Who Faces Higher Prices? An Empirical Analysis Based on Japanese Homescan Data

Naohito Abe[1, 3] and Kyosuke Shiotani[2]
[1] Institute of Economic Research, Hitotsubashi University, Tokyo, JAPAN
[2] Bank of Japan, JAPAN
[3] Corresponding Author: Naohito Abe; email: nabe@ier.hit-u.ac.jp

Abstract

On the basis of household-level scanner data (called homescan) for Japan over a three-year period, we construct a household-level price index and investigate the causes of price differences across households. Similar to the result of Aguiar and Hurst (2007), we observe large price differentials across households. However, the differences across age and income groups are small. In addition, we find that elderly people face higher prices than the younger ones, which is contrary to the results of Aguiar and Hurst (2007). The most important determinant of the price level is the reliance on bargain sales; an increase in the purchase of goods at bargain sales by one standard deviation decreases the price level by more than 0.9%, while shopping frequency only has limited effects on the price level.

Keywords: homescan data, household-level price index, shopping behavior

1. Introduction

Owing to recent technological developments in data creation, numerous researchers of commodity prices have begun to use not only traditional aggregates, such as the consumer price index, but also information on micro-level commodity prices. To date, commodity-level price information is used in various economic fields, such as macroeconomics (Nakamura and Steinsson (2007)), international economics (Haskel and Wolf (2001)), and industrial economics (Bay et al. (2004)). Recently, on the basis of commodity-level homescan data, Aguiar and Hurst (2007) (hereafter AH) found that there is a violation of the law of one price across different age groups. More precisely, in the US, elderly families face lower prices for the same commodities than younger families. AH interpret their results in line with the standard life cycle model of consumption with endogenous decisions of shopping time. The mechanism is simple. Since the opportunity costs of shopping for retired people are lower than those for the younger ones, the elderly tend to shop more to find lower prices, which results in the violation of the law of one price.

This study considers the relationship between shopping behaviors and price level on the basis of commodity-level homescan data for Japan. Although we find that each commodity is traded at different prices, the differences in price levels across age or income groups are very small. Other than income and age, we find several important determinants of the price index. The most important of these is the ratio of purchases at bargain sales. By increasing their purchase at bargain sales by one standard deviation, people can enjoy a reduction of 0.9% in the price level, which is consistent with Griffith et al. (2009) who found a significant amount of savings from purchasing at bargain sales in the UK. Although other shopping behaviors—such as frequency of shopping, degree of mass purchasing, or preference for high-quality goods—are all statistically significant, these behaviors are not quantitatively important. Our empirical results suggest that further investigation into shopping strategy, particularly, the determinants of purchasing at bargain sales, is necessary to understand the mechanism underlying the price level differential across families.

2. Data

The data for this study are from the "Household Consumer Panel Research" (hereafter SCI) data set compiled by Intage, a marketing company in Japan. The SCI contains the daily shopping information of approximately 12,000 households, randomly selected from all prefectures (except Okinawa) in Japan. The sample households are restricted to married couples. Using a barcode reader, households are asked to scan the barcode of every commodity they purchase. In the SCI, the following information is available for every commodity purchased: (1) Japanese Article Number (JAN)—a unique commodity identifier, (2) date of purchase, (3) price and quantity, and (4) store name from which the commodity was purchased. Fresh foods (e.g., meat, fish, and vegetables) without barcodes are excluded. This limitation is shared by the homescan data of AC Nielsen in the US. The data we use in this study encompasses the period from 2004 to 2006. Table 1 represents the family composition.

3. Relative Price Index

Following AH, we construct the price index in the following manner. Let us consider a commodity that belongs to a product category $c \in C$. We denote the price of good $i \in I_c$ purchased by household $j \in J$ on date $t \in T$ by $p_{i,t}^{j,c}$ and the quantity by $y_{i,t}^{j,c}$. Then, the total expenditure by the household during time interval m can be written as

$$X_m^j = \sum_{c \in C, i \in I_c, t \in m} p_{i,t}^{j,c} y_{i,t}^{j,c}.$$

If the household purchases each product at the average price, the expenditure would be

$$\overline{X}_m^j = \sum_{c \in C, i \in I_c, t \in m} \overline{p}^c{}_{i,m} y_{i,t}^{j,c},$$

