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

October 2013

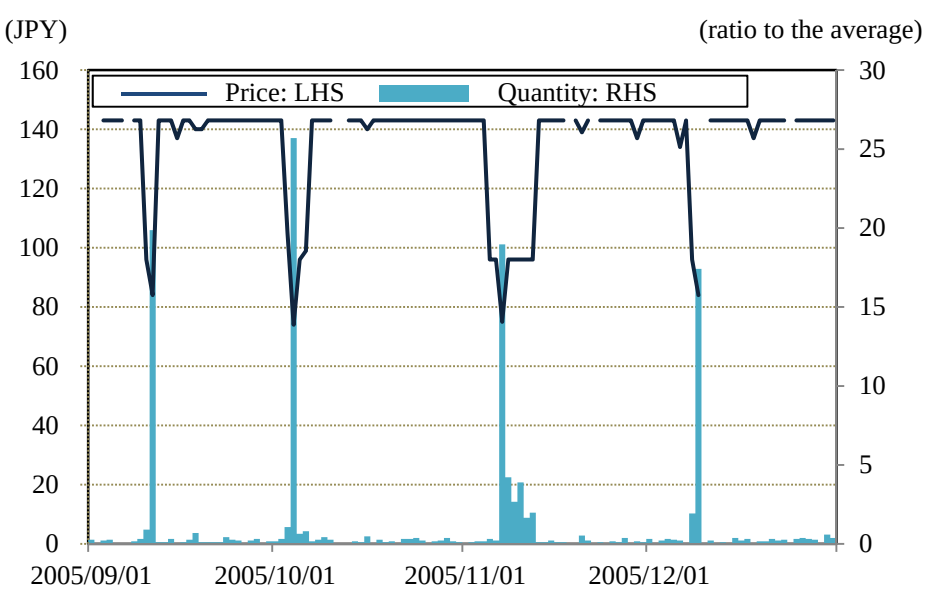
Naohito Abe (Hitotsubashi University) &
Kyosuke Shiotani (Bank of Japan*)

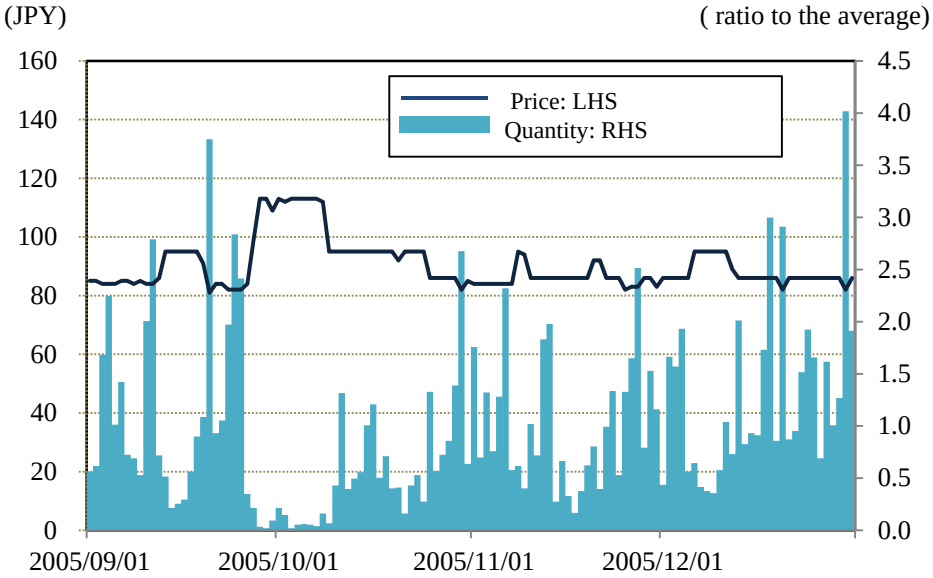
*The views expressed in this paper are those of the authors and are not reflective of those of the Bank of Japan.

Are prices heterogeneous?

- In many occasions, we assume the law of one price holds: the same commodities are traded at the same price.
- For example, in many consumption or firm analyses, we use the official CPI as the unique CPI to deflate nominal income, expenditure, and sales.
- Some researchers introduce heterogeneity by using different weights: Kitamura (2008), Kuroda and Yamamoto (2010) , Unayama and Keida (2011).
- But they do not consider heterogeneity in prices across the same commodities. They assume “The law of one price”.

