Who faces higher prices?

An empirical analysis based on Japanese homescan data[1]

Naohito Abe[2] (Hitotsubashi University) and

Kyosuke Shiotani (Bank of Japan[3])

Abstract

On the basis of Japanese household-level scanner data (called homescan), we construct a household-level price index and investigate the causes of price differences across households. Similar to the results 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 younger ones do, which contradicts Aguiar and Hurst (2007). The most important determinant of the price level is the reliance on bargain sales; doubling the proportion of purchases at bargain sales decreases the price level by about 2%, while shopping frequency has limited effects on the price level.

JEL Classification Codes: D12, D30

Keywords: Price Index, Household Consumption, Scanner data

1. Introduction

Owing to recent technological developments in data creation, numerous

1 We are grateful to comments from Andrew Leicester, Yukiko Abe, Sachiko Kuroda, Colin McKenzie, Tsutomu Watanabe, Yukinobu Kitamura, Michio Suzuki, and seminar participants at Osaka University, annual meeting of the Japanese Economic Association, and Asian Economic Policy Review Conference. This research is an outcome of the JSPS Grants-in-Aid for Young Scientist (S) 21673001.

2 Naohito Abe: The Institute of Economic Research, Hitotsubashi University. Naka, Kunitachi Tokyo. E-mail: nabe@ier.hit-u.ac.jp. Phone: +81-425-80-8347. b

3 The views expressed in this paper are those of the authors and are not reflective of those of the Bank of Japan.

commodity-price researchers 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 data are used in various economic fields, such as macroeconomics (Nakamura and Steinsson, 2008), international economics (Haskel and Wolf, 2001), and industrial economics (Baye et al., 2004). Recently, on the basis of commodity-level homescan data, Aguiar and Hurst (2007) (hereafter, AH) found a violation of the law of one price across different age groups. More precisely, in the United States, elderly families face lower prices for the same commodities than younger families do. AH interpret their results as consistent with the standard life-cycle model of consumption with endogenous decisions of shopping time. The mechanism is straightforward. Since the opportunity costs of shopping for retired people are lower than they are for younger people, the elderly can spend more time searching for lower prices. AH's findings have provided us with answers to several famous puzzles such as the retirement-savings puzzle and the excess sensitivity of expenditure to predicted income shocks.

Compared to standard consumption panel data such as the Panel Study of Income and Dynamics, homescan data by AC Nielsen and Kantar provide us with detailed and frequent information on purchases at the household-commodity level. However, most homescan data sets do not update household characteristics such as income and employment status regularly. In other words, for household data on income and employment status, homescan data lack within variation. For this reason, AH did not control for household-level fixed effects, rather, they used age and income as instrumental variables (IVs) to deal with endogeneity in the determination of shopping frequency. Because IV estimates by AH are about 20 times larger than estimates without IVs, careful examinations of the effects of households' unobservable characteristics are required.

This study considers the relationship between shopping behaviors and price levels on the basis of Japanese commodity-level homescan data. The advantage of using Japanese homescan over those in the US is that in Japanese data, household characteristics are updated every year, enabling us to make more robust estimates. It is also worth noting that many previous researchers have confirmed the existence of the retirement-savings puzzle or excess sensitivity in Japan[4]. Therefore, by investigating the relationship between shopping behaviors and price levels in Japan, we can check whether the mechanisms proposed by AH play important roles in economies outside the US.

As in the US, we find that commodities are traded at various different prices in Japan. Figure 1 illustrates the distribution of the relative commodity price index, constructed following AH.[5] The index takes a value of unity if the recorded price is equal to the regional average price. A value of 1.2 implies that the price is 20% higher than the average.[6] The figure clearly shows that the same products are sold at very different prices. We also found that the price level increases with age, in sharp contrast to AH's findings. Among several potentially important determinants of the price index, the proportion of expenditures spent on bargain sales is the most important factor. By doubling the proportion of purchases at bargain sales, people can enjoy a reduction in their price level of about 2%, which is consistent with Griffith et al. (2009), who find significant savings from purchasing at bargain sales in the United Kingdom. Other shopping behaviors, such as the frequency of shopping, the degree of mass purchasing, and a preference for high quality goods are all statistically significant. However, these

---

[4] See Wakabayashi (2008) and Ogawa (1990) for studies on Japanese consumption.

[5] The figure shows the distribution of the household-level monthly price index. The definition of the index will be given in the next section.

[6] In another strand of research on household-level heterogeneity in the price level, rather than the differences in price, but differences in weights are considered. See Kitamura (2008) and Kuroda and Yamamoto (2010) for the Japanese case.

behaviors are not quantitatively important.

Our empirical results suggest that the price-reduction mechanism based on the opportunity costs of shopping provided by AH is not observed in Japan. In contrast to the US, elderly people in Japan, who are supposed to have lower opportunity costs for shopping, tend not to use bargain sales and thus face higher prices than do the young. This suggests that further investigation into shopping strategy, particularly the determinants of purchasing at bargain sales, is necessary to understand the mechanism behind the price-level differential across families.

2. Data

The data are from the "Household Consumer Panel Research" (hereafter SCI) data set compiled by Intage, a marketing company in Japan. SCI contains the daily shopping information of approximately 12,000 households, randomly selected from all prefectures (except Okinawa) in Japan.[7] The sample households are restricted to married couples. Using a barcode reader, households are asked to scan the barcode of every commodity they purchase.[8] In SCI, for every commodity purchased, we can observe (1) the Japanese Article Number (JAN), a unique commodity identifier,[9] (2) the date of purchase, (3) the price and quantity, and (4) the name of the store from which the commodity was purchased. Fresh foods (e.g., meat, fish, and vegetables) without

[7] The sample households are selected based on three-stage stratified area sampling.
[8] Compared to the standard panel survey based on recollection, SCI requires sample households to spend more time for data creation. Therefore, one might suspect that SCI is subject to large sample selection bias. To check the potential bias, Intage uses the information of sales based on point of sales data of more than 4,000 retailers. According to Intage, the differences between the two scanner data, one from households, and one from retailers, are not large.
[9] JAN (Japanese Article Number) code is managed by The Distribution Systems Research Institute. The code is compatible with the Universal Product Code (UPC). Although the JAN code is supposedly a unique identifier, some companies use the same JAN code for different products. Intage creates its own additional code to deal with the repeated use of JAN codes. We use both JAN and Intage codes to identify commodities.

