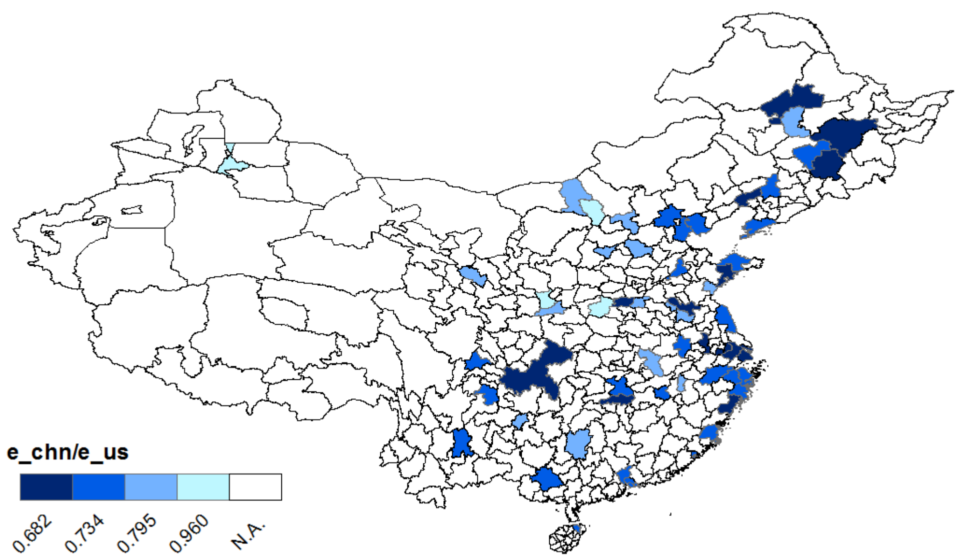


Figure 6: Relative Cost-of-living in China: **Toothpaste**

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

[14]The price follows an ascending order for colors ranking as dark blue, light blue, light red, and dark red.