The Difference of Price Movements within the Same Commodities

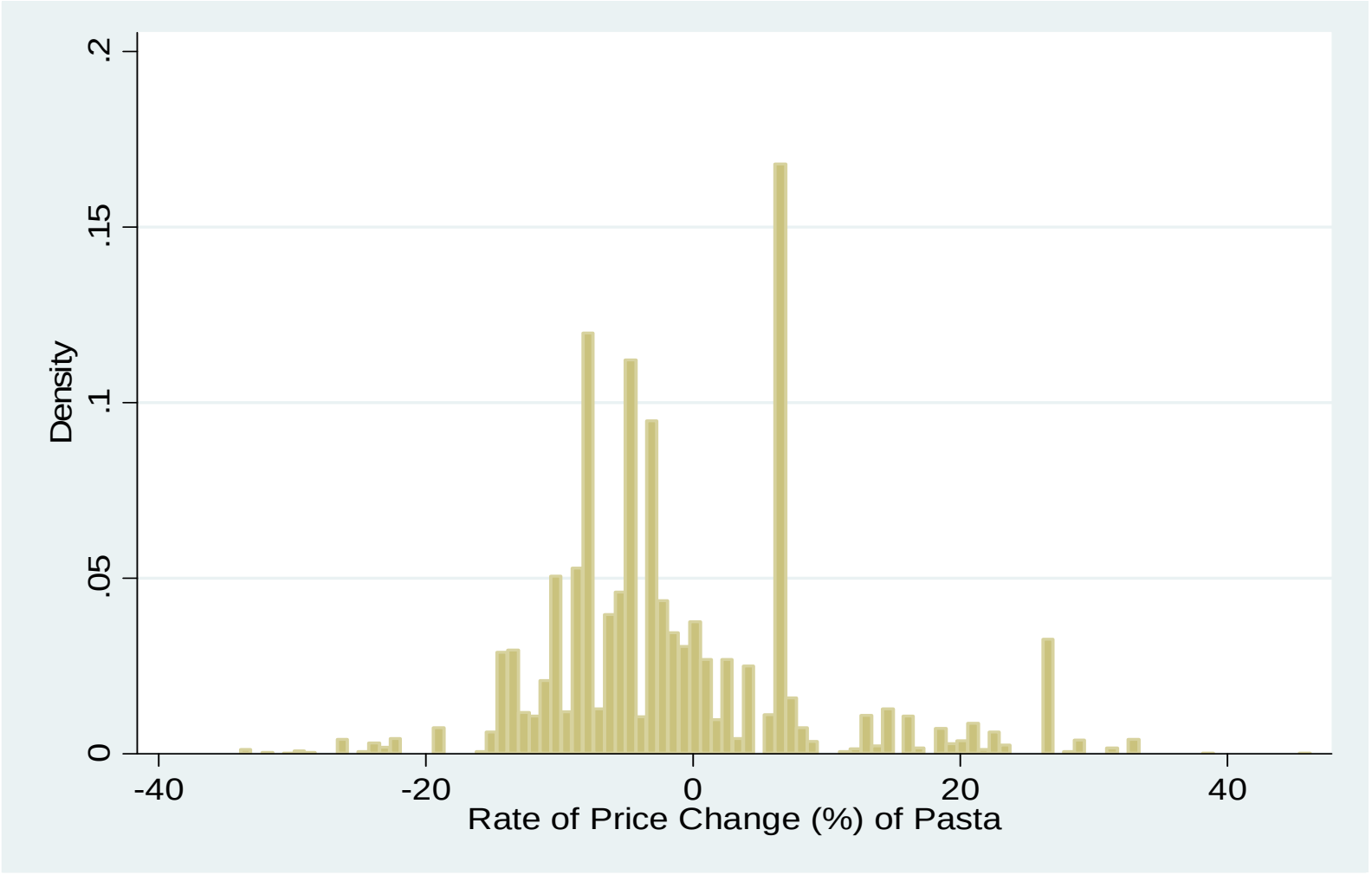




	Average	STD	Max	90th	median	10th	Minimum
Price	106.18	14.32	143	125.29	101.11	90.55	83.05