barcodes are excluded. This limitation is shared by AC Nielsen's US homescan data. The data we use in this paper cover three years, from 2004 to 2006. Table 1 shows the distribution of the family composition and comparisons with the Census and the KHPS (Keio Household Panel Survey)[10]. Compared with the Census, the sample households of SCI contain more family members. A similar bias can be found in the KHPS. Table 2 shows the age distribution of the sample wives in each surveyed year, which shows that the age distribution of SCI is close to that in the Census. Table 3 reports the employment status of the wives, which also shows that the proportion of housewives who have any paid job is close to the figure in the Census. From Tables 1-3, it is safe to say that the distribution of households in SCI is not so different from the Census or other survey data.[11]

3. Relative Price Index

Following AH, we construct the price index as follows. Let us consider a commodity that belongs to a product category $c \in C$. 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 $y_{i,t}^{j,c}$. The total expenditure by the household during time interval m can be written as,

$$X_m^j = \sum_{c \in C, i \in I_c, t \in m} p_{i,t}^{j,c} y_{i,t}^{j,c} . \tag{1}$$

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

[10] For details of KHPS, see http://www.gcoe-econbus.keio.ac.jp/english/publicdata1.html.

[11] Abe and Niizeki (2010) compared expenditures in SCI with those in the micro data of the Family Income and Expenditure Survey by the Statistical Bureau. Consistent with the previous studies in the United Kingdom (Leicester and Oldfield, 2009), although the amount of expenditure from scanner data is smaller than that in the diary base survey, the correlation between the expenditure and age or employment status are quite similar to each other.

$$\overline{X}^{j}_{m} = \sum_{c\in C, i\in I_c, t\in m} \overline{p}^{c}{}_{i,m} y^{j,c}_{i,t} \ , \tag{2}$$

where

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

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

$$\tilde{p}^{j}_{m} \equiv \frac{X^{j}_{m}}{\overline{X}^{j}_{m}} . \tag{3}$$

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

$$p^{j}_{m} \equiv \frac{\tilde{p}^{j}_{m}}{\frac{1}{J}\sum_{j} \tilde{p}^{j}_{m}} . \tag{4}$$

This household-level price index shows the price level each household faces relative to the average price level.[12] Figure 2 shows the life-cycle profile of this price index. The horizontal axis shows the age of the wife, while the vertical axis indicates the price index. As is clear from the figure, the price index increases with age; it does not decrease, as found by AH. Moreover, the slope is very small, which implies that the differences in prices across age groups are extremely limited; the absolute value of the slope is approximately one-third that estimated in the US. Figure 3 also shows the relationship between the price index and household income. Similar to Figure 2, we can observe a slightly upward line of price over income, which implies that households with

[12] When calculating the average price for each commodity, we use the regional average that divides entire Japan into 10 different regions.

greater incomes face moderately higher prices than low-income families.

Columns (1)–(2) in Table 4 shows the regression coefficients for income and age dummies when the dependent variable is the natural logarithm of the price index.[13] The effects of age and income group dummies on the price index are stable and highly significant. However, the values of the coefficients are generally not large. According to column (1) in Table 4, households whose income is over 9 million yen face 0.013-point higher prices than do the lowest income group.

It is worth noting that this price index cannot capture the movements of prices over time because the average of the price index is always unity.

4. Shopping Behaviors

One of AH's main results is that elderly people can lower their prices by increasing how frequently they shop. In this section, in addition to the shopping frequency, we introduce six other shopping behaviors that might affect the relative price index introduced in the previous section.

As the measure of shopping frequency, we use the number of stores households use, (ln_trip). More precisely, we first count the number of different stores that a sample household visits each day. Next, we calculate the sum of the number for each month,

[13] In their regression analyses, AH controlled for "shopping needs" because a shopper who spreads shopping time over numerous goods have less time to find the best bargain price. Following AH, in subsequent regression analyses we include the natural logarithms of the number of commodities, the number of product categories, and the total expenditure per month as the "shopping needs."

which gives the index for shopping frequency.

Ln_store captures the variety of shops each household patronizes. Note that this variable does not include information regarding frequent shopping at the same store. This variable can be used as a proxy for search intensity, which might lead to a lower price index, to find the people who use some stores in search of better prices.

Next, we construct the Herfindahl–Hirschman Index (HHI) to capture the concentration of spending. The HHI is a measure of the amount of competition in an industry. We use it as an indicator of the degree of concentration of stores where the households purchase goods. For example, consider two households. Both families go to three stores a month. One of the families relies on a large supermarket and spends 90 % of the monthly expenditure at the supermarket, while the other family spends evenly across the three stores. Our HHI captures the difference in such shopping behaviors. The HHI is defined as

$$HHI_m^j \equiv \sum_{k=1}^{K} {S_{k,m}^j}^2, \tag{5}$$

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

Next, we consider the monthly total number of goods a household buys:

$$Quantity_m^j = \sum_{c \in C, i \in I_c, t \in m} y_{i,t}^{j,c}. \tag{6}$$

It is reasonable to suppose that a family buying many goods can enjoy volume discounts more, thus decreasing the price level.

To observe the effect of buying at bargain sales, we construct a measure for bargains. As one might expect, a household can decrease its price index by purchasing more goods at bargain sales. Because of the lack of store-level flags for bargain sales in our dataset, it is necessary to define the price at bargain sales based on information regarding the movements of store-level prices. In this paper, we adopt the store-level monthly minimum price for each good, $min\ P_{i,t}^{c}$, as the price at bargain sales. Then, the following index is used,

$$bargain_{m}^{j} = \frac{\sum_{c\in C, i\in I_c, t\in m} I\left(P_{i,t}^{j,c}\right) p_{i,t}^{j,c} y_{i,t}^{j,c}}{\sum_{c\in C, i\in I_c, t\in m} p_{i,t}^{j,c} y_{i,t}^{j,c}}, \tag{7}$$

where

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

shows the proportion of expenditures at bargain prices. A household with a large bargain index is purchasing products at lower-than-normal prices, which lowers the relative price index. It is worth noting that this measure captures the importance of temporal reduction within a month. If prices are stable for several months, or if bargain sales last more than one month, this index fails to capture the importance of bargain sales[14].