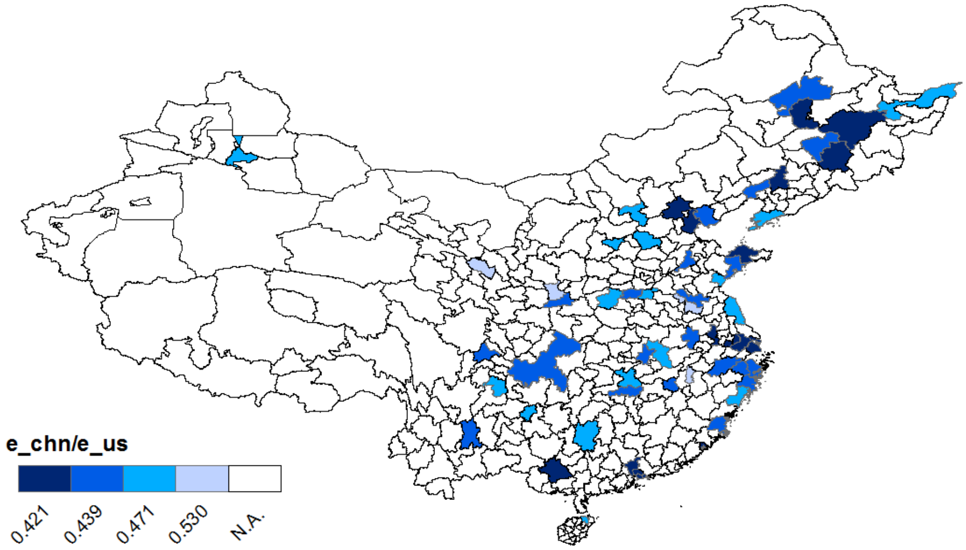


Figure 7: Relative Cost-of-living in China: **Laundry Detergent**

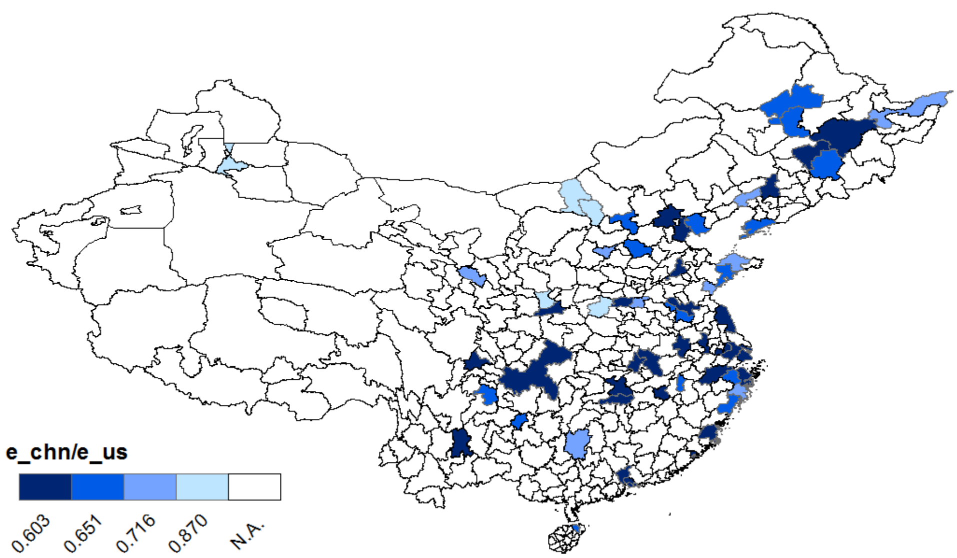


Figure 8: Relative Cost-of-living in China: **Personal Wash Items**

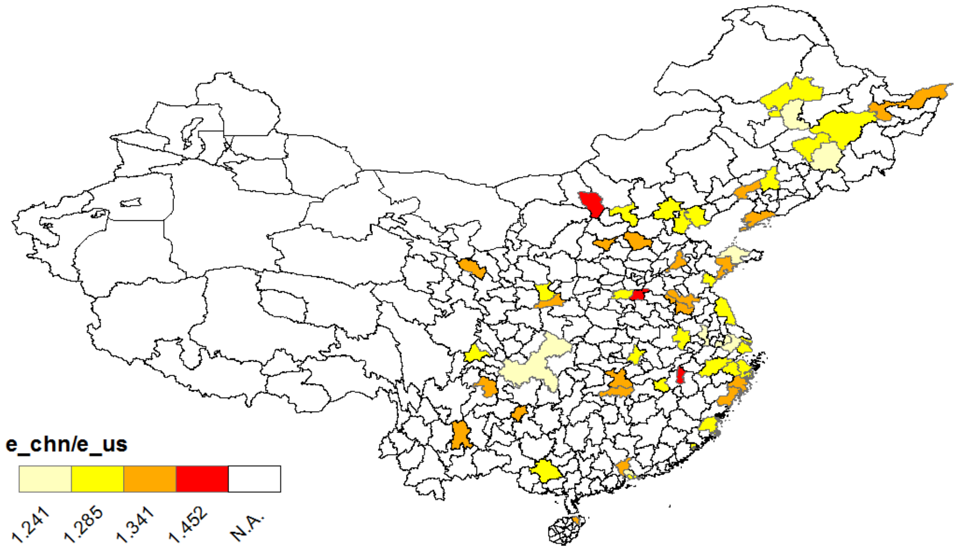


Figure 9: Relative Cost-of-living in China: **Shampoo**

5.3 Decomposition of Cost of Living Indexes

Motivated by our cost-of-living index as shown in the second line of (18), for each product category we explore to what extent the three components (denoted as P_c^g, $\overline{\lambda}_{fc}$ and λ_c) contribute to the overall cost of living. Tables 5 presents the regression results of the three components on total market size (GDP),[15] using the predicted shares. For all product categories, each observation is one city.[16] The first row shows that cities with twice the market size have a 5.2% lower cost of living for Toothpaste, 4.2% lower for Laundry detergent, 5.5% lower for Personal Wash items, and 2% lower for Shampoo. The second to fourth rows indicate the percentage that each margin contributes to the lower cost of living in larger cities. For example, in Toothpaste, 82% of the drop in the cost of living occurs because of the decrease in the prices of common goods P_c^g (pro-competitive effect), 6% due to the average spending on the common products from firms, $\overline{\lambda}_{fc}$, and 12% due to the spending on the common set of firms, λ_c. For the other products, the price drop in larger cities account for 14%, 7% and 15% of the overall lower cost of living for Laundry Detergent, Personal Wash items and Shampoo, respectively. We also obverve significant variety benefit in larger cities. The contribution of $\overline{\lambda}_{fc}$ to the lower cost of living in larger cities varies from 6% to 110%.

[15] In each regression, we also include a dummy variable for the capital city which is omitted in the outcome table.

[16] Besides the 22 cities included in the Nielsen (China) database, we have predicted the firms shares for the other 38 cities as explained in Appendix. In Table 5, however, we exclude cities that have the very imprecise predictions for their firm shares, so that the number of cities used in each regression ranges from 55 to 60.