where

$$\overline{p}^c{}_{i,m} = \sum_{j \in J, t \in m} p_{i,t}^{j,c} \frac{y_{i,t}^{j,c}}{\sum_{j \in J, t \in m} y_{i,t}^{j,c}}$$

is the weighted average price paid for a good i in category c during time interval m. Further, we define the price index for the household as the ratio of actual expenditure divided by the expenditure at the average price $\overline{p}^c{}_{i,m}$ in the following manner:

$$\tilde{p}_m^j \equiv \frac{X_m^j}{\overline{X}_m^j}.$$

Finally, we normalize the index by dividing by the average price index within the month to obtain

$$p_m^j \equiv \frac{\tilde{p}_m^j}{\frac{1}{J}\sum_j \tilde{p}_m^j}.$$

This household-level price index shows the relative price faced by each household to the average price.

4. Shopping Behavior

One of the main results of AH is that elderly people can lower their prices by increasing their shopping frequency. In this section, in addition to the shopping frequency, we introduce other types of shopping behavior that might affect the relative price index introduced in the previous section.

Shopping frequency: (ln trip)

We used the number of stores households use as the measure of shopping frequency. More precisely, we first counted the number of different stores visited by a sample household each day. Then, we calculated the sum of the number for each month, which yields the index for the degree of shopping frequency.

The number of different stores: (ln stores)

This measure captures the variety of shops used for shopping by each household. Note that this variable does not include information regarding frequent shopping at the same store.

HHI: (ln HHI)

Next, we constructed the Herfindahl–Hirschman Index (HHI), such that

$$HHI_m^j \equiv \sum_{k=1}^{K} {S_{k,m}^j}^2,$$

where $S_{k,m}^j$ is the share of store $k \in \mathrm{K}$ in the monthly total purchases of household j.

The total number of goods bought by a household: (ln quantity)

The monthly total number of goods purchased by a household is expressed in the following manner.

$$Quantity_m^j = \sum_{c \in C, i \in I_c, t \in m} y_{i,t}^{j,c}.$$

It is reasonable to assume that a family purchasing numerous goods can enjoy higher volume discounts, thereby decreasing the price level.

Non-bargain index: (non_bargain)

To observe the effect of purchasing at bargain sales, we constructed a measure for bargains. As expected, a household can decrease its price index by purchasing more goods at bargain sales. The index of purchasing at a non-bargain price is defined as

$$non\ bargain_m^j = \frac{\sum_{c \in C, i \in I_c, t \in m} I\left(P_{i,t}^{j,c}\right) p_{i,t}^{j,c} y_{i,t}^{j,c}}{\sum_{c \in C, i \in I_c, t \in m} p_{i,t}^{j,c} y_{i,t}^{j,c}},$$

where $min\ P_{i,t}^c$ is the store-level monthly minimum price of commodity i, and

$$I\left(P_{i,t}^{j,c}\right) = \begin{cases} 1, & P_{i,t}^{j,c} > min\ P_{i,t}^c \\ 0, & Otherwise \end{cases}.$$

Store choice index: (ln store_choice)

We define the index for the quality of each store, $k \in \mathrm{K}$, by basically following the procedure for the relative price index. Store quality index is the ratio of hypothetical sales when the store sells the goods at the average commodity-level price $\bar{P}_{i,m}^c$ to the sales when the goods are sold at their average categorical price. More precisely, first, we obtain the average price for a given good in category $c \in C$ as

$$\bar{P}_m^c = \sum_{i \in I_c, k \in K, t \in m} p_{i,t}^{k,c} \frac{y_{i,t}^{k,c}}{\sum_{i \in I_c, k \in K, t \in m} y_{i,t}^{k,c}}.$$

Note that $\bar{P}_m^c$ does not contain the subscript *i*. Next, assuming that stores sell the goods in each category at the average categorical price, we calculate the hypothetical total sales in the following manner:

$$\bar{Z}_m^k = \sum_{c \in C, i \in I_c, t \in m} \bar{p}_m^c y_{i,t}^{k,c}.$$

Then, we calculate the total sales of store *k* if it sells goods at the average commodity-level price:

$$\bar{p}_{i,m}^c = \sum_{k \in K, t \in m} p_{i,t}^{k,c} \frac{y_{i,t}^{k,c}}{\sum_{k \in K, t \in m} y_{i,t}^{k,c}},$$
$$Z_m^k = \sum_{c \in C, i \in I_c, t \in m} \bar{p}_{i,m}^c y_{i,t}^{k,c}.$$

Note that $\bar{p}_{i,m}^c$ contains the subscript *i*. Now, the index for the quality of goods sold in store *k* is defined as

$$\tilde{q}_m^k \equiv \frac{Z_m^k}{\bar{Z}_m^k}.$$

Finally, we normalize the index by dividing by the average monthly quality index in the following manner:

$$q_m^{\mathrm{k}} \equiv \frac{\tilde{q}_m^k}{\sum_{k \in K} \tilde{q}_m^k},$$

which yields the quality index of store *k* during time interval *m*.