Generally, most products can be purchased at both luxury stores and discount stores. The movement of prices differs across stores to a great extent. Abe and Tonogi (2009) show that prices move very differently across stores based on a large point-of-sale database of Japanese stores. Suppose a high-income family has greater opportunity costs for shopping than do low-income families. Also, suppose that a high-income family tends to use luxury stores. Then it is probable that luxury stores sell commodities at higher prices than standard supermarkets because customers can reduce their shopping costs by buying goods at one shop even if they know other stores have set lower prices for exactly the same goods. However, discount shops cannot set higher prices for

[14] When a household purchases very rare items that are sold only once in a given month and store, we cannot identify whether this price is a bargain price or a regular price. Our variable regards this as a regular price. To check the importance of the rarely traded goods, we try several different definitions of bargain sales and find that our main results in later sections do not depend on rarely traded goods.

common goods because common goods are their main products, so these shops expect customers to change their favorite shops if one shop increases its prices for commonly used goods. Thus, it is worth examining the effects of the quality of stores on the price index. We define the index for the quality of each store, $k \in \mathrm{K}$, by following the same basic procedure as we did for the relative price index. The store quality index is the ratio of the hypothetical sales if the store sells the goods at their average price $\bar{P}_{i,m}^{c}$ to the sales if the store sells the goods at their categorical average price. More precisely, we first obtain the average price for a given good in category $c \in C$:

$$\bar{P}_{m}^{c} = \sum_{i \in I_c, k \in K, t \in m} p_{i,t}^{k,c} \frac{y_{i,t}^{k,c}}{\sum_{i \in I_c, k \in K, t \in m} y_{i,t}^{k,c}}. \tag{8}$$

Next, assuming that the stores sell the average goods in each category at the average price, we obtain the total sales:

$$\bar{Z}_{m}^{k} = \sum_{c \in C, i \in I_c, t \in m} \bar{p}_{m}^{c} y_{i,t}^{k,c}. \tag{9}$$

Then, we calculate the total sales of store k if it sells the goods at their average prices

$$\bar{p}_{i,m}^{c} = \sum_{k \in K, t \in m} p_{i,t}^{k,c} \frac{y_{i,t}^{k,c}}{\sum_{k \in K, t \in m} y_{i,t}^{k,c}}, \tag{10}$$

$$Z_{m}^{k} = \sum_{c \in C, i \in I_c, t \in m} \bar{p}_{i,m}^{c} y_{i,t}^{k,c}. \tag{11}$$

Now, the index for the quality of goods sold at store k is defined:

$$\tilde{q}_{m}^{k} \equiv \frac{Z_{m}^{k}}{\bar{Z}_{m}^{k}} . \tag{12}$$

Finally, we normalize the index by dividing by the average monthly quality index,

$$q_{m}^{\mathrm{k}} \equiv \frac{\tilde{q}_{m}^{k}}{\sum_{k \in K} \tilde{q}_{m}^{k}} , \tag{13}$$

which gives us the quality index of a store k during the time interval m.

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

$$Store\ \mathrm{choice}_m^j \equiv \sum_{k \in K} S_{k,m}^j q_m^k \mathrm{T}. \quad (14)$$

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

By changing stores to households in the previous index, we can create a household-level monthly average quality index. The quality index for households is defined as the ratio of the hypothetical expenditure if the household purchases the goods at their average price, $\bar{p}_{i,m}^c$, to the expenditure if the household purchases the goods at their categorical average price, $\bar{p}_m^c$. Formerly, define the total expenditure by household j when assuming the household purchase goods at the categorical average price:

$$\bar{Z}_m^j = \sum_{c \in C, i \in I_c, t \in m} \bar{p}_m^c y_{i,t}^{j,c}. \quad (15)$$

Next, define the hypothetical expenditure if the household purchases the goods at their commodity level average price:

$$Z_m^j = \sum_{c \in C, i \in I_c, t \in m} \bar{p}_{i,m}^c y_{i,t}^{j,c}, \quad (16)$$

The index for quality of goods bought by household *j* is defined as follows:

$$\tilde{q}_m^j \equiv \frac{Z_m^j}{\bar{Z}_m^j}. \quad (17)$$

We normalize the index by dividing by the average monthly quality index as follows:

$$q_m^j \equiv \frac{\tilde{q}_m^j}{\sum_{j \in J} \tilde{q}_m^j}. \quad (18)$$

It is expected that the greater this quality index, the higher the price index will be.

As noted previously, this measure is not affected by other shopping strategies of each household, such as buying at sales, because it uses the average price of each good. In this index, we assume all households encounter the same prices for specific goods, so a higher index value does not imply that a household buys goods at higher prices than another household does.

Table 5 reports the descriptive statistics of these shopping behavior variables and the relative price index across different age and income groups. On average, Japanese families shop 14.4 times a month. The standard deviation of the number of trips is large, at 9.5, which implies that families are highly heterogeneous in their shopping frequency. Figure 4 confirms the heterogeneity. Some families shop more than 100 times a month. It is important to note that this index counts multiple trips to the same store within the same day as only one trip, so the number of shopping trips in this table is the lower bound of the actual number of trips.

According to Table 5, households with a wife of 50–54 years old shop more frequently than do younger households, which is consistent with the results found by AH. We can also observe that the shopping frequency increases with income. Not surprisingly, the proportion of bargain purchases decreases with age and income. The standard deviation is also large. Figure 5 shows the distribution of the proportion of expenditures spent on bargain sales. We can observe a mass point at zero, which implies that many families always purchase goods at higher prices rather than at the monthly minimum price.[15] The shopping concentration measure, HHI, decreases with age and income, implying that elderly and high-income families tend to disperse their expenditures across different stores.

---

[15] The monthly minimum price is defined as the commodity-store level minimum price each month.

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

Columns (3)‐(5) in Table 4 report the results of ordinary least squares. Because of the endogeneity in shopping behaviors, we should be careful in our interpretations of the coefficients of shopping behaviors, such as the frequency of trips. Because of the large sample size, some of the t-values exceed fifty. Except for the Hafindahl–‐Hirschman Index (ln_HHI) and the number of different stores (ln_store), the sign of the shopping behaviors are generally consistent with the casual hypotheses raised in the previous section. For example, the coefficient of the frequency of trips (ln_trip) is negative, which implies that households that shop often face lower prices. Moreover, the size of the coefficient, -0.0135 in column (3), is similar to the same result in the ordinary least squares (OLS) regression performed by AH.