Table 5: Decomposition of Cost-of-living with Aggregate Market Size

	Toothpaste		Laundry Detergent		Personal Wash		Shampoo	
Dep.↓\Indep.→	$\ln Y$	R^2	$\ln Y$	R^2	$\ln Y$	R^2	$\ln Y$	R^2
$\ln E(P_{fc}, F_c)$	-0.052***	0.233	-0.042***	0.422	-0.055**	0.214	-0.020***	0.234
	(0.013)		(0.007)		(0.016)		(0.005)	
$\ln P_c^g$	-0.042***	0.260	-0.006	0.056	-0.004	0.060	-0.002	0.139
	(0.010)		(0.006)		(0.005)		(0.004)	
	82%		*14%*		*7%*		*15%*	
$\ln \lambda_{fc}$	-0.003	0.005	-0.034***	0.517	-0.049***	0.471	-0.022***	0.462
	(0.009)		(0.007)		(0.009)		(0.004)	
	6%		*81%*		*88%*		*110%*	
$\ln \lambda_c$	-0.006*	0.145	-0.001	0.098	-0.003	0.021	0.005	0.018
	(0.003)		(0.008)		(0.006)		(0.005)	
	12%		*5%*		*55%*		*-25%*	

Notes: All regregions include a *Capital* dummy 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 in units of $1e8$ U.S. dollar. Standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

6 Conclusions

[to be added]

Appendix

A. Theory Appendix

A1. Choice of price by the firm

The solution to the problem (12) w.r.t. the price p_{fid} is

$$x_{fid} + \sum_{j \in I_{fd}} [p_{fid} - g_{fi}(w_c) - T_{cd}] \frac{dx_{fjd}}{dp_{fjd}} = 0, \quad \forall i \in I_{fd} \tag{25}$$

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(25) and dividing by demand x_{fid}, we can re-express the condition as

$$1 + \sum_{j \in I_{fd}} \left[1 - \frac{g_{fj}(w_c) + T_{cd}}{p_{fjd}}\right] \frac{d \ln x_{fid}}{d \ln p_{fjd}} = 0, \quad \forall i \in I_{fd} \tag{26}$$

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 (26) 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 (26) is

$$\left(\frac{\mu_{fd} - 1}{\mu_{fd}}\right) = -\left(\frac{d \ln x_{fid}}{d \ln p_{fjd}}\right)^{-1} \tag{27}$$

Expression (27) 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 (27) 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[17]

$$-\sum_{j \in I_{fd}} \frac{d \ln x_{fid}}{d \ln p_{fjd}} = \eta - (\eta - 1) S_{fd} > 1 \tag{28}$$

Notice that the expression on the right-hand side of (28) is precisely what we get from (11) 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 (13).

[17] 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(11), (28) is obtained.

A2. Choice of variety by the firm

The first-order condition of (14) with respect to N_{fd} yields[18]

$$[p_{fd} - g_f(w_c) - T_{cd}]x_{fd} + N_{fd}[p_{fd} - g_{fi}(w_c) - T_{cd}]\frac{dx_{fd}}{dN_{fd}} = k_{fd} \tag{29}$$

Using demand function, the elasticity of demand w.r.t. product scope $N_{fd}$is given in (30):

$$\begin{aligned}\frac{d\ln x_{fd}}{d\ln N_{fd}} &= (\sigma - \eta)\frac{d\ln P_{fd}}{d\ln N_{fd}} + (\eta - 1)\frac{d\ln P_d}{d\ln P_{fd}}\frac{d\ln P_{fd}}{d\ln N_{fd}} \\ &= \frac{\sigma - \eta + (\eta - 1)S_{fd}}{1-\sigma}\end{aligned} \tag{30}$$

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 (29) can be rewritten as(15),

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

A3. Model of core-competency and iceberg trade cost detailed

In this section, we resolve firm's problem with heterogenous 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,

$$\max_{N_{fd}, \{p_{fd}(\omega)|\omega\in[0,N_{fd}]\}} \sum_{d=1}^{D}\left\{\int_0^{N_{fd}}[p_{fd}(\omega) - \tau_{cd}g_f(\omega)]x_{fd}(\omega)d\omega - k_{fd}N_{fd} - K_{fd}\right\}\mathbf{1}(N_{fd} > 0). \tag{31}$$

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

$$p_{fd}(\omega) = \left[1 + \frac{1}{(\eta - 1)(1 - S_{fd})}\right]\tau_{cd}g_f(\omega)RelativeCost-of-livinginChina:Shampoo, \quad \forall\omega\in[0, N_{fd}]. \tag{32}$$

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

$$[p_{fd}(N_{fd}) - \tau_{cd}g_f(N_{fd})]x_{fd}(N_{fd}) + \int_0^{N_{fd}}[p_{fd}(\omega) - \tau_{cd}g_f(\omega)]\frac{d\ln x_{fd}(\omega)}{dN_{fd}}d\omega = k_{fd}. \tag{33}$$

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

$$\begin{aligned}\frac{d\ln x_{fd}}{d\ln N_{fd}} &= (\sigma - \eta)\frac{d\ln P_{fd}}{d\ln N_{fd}} + (\eta - 1)\frac{d\ln P_d}{d\ln P_{fd}}\frac{d\ln P_{fd}}{d\ln N_{fd}} \\ &= \frac{\sigma - \eta + (\eta - 1)S_{fd}}{1-\sigma}s_{fd}(N_{fd})N_{fd}.\end{aligned} \tag{34}$$