We employ the average of the store quality index weighted by the share of each store in monthly total purchases of a household *j*:

$$Store\ \mathrm{choice}_m^j \equiv \sum_{k \in K} S_{k,m}^j q_m^k.$$

The higher the store choice index, the greater is the likelihood of using luxury stores, which leads to a higher price index.

Quality index: (ln quality)

By changing “store” to “household” in the previous index, we created the household-level monthly average quality index. Quality index for households is defined as the ratio of the hypothetical expenditure when the household purchases goods at the average commodity-level price $\bar{P}_{i,m}^c$ to the expenditure when the goods are purchased at their categorical average price $\bar{P}_m^c$. Because this is straightforward, we do not provide the precise definition of the household quality index.

$$q_m^{\mathrm{j}} \equiv \frac{\tilde{q}_m^j}{\sum_{j \in J} \tilde{q}_m^j}.$$

5. The Relationship Between the Relative Price Index and Shopping Behavior

Table 1: Household Price Index and Shopping Strategy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lntrip	0.0027	-0.0024						
	(8.603)	(-12.230)						
lnstore	0.0019		-0.0008					
	(6.523)		(-4.196)					
lnHHI	-0.0014			0.0012				
	(-4.663)			(5.105)				
lnquantity	-0.0108				-0.0098			
	(-37.399)				(-45.968)			
non_bargain	0.0594					0.0567		
	(99.182)					(96.201)		
lnstore_choice	0.0163						0.0205	
	(13.531)						(17.452)	
lnquality	0.0190							0.0236
	(34.243)							(44.338)
Dummy for Income (1)								
4,000–5,490	0.0001	-0.0001	-0.0001	-0.0001	-0.0001	0.0001	-0.0001	-0.0001
	(0.171)	(-0.111)	(-0.103)	(-0.098)	(-0.096)	(0.153)	(-0.085)	(-0.128)
5,500–6,990	0.0003	0.0001	0.0002	0.0002	0.0001	0.0003	0.0002	0.0002
	(0.435)	(0.191)	(0.240)	(0.229)	(0.111)	(0.464)	(0.270)	(0.257)
7,000–8,990	0.0007	0.0006	0.0006	0.0006	0.0005	0.0008	0.0007	0.0005
	(0.898)	(0.687)	(0.751)	(0.749)	(0.565)	(0.998)	(0.817)	(0.676)
9,000+	0.0001	0.0000	0.0001	0.0001	-0.0001	0.0002	0.0001	0.0001
	(0.157)	(0.036)	(0.121)	(0.118)	(-0.055)	(0.266)	(0.142)	(0.056)
Dummy for Age (2)								
30–34	0.0010	0.0011	0.0012	0.0013	0.0013	0.0008	0.0012	0.0013
	(1.403)	(1.636)	(1.776)	(1.784)	(1.814)	(1.141)	(1.759)	(1.804)
35–39	0.0017	0.0018	0.0019	0.0020	0.0020	0.0016	0.0019	0.0018
	(1.692)	(1.762)	(1.889)	(1.907)	(1.933)	(1.611)	(1.830)	(1.778)
40–44	0.0021	0.0021	0.0022	0.0022	0.0024	0.0020	0.0021	0.0020
	(1.787)	(1.729)	(1.809)	(1.828)	(2.001)	(1.695)	(1.722)	(1.640)
45–49	0.0029	0.0029	0.0030	0.0030	0.0031	0.0030	0.0029	0.0028
	(2.129)	(2.092)	(2.158)	(2.181)	(2.243)	(2.184)	(2.068)	(1.998)
50–54	0.0033	0.0030	0.0031	0.0031	0.0032	0.0034	0.0030	0.0029
	(2.062)	(1.862)	(1.912)	(1.938)	(1.958)	(2.120)	(1.825)	(1.819)
55–59	0.0026	0.0025	0.0025	0.0026	0.0027	0.0029	0.0023	0.0023
	(1.491)	(1.410)	(1.421)	(1.446)	(1.517)	(1.638)	(1.277)	(1.293)
60+	0.0036	0.0035	0.0035	0.0035	0.0036	0.0039	0.0032	0.0032
	(1.796)	(1.716)	(1.706)	(1.731)	(1.806)	(1.926)	(1.578)	(1.595)
Observations	371,367	371,367	371,367	371,367	371,367	371,367	371,367	371,367
R-squared	0.038	0.001	0.001	0.001	0.006	0.026	0.001	0.006
Number of monitor_code	14,442	14,442	14,442	14,442	14,442	14,442	14,442	14,442