AH used dummies for income and age as instrumental variables to control for the endogeneity of shopping behavior. Unfortunately, in our dataset, the two-stage least squares estimates with these instrumental variables are quite unstable and cannot pass the over-identification tests. Thus, rather than relying on instrumental variables, we adopt a fixed-effects model, which enables us to omit the biases due to unobservable family-level effects.

Columns (6)–(8) in Table 4 show the estimation results. Robust and stable relationships between shopping behaviors and the relative price index can be found in ln_quantity, bargain, ln_store_choice, and ln_quality. The effects of age and income become much smaller than those reported in columns (1)‐(5), probably because the fixed effects absorb the age effect. The effect of shopping frequency becomes positive significant. Although the positive effect is difficult to interpret, the magnitude of the effect is negligible. The only non-negligible effect comes from bargain sales. The average proportion of purchases at bargain sales is about 14%. Thus, by doubling this proportion, households can enjoy about a 2% reduction in the price level. It is worth noting that the R-squared in column (3) of Table 4 is approximately 20%, which implies that approximately 80% of the differences in the relative price index cannot be explained by the observed variables. As shown in Figure 1 and Table 4, there is a notable amount of heterogeneity in the relative price index across households. We need more information on the households' shopping behavior and preferences to study the cause of this heterogeneity in more detail.

6. Why do the Elderly Face Higher Prices?

Table 4 shows a very stable and robust effect of age on the price level. The positive significant effects of the dummies on elderly households remain even after controlling for various shopping strategies. Thus, we have to seek different mechanisms behind the age effect. In this section, we consider an alternative shopping strategy whereby younger households may shop at discount stores more than older households, which might explain the price level differentials between age groups. Table 6 reports the share of total expenditures purchased at different types of store by household age. We observe that elderly households spend more at specialized stores and supermarkets, and less at drug stores and home-improvement stores than young households do. Table 7 compares the OLS estimates with and without store choice variables. Although the age dummies remain positive and significant in columns (3)-(4), the dummy for 60 years or older becomes smaller by more than 10%. Therefore, the store choice can partially explain price-level differentials between households. However, Table 8 also shows that the elderly face higher prices than the young do even within the same type of store, which might appear in columns (3)–(4) in Table 7. Thus, the mechanism behind the higher prices paid by the elderly needs further investigation. This could possibly be explained by the existence of unobserved variables that correlate with age. Financial assets are one such example. Since the homescan data we use do not contain such information, we need to seek different data sets, a task for future research.

7. Policy Implications

Consumer price indices are some of the most important economic indicators for monetary policy. When the inflation rate is close to zero, a slight change in the rate, perhaps as small as a tenth of a percent, attract strong attention among policy makers and investors. However, it is well known that price indices are subject to various types of measurement error (see Boskin et al., 1996, Shiratsuka 1999, Lebow and Rudd 2003). Our research shows that the law of one price, which the standard consumer price index theory assumes, is violated. Price differentials across households are large and correlated with household characteristics such as income and age, and with their shopping strategies. Abe and Tonogi (2010) observed that price change rates and bargain ratios are extremely heterogeneous across stores, even for the same commodity. These results suggest that inflation can have different effects on different types of household. It is likely that the welfare loss caused by inflation for households that can find better bargain prices or discount shops is smaller than the welfare loss for households that have limited choices to seek lower prices. According to our results, it is probable that inflationary monetary policy creates a bigger loss for elderly and high-income people. When artificially changing the inflation rate, policy makers should consider that such polices have effects not only on the average of household welfare, but also on its distribution.

8. Conclusion

This study used Japanese scanner data to investigate household-level price differences. The data reveal that the law of one price is violated to a great extent; differences in prices across households exist for the same commodity. These results are consistent with previous studies based on US data. However, the price level is negatively correlated with shopping frequency, while it is positively correlated with income and age, which is inconsistent with US results. The proportion of purchases at bargain sales is a declining function of age. After controlling for purchases at bargain sales, the age effects on the price level become very small, suggesting that elderly households face higher prices than young households do because elderly people use less bargain sales. The fixed effects estimates show very small significant effects of the shopping frequency on the price level, which is inconsistent with previous studies based on US data.

Many tasks remain to be completed in this line of research. In this study, the product-level information is not fully utilized. The variation in household characteristics, such as employment status and family composition, may also be important in explaining the differences in prices across households. Finally, following Broda and Romalis (2009), the heterogeneity in the movements of the price level, that is, the heterogeneity in household-level inflation, needs to be investigated.

References

Abe, N. and T. Niizeiki (2010) "Consumption Data on Homescan: Comparisons with Other Consumption data," (in Japanese) *The Economic Review*, 61(3), pp. 224-236.

Abe, N., and A. Tonogi (2010) "Micro and Macro Price Dynamics in Daily Data," *Journal of Monetary Economics*, 57, pp. 716-728.

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.

Boskin, Michael J., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson (1996) "Toward a More Accurate Measure of the Cost of Living," Final Report to the Senate Finance Committee.

Broda, C. and J. Romalis (2009) "The Welfare Implications of Rising Price Dispersion," mimeo.

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.

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.

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

Kitamura, Y. (2008) "Kakeibetsu Bukkashisu no Kouchiku to Bunseki," (in Japanese) *Kinyukennkyux*, 27(3), pp. 91-150.

Kuroda, S., and I. Yamanoto (2010) "Kakeibetsu Infle Ritsu No Bonpu To Sono Jizokusei," (in Japanese) in Seko et al. ed., *Nihon Kakei Kodo no Dainamizumu*, Keio University Press, pp. 217-244.

Lebow, D. E. and J. B. Rudd (2003) "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature*, 41(1), pp. 159–201.

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.

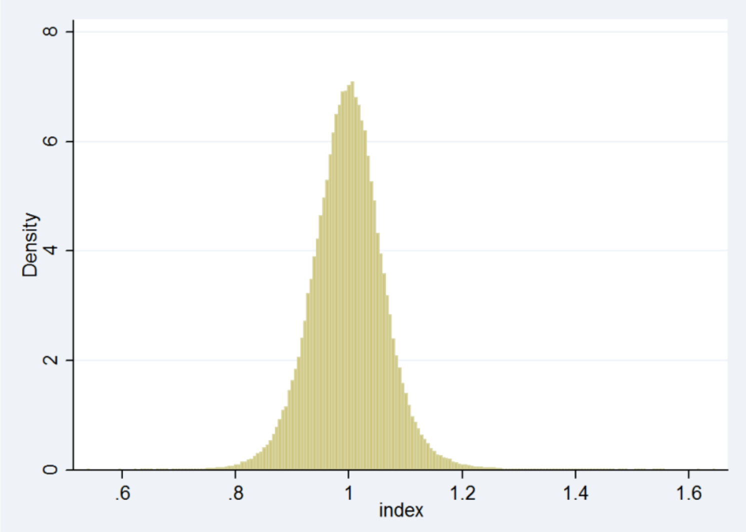
Ogawa, K. (1990) "Cyclical Variations in Liquidity Constrained Consumers: Evidence from Macrodata in Japan," *Journal of Japanese and International Economies*, 4, pp. 173-193.