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

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 (33) and (34), the optimal scope is derived as,

$$H(N_{fd}) = \frac{\eta - 1}{\sigma - 1}\left[\frac{S_{fd}(1-S_{fd})}{\eta - (\eta-1)S_{fd}}\right]\frac{Y_d}{k_{fd}}. \tag{35}$$

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

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

B. Collection Method of 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.

B1. Introduction to *Wochacha APP*

The mobile phone application *Wochacha* was developed by Wochacha Info Tech Co. Ltd in January 2010, which received a capital injection from American Sequoia Capital, Ivy Capital, and others.. Detailed information regarding *Wochacha* can be obtained from its webpage.[19] 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 ithe lowest price. Consumers could also search for price quotes based on the UPC (or EAN) of product in their resident city.[20] 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 ways to collecting the retailing 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 the partnership with the supermarket, based on which *Wochacha* directly imports the scanner data from them. In either way, the collected prices are not published until *Wochacha* data auditors have rexamined them.

[19] The official webpage: http://www.wochacha.com/about/index.

[20] 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.

B2. Data Collection Example

Figure 10 exhibits the results when one search for prices quotes of Pampers[21]in Shanghai. Based on the results, there are eight retail price quotes varying form 112*RMB* by Walmart to 165*RMB* 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.

While it is a public source of retail price information, it is extremely time-consuming to directly collect these data by hand (which is how we started when we initially collected these data with undergraduates research assistants). Instead, 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 the sampled cities to construct our retailing price database. The collection took place from October to December in 2015.

Figure 10: Interface of *Wochacha* Application

B3. Cities for Analysis

Cities that we scraped for price data are listed in Table 6. We include 60 cities from 28 provinces (or municipalities), and they are widely spread across China.

C. Elasticity of Substitution

We directly use the elasticities of substitution for Toothpaste, Laundry Detergent, Personal Wash items and Shampoo from Hottman, Redding, and Weinstein (2014), which are listed in Table 7.

[21]The EAN-13 is 6903148040737.

D. Prediction of the Product Market Share

D1. Prediction Method

The cost-of-living calculation requires firm sales information, whereas the Nielsen (China) sales database only provide us with 22 cities. To allow our analysis also covering the small cities that only have price information from scraping mobile phone application, we implement the Heckman two-step approach to estimate the product share in those cities. To do so, we firstly apply a two-step Heckman approach to the estimation of market share using the Nielsen (China) sales database, as shown below:

$$
\begin{aligned}
Prob(Share_{ic} > 0) = \alpha_0 + \alpha_1 \ln Y_c + \alpha_2 \ln L_c + \alpha_3 \ln N_c + \alpha_4 Capital_c \\
+ \delta_i Variety_i + \gamma_{1r} Region_r + \epsilon_{ic},
\end{aligned} \tag{36}
$$

$$
\begin{aligned}
E(\ln Share_{ic} | Share_{ic} > 0) = \beta_0 + \beta_1 Capital_c + \phi_{ic} Brand_i \times \ln Y_c \\
+ \mu_{ic} Brand_i \times \ln L_c + \psi_{ic} Brand_i \times \ln N_c + \gamma_{2r} Region_r,
\end{aligned} \tag{37}
$$

where equation (36) is the selection equation and equation (37) 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-13 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 e.t.c.; $Variety_i$ is a collection of variety categories which is constructed by jointly combining several product characteristics[22]; $Share_{ic}$ is calculated as product i's sales share in city c. We use the two-step Heckman approach 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 also do not include product price in the share specification (37).

For each of the four product categories, we then implement the two-step Heckman method in the Nielsen (China) sales database of 22 cities. After obtaining the coefficients $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}, \hat{\phi}, \hat{\mu}, \hat{\psi})$, we predict the market share for the products with scraped price and 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).

D2. Filter Criteria

Though the above 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

[22] 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*

this situation, we exclude cities, which have the extremely low city's average firm sales share of common products ($\prod_{f\in F}(\lambda_{fc})^{W_f(F)}$) from our studied sample. 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 exclude the cities from our sample if its prediction error is twice as large than the standard deviation of the prediction errors, i.e., we exclude city c if $|\hat{u}_c| > 2 \times std(\hat{u})$[23]. After implementing the criteria, we finally have 58 cities for Toothpaste and Laundry Detergent, 59 cities for Personal Wash items, and 55 cities for Shampoo (out of 60 cities).

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

Table 6: 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.

Table 7: 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

References

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*, pp. 157–177.

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 (2014): "Quantifying the Sources of Firm Heterogeneity," NBER Working Paper 20436.

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*.

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.