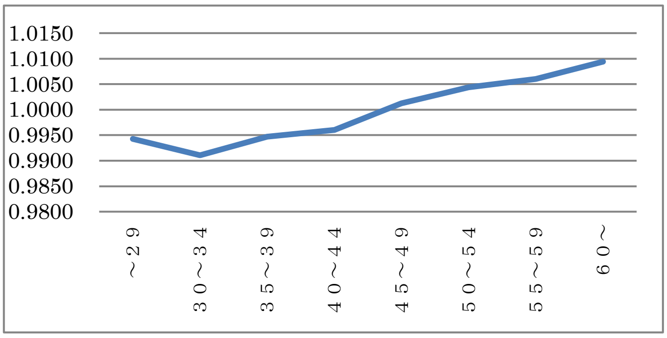
Effects of a change by One Standard Deivation on Prices

	lntrip	lnstore	lnHHI	lnquantity	non_bargain	lnstore_choice	lnquality
SD	0.73807	0.60218	0.48529	0.65390	0.16107	0.08413	0.17129
Coefficients	-0.0024	-0.0008	0.0012	-0.0098	0.0567	0.0205	0.0236
Effects on ln Prices	-0.00177	-0.00048	0.00058	-0.00641	0.00913	0.00172	0.00404

The above table shows the relationship between the price index and household characteristics as well as their shopping strategy obtained from a fixed-effects linear panel regression. Note that time effects and the number of household members are controlled and that fixed-effects estimates yield more robust results than the instrumental variable (IV) estimates employed by AH. From the table, it is evident that (1) the number of goods shopped for is negative in the second specification, but the magnitude is very small—one-tenth of that by AH and (2) the non-bargain index has relatively large effects on the household price index. Households can enjoy a decrease of 0.9% in the price level by increasing their purchasing during bargain sales by 10%.

Figure 1 depicts the relationship between the price index faced by each household and the age of wife. Evidently, price level is an increasing function of age, which is contrary to the result reported by AH based on US data.

Figure 1: Life-Cycle Profile of the Price Index



6. Conclusion

In this study, we investigated household-level price differences on the basis of Japanese homescan data. We found that the price level has a negative correlation with shopping frequency. However, the fixed-effects estimates show very small significant effects of shopping frequency on price level. The largest effects come from the non-bargain ratio.

However, there are many remaining tasks. In this study, the product-level information is not fully utilized. Further, variation in household characteristics, such as employment status and family composition, may be important in explaining the differences in inflation rates across households.

7. References

Aguiar, M. and E. Hurst (2007) "Life-Cycle Prices and Production," *American Economic Review*, 97(5), pp. 1533-1559.

Baye, M., J. Morgan, and P. Scholten (2004) "Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site," *Journal of Industrial Economics*, 52, pp. 463-496.

Goldberg, P.K. and V. Frank (2005) "Market Integration and Convergence to the Law of One Price: Evidence from the European Car Market," *Journal of International Economics*, 65(1), pp. 49-73.

Haskel, J. and H. Wolf (2001) "The Law of One Price - A Case Study," S*candinavian Journal of Economics*, 103, pp. 545-558.

Griffith, R., E. Leibtag, A. Leicester, and A. Nevo (2009) "Consumer Shopping Behavior: How Much Do Consumers Save?" *Journal of Economic Perspectives*, 23(2), pp. 99-120.

Nakamura, E. and J. Steinsson (2008) "Five Facts about Prices: A Reevaluation of Menu Cost Model," *Quarterly Journal of Economics*, 123 (4), pp. 1415-1464.