Shiratsuka, Shigenori (1999) "Measurement Errors in the Japanese Consumer Price Index," *Monetary and Economic Studies*, 17(3), pp. 69-102.

Unayama, T. and M. Keida, (2011) "Koureisha-Setai no Shouhi Koudou To Bukka Shisuu," (in Japanese) RIETI Discussion Paper Series 11-J-047/.

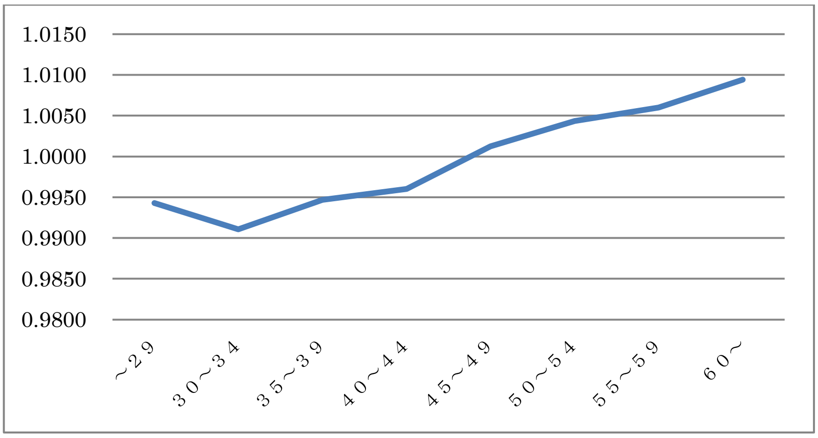
Wakabayashi, M. (2008) "The Retirement Consumption Puzzle in Japan," *Journal of Population Economics*, 21, pp. 983-1005.

Figure 1: Distribution of the relative price index across households



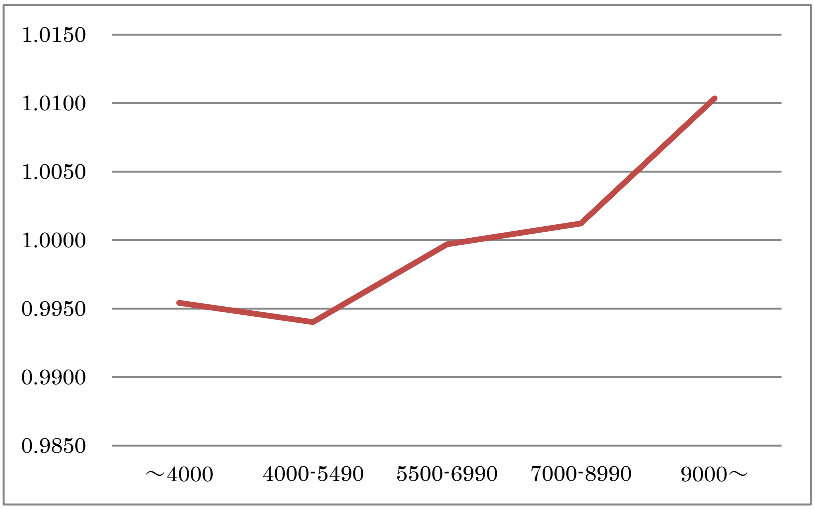
Note: The definition of the price index is given in Section 3.

Figure 2: Life-Cycle Profile of the Price Index



Note: The horizontal axis is the age of the wife in the household.

Figure 3: Household Income and Price Index



Note: The horizontal axis is household income in units of 1,000 yen.

Figure 4: Distribution of the Frequency of Shopping per Month

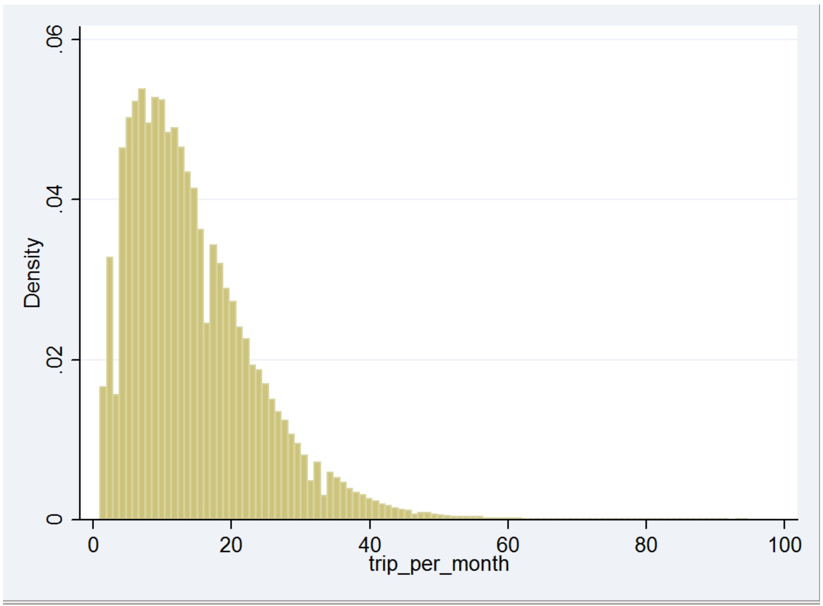


Figure 5: Distribution of the Proportion of Expenditures on Bargain Sales

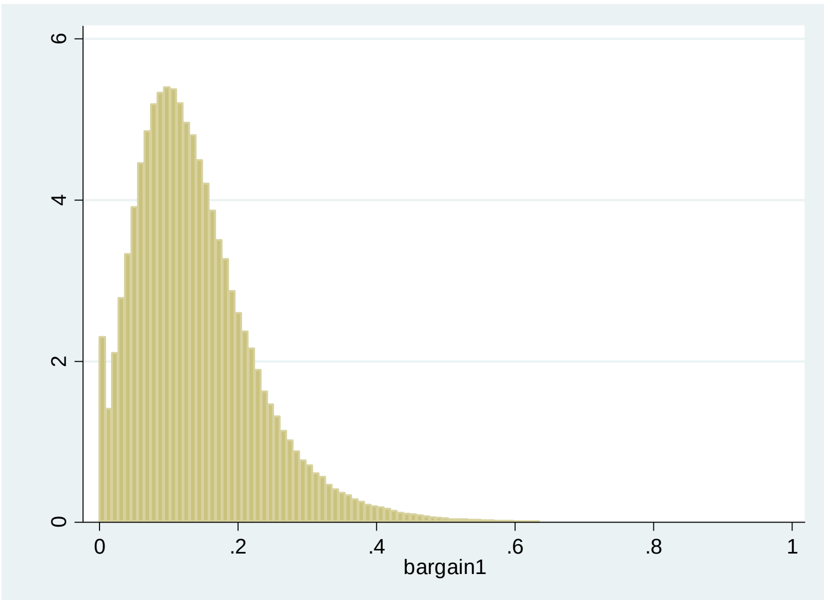


Table 1: Family Composition of SCI and Other Data

Family members			2	3	4	5	6
	2004		0.16	0.24	0.38	0.15	0.07
year	2005	SCI	0.16	0.24	0.38	0.15	0.07
	2006		0.17	0.24	0.38	0.14	0.07
	2005	Census	0.38	0.27	0.22	0.08	0.05
	2004-2009	KHPS	0.22	0.24	0.29	0.14	0.11

Note: SCI is homescan data by Intage. KHPS stands for Keio Household Panel Survey.

Table 2: Distribution of Wife's Age

Wife Age			～29	30～34	35～39	40～44
	2004		0.08	0.11	0.16	0.16
year	2005	SCI	0.08	0.12	0.14	0.16
	2006		0.08	0.12	0.14	0.15
	2005	Census	0.068	0.107	0.111	0.11
Wife Age			45～49	50～54	55～59	60～
	2004		0.12	0.14	0.11	0.12
year	2005	SCI	0.12	0.13	0.13	0.11
	2006		0.12	0.12	0.14	0.11
	2005	Census	0.111	0.128	0.148	0.218

Table 3: The Relation between Wife's Age and Job Status

Age	SCI						Census 2005
	Full Time	Part Time	Self Employed	Agriculture	Sideline	Non Working	Non Working
～29	0.13	0.23	0.01	0.00	0.02	0.61	0.55
30～34	0.10	0.33	0.02	0.00	0.04	0.52	0.54
35～39	0.14	0.41	0.02	0.00	0.03	0.40	0.47
40～44	0.14	0.50	0.03	0.00	0.04	0.29	0.35
45～49	0.20	0.47	0.04	0.00	0.03	0.27	0.30
50～54	0.19	0.44	0.04	0.00	0.02	0.31	0.34
55～59	0.18	0.33	0.04	0.00	0.02	0.42	0.41
60～	0.09	0.19	0.06	0.00	0.02	0.64	0.50

Note: The figures are average for the sample period: 2004-2006.

Table 4: The Relation between Price and Shopping Behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regression Type	OLS	OLS	OLS	OLS	OLS	FE	FE	FE
ln_trip			−0.0095	−0.0135		0.0042	0.0036	
			(−38.224)	(−61.168)		(9.590)	(9.533)	
ln_store			−0.0017			0.0006		
			(−5.582)			(1.647)		
ln_HHI			0.0035			0.004		
			(11.712)			(10.571)		
ln_quantity			−0.0019			0.0003		
			(−3.254)			(0.257)		
bargain			−0.2401		−0.2489	−0.1288		−0.1311
			(−209.735)		(−217.255)	(−102.358)		(−103.887)
ln_store_choice			0.0305			0.021		
			−(23.159)			(10.046)		
ln_quality			0.0388			0.0264		
			(54.630)			(27.956)		
Dummy for Income (1)								
4,000−5,490	0.0019	0.0016	0.0007	0.0016	0.0015	0.0004	0.0001	0.0003
	(6.503)	(5.262)	(2.472)	(5.368)	(5.421)	(0.403)	(0.149)	(0.294)
5,500−6,990	0.0059	0.0059	0.0029	0.0051	0.0044	0.0006	0.0004	0.0004
	(18.546)	(18.676)	(10.086)	(16.128)	(14.845)	(0.521)	(0.318)	(0.378)
7,000−8,990	.0073	.0072	0.0039	0.0065	0.0057	0.001	0.0008	0.0009
	(22.234)	(22.389)	(12.938)	(19.650)	(18.561)	(0.825)	(0.602)	(0.756)
9,000−	0.0132	0.0133	0.0076	0.0121	0.0098	0.0006	0.0003	0.0005
	(38.275)	(40.347)	(23.881)	(34.824)	(30.196)	(0.423)	(0.197)	(0.354)
Dummy for Age (2)								
30−34	−0.0015	−0.0021	−0.0015	−0.0015	−0.0012	0.0012	0.0014	0.0011
	(−3.509)	(−4.859)	(−3.820)	(−3.492)	(−2.980)	(1.027)	(1.151)	(0.959)
35−39	0.0018	0.0020	0.0014	0.0019	0.0017	0.0016	0.0023	0.0015
	(3.931)	(4.833)	(3.356)	(4.267)	(4.082)	(0.976)	(1.353)	(0.944)
40−44	0.0013	0.0036	0.0012	0.0017	0.0015	0.002	0.0027	0.002
	(2.721)	(8.522)	(2.590)	(3.372)	(3.211)	(1.084)	(1.370)	(1.097)
45−49	0.0035	0.0078	0.003	0.0039	0.0034	0.0025	0.0034	0.0025
	(6.571)	(17.705)	(6.100)	(7.229)	(6.952)	(1.176)	(1.554)	(1.174)
50−54	0.0039	0.0085	0.0047	0.0052	0.004	0.0025	0.0034	0.0024
	(7.051)	(19.368)	(9.366)	(9.576)	(8.004)	(1.029)	(1.337)	(1.003)
55−59	0.0056	0.0101	0.0073	0.0075	0.0062	0.0019	0.0029	0.002
	(10.121)	(23.035)	(14.360)	(13.660)	(12.157)	(0.733)	(1.043)	(0.767)
60−	0.0107	0.0152	0.012	0.0131	0.0109	0.0028	0.0038	0.003
	(19.154)	(33.644)	(23.030)	(23.197)	(20.894)	(0.927)	(1.206)	(0.977)
Constant	0.0326	0.0342	0.0238	−0.0076	0.0471	0.119	0.1118	0.1117
	(16.107)	(18.527)	(6.135)	(−3.524)	(24.422)	(19.655)	(26.536)	(28.340)
Household Characteristics	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Needs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Dummies	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	371,367	371,367	371,367	371,367	371,367	371,367	371,367	371,367
R−squared	0.048	0.035	0.215	0.059	0.19	0.084	0.008	0.075
Number of monitor_code	14,442	14,442	14,442	14,442	14,442	14,442	14,442	14,442

Note:

Ordinary least squares and fixed effects estimations based on Japanese homescan provided by Intage.

The dependent variable is the Household−Level Price Index

Clustering t−statistics are in parentheses.

Household characteristics include dummy variables for the number of family members and the number of

Location Dummies include dummy variables for city size dummies and prefecture dummies

Needs include the natural logarithms of the number of commodities, the number of product categories, and the total expenditure per month.

The data is converted to household−level monthly data.

(1) The unit is 1000yen. The base is the income below 4,000.

(2) The age of the wife in the household. The base is the dummy for below 30.

Table 5: Descriptive Statistics of Shopping Behaviors

		lnprice		Price Index (Level)		ln_trip		Number of Trips		Number of Stores (ln)	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
age of wife	～29	−0.0075	0.0595	0.9943	0.0590	2.0662	0.7512	10.1810	7.3054	1.2776	0.6075
	30～34	−0.0106	0.0572	0.9911	0.0566	2.1968	0.7486	11.4913	7.7730	1.3448	0.6102
	35～39	−0.0070	0.0572	0.9947	0.0567	2.3254	0.7387	12.9688	8.5863	1.4037	0.5978
	40～44	−0.0056	0.0571	0.9960	0.0567	2.4589	0.7505	14.8864	9.8937	1.4556	0.6031
	45～49	−0.0003	0.0564	1.0013	0.0563	2.5284	0.7340	15.7767	10.1449	1.4748	0.5982
	50～54	0.0027	0.0571	1.0044	0.0571	2.6112	0.6869	16.7191	10.3619	1.5572	0.5727
	55～59	0.0044	0.0573	1.0060	0.0573	2.5979	0.6526	16.1623	9.4398	1.5627	0.5759
	60～	0.0076	0.0584	1.0094	0.0587	2.5977	0.6516	16.1627	9.4465	1.5291	0.5787
	Total	−0.0021	0.0577	0.9995	0.0575	2.4349	0.7381	14.4381	9.5046	1.4565	0.5997
income	～4000	−0.0062	0.0594	0.9954	0.0669	2.3426	0.7376	13.1376	8.8552	1.3692	0.6015
	4000−5490	−0.0070	0.0577	0.9940	0.0638	2.3895	0.7294	13.6798	8.9383	1.4247	0.5914
	5500−6990	−0.0026	0.0573	0.9997	0.0636	2.4217	0.7432	14.2156	9.4528	1.4552	0.5980
	7000−8990	−0.0007	0.0567	1.0012	0.0619	2.4921	0.7453	15.2501	10.1186	1.4998	0.5989
	9000～	0.0073	0.0563	1.0104	0.0625	2.5430	0.7184	15.7681	10.0189	1.5428	0.5963
	Total	−0.0021	0.0577	0.9998	0.0640	2.4349	0.7381	14.3654	9.5090	1.4565	0.5997

		Bargain		ln_quantity		ln_HHI		ln_store_choice		ln_quality	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
age of wife (1)	～29	0.1394	0.0967	4.0852	0.6789	8.4072	0.4752	−0.1816	0.0753	−0.0518	0.1751
	30～34	0.1433	0.0924	4.2796	0.6544	8.3691	0.4848	−0.1784	0.0739	−0.0355	0.1753
	35～39	0.1433	0.0859	4.4645	0.6417	8.3649	0.4791	−0.1748	0.0768	−0.0244	0.1605
	40～44	0.1441	0.0830	4.6017	0.6465	8.3466	0.4837	−0.1714	0.0730	−0.0157	0.1548
	45～49	0.1402	0.0825	4.6544	0.6393	8.3459	0.4851	−0.1664	0.0779	−0.0008	0.1599
	50～54	0.1371	0.0861	4.6235	0.6232	8.2837	0.4829	−0.1584	0.0841	0.0005	0.1749
	55～59	0.1364	0.0908	4.5536	0.5903	8.2729	0.4854	−0.1541	0.0917	0.0016	0.1854
	60～	0.1364	0.0937	4.5348	0.5836	8.2979	0.4912	−0.1501	0.1133	−0.0034	0.1842
	Total	0.1403	0.0884	4.4945	0.6539	8.3344	0.4853	−0.1668	0.0841	−0.0150	0.1713
income (2)	～4000	0.1436	0.0951	4.3503	0.6523	8.3812	0.4816	−0.1745	0.0839	−0.0494	0.1730
	4000−5490	0.1443	0.0916	4.4441	0.6405	8.3507	0.4774	−0.1729	0.0808	−0.0305	0.1685
	5500−6990	0.1407	0.0876	4.5026	0.6582	8.3399	0.4824	−0.1682	0.0766	−0.0130	0.1640
	7000−8990	0.1412	0.0879	4.5772	0.6418	8.3129	0.4878	−0.1618	0.0903	0.0006	0.1670
	9000～	0.1333	0.0857	4.6117	0.6462	8.2818	0.4938	−0.1548	0.0883	0.0218	0.1762
	Total	0.1403	0.0884	4.4945	0.6539	8.3344	0.4853	−0.1668	0.0841	−0.0150	0.1713

Note

(1) The age of the wife in the household. The base is the dummy for below 30.

(2) The unit is 1000yen. The base is the income below 4,000 thousand yen.

Table 6: The Expenditure Shares in Different Store Types and Age

		Convenience store		Specialized store		Pharmacy		Home Improvement Store		Home delivery & Door to door sales	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
age of wife	～29	0.0107	0.0016	0.0666	0.0085	0.1778	0.0098	0.0619	0.0053	0.0487	0.0046
	30～34	0.0088	0.0014	0.0657	0.0045	0.1448	0.0080	0.0618	0.0042	0.0678	0.0053
	35～39	0.0089	0.0010	0.0526	0.0035	0.1116	0.0074	0.0568	0.0042	0.0780	0.0049
	40～44	0.0101	0.0015	0.0522	0.0038	0.1005	0.0097	0.0501	0.0040	0.0843	0.0066
	45～49	0.0078	0.0009	0.0598	0.0079	0.0934	0.0073	0.0560	0.0029	0.0843	0.0051
	50～54	0.0106	0.0016	0.0785	0.0086	0.0922	0.0069	0.0593	0.0037	0.0798	0.0043
	55～59	0.0102	0.0012	0.0920	0.0104	0.0848	0.0064	0.0547	0.0034	0.0750	0.0046
	60～	0.0092	0.0019	0.1061	0.0123	0.0790	0.0036	0.0453	0.0029	0.0594	0.0045
	Total	0.0096	0.0017	0.0702	0.0196	0.1111	0.0300	0.0558	0.0061	0.0765	0.0107

		Supermarket		Others	
		mean	sd	mean	sd
age of wife	～29	0.6254	0.0125	0.0115	0.0019
	30～34	0.6420	0.0088	0.0109	0.0012
	35～39	0.6860	0.0085	0.0075	0.0007
	40～44	0.6968	0.0096	0.0081	0.0008
	45～49	0.6926	0.0097	0.0083	0.0015
	50～54	0.6717	0.0084	0.0100	0.0011
	55～59	0.6718	0.0065	0.0133	0.0013
	60～	0.6816	0.0108	0.0216	0.0022
	Total	0.6766	0.0224	0.0119	0.0050

Table 7: The Relative Price by Store Type and Age

		Convenience store		Specialized store		Pharmacy		Home improvement store	
		price	share	price	share	price	share	price	share
age of wife	～29	109.9671	0.0107	99.1886	0.0666	99.2837	0.1778	97.8523	0.0619
	30～34	109.0378	0.0088	98.8977	0.0657	98.8687	0.1448	97.5777	0.0618
	35～39	108.7232	0.0089	99.2706	0.0526	98.8541	0.1116	97.5827	0.0568
	40～44	108.5826	0.0101	99.5857	0.0522	98.6101	0.1005	97.8275	0.0501
	45～49	107.8898	0.0078	99.6815	0.0598	99.1955	0.0934	98.3455	0.0560
	50～54	109.2453	0.0106	99.8444	0.0785	99.1625	0.0922	98.8271	0.0593
	55～59	108.4516	0.0102	100.6701	0.0920	99.4394	0.0848	98.8922	0.0547
	60～	107.8993	0.0092	100.8410	0.1061	99.8501	0.0790	99.2877	0.0453
	Total	108.6462	0.0096	99.9079	0.0702	99.0840	0.1111	98.2684	0.0558

		Home delivery & Door to door sales		Supermarket		Others	
		price	share	price	share	price	share
age of wife	～29	100.3027	0.0487	99.6685	0.6254	99.5327	0.0115
	30～34	100.4226	0.0678	99.3710	0.6420	99.8055	0.0109
	35～39	100.4065	0.0780	97.2600	0.6860	99.8625	0.0075
	40～44	100.4145	0.0843	99.8271	0.6968	99.5230	0.0081
	45～49	100.5679	0.0843	100.1899	0.6926	99.9938	0.0083
	50～54	100.8512	0.0798	100.3565	0.6717	101.1497	0.0100
	55～59	100.9304	0.0750	100.5988	0.6718	100.7426	0.0133
	60～	100.9685	0.0594	100.8945	0.6816	101.7224	0.0216
	Total	100.6175	0.0765	100.1078	0.6766	100.5027	0.0119

Table 8: The Effects of Store Choice

	(1)	(2)	(3)	(4)
Dummy for Income (1)				
4,000–5,490	0.0019	0.0016	0.0018	0.0014
	(6.503)	(5.262)	(6.123)	(4.833)
5,500–6,990	0.0059	0.0059	0.0056	0.0056
	(18.546)	(18.676)	(17.902)	(17.813)
7,000–8,990	.0073	.0072	.0069	.0068
	(22.234)	(22.389)	(21.141)	(21.320)
9,000–	0.0132	0.0133	0.0127	0.0129
	(38.275)	(40.347)	(37.060)	(39.190)
Dummy for Age (2)				
30–34	−0.0015	−0.0021	−0.0018	−0.0025
	(−3.509)	(−4.859)	(−4.041)	(−5.820)
35–39	0.0018	0.0020	0.0012	0.0011
	(3.931)	(4.833)	(2.582)	(2.655)
40–44	0.0013	0.0036	0.0004	0.0023
	(2.721)	(8.522)	(0.786)	(5.536)
45–49	0.0035	0.0078	0.0028	0.0068
	(6.571)	(17.705)	(5.197)	(15.412)
50–54	0.0039	0.0085	0.0027	0.0071
	(7.051)	(19.368)	(4.965)	(16.078)
55–59	0.0056	0.0101	0.0044	0.0086
	(10.121)	(23.035)	(8.006)	(19.570)
60–	0.0107	0.0152	0.0096	0.0135
	(19.154)	(33.644)	(17.013)	(29.676)
Constant	0.0326	0.0342	0.0475	0.0529
	(16.107)	(18.527)	(18.129)	(21.406)
Household Characteristics	Yes	No	Yes	No
Needs	Yes	Yes	Yes	Yes
Location Dummies	Yes	No	Yes	No
Time Dummies	Yes	No	Yes	No
Store Choice	No	No	Yes	Yes
Observations	371,367	371,367	371,367	371,367
R-squared	0.048	0.035	0.064	0.052
Number of monitor_code	14,442	14,442	14,442	14,442

Note:
Ordinary least squares estimates based on Japanese homescan provided by Intage.
The dependent variable is the Household-Level Price Ind.

Clustering t-statistics are in parentheses.

Household characteristics include dummy variables for the number of family members and the number of children.

Needs include the natural logarithms of the number of commodities, the number of product categories, and the total expenditure per month.

Location Dummies include dummy variables for city size dummies and prefecture dummies.
Store Choice is the expenditure share by store type.
The data is converted to household-level monthly data.
(1) The unit is 1000 yen. The base is the income below 4,000.
(2) The age of the wife in the household. The base is the dummy for below 30.