From Abe and Tonogi (2010)

In reality: The Distribution of the Rate of Price change of Pasta



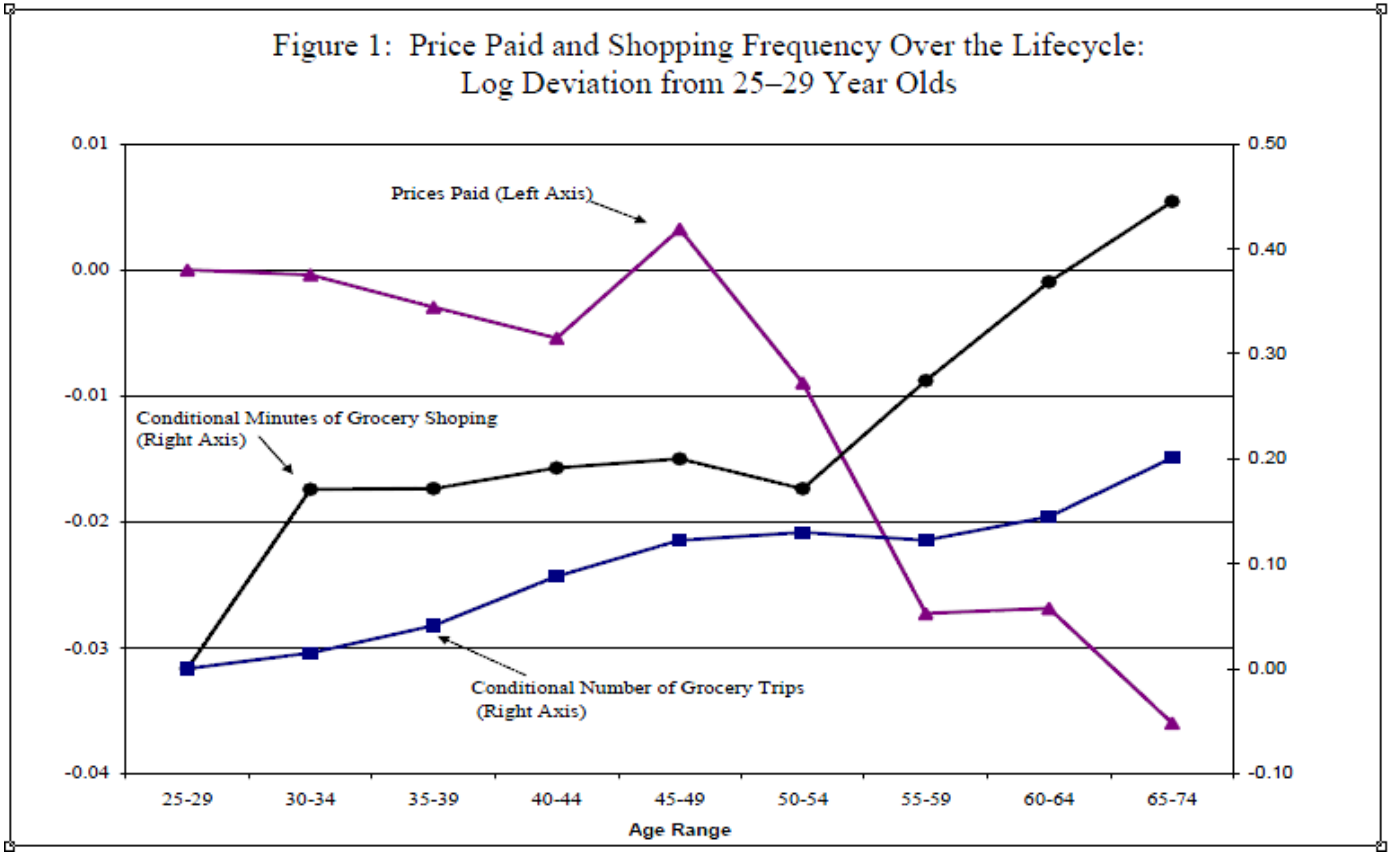
Violation of the Law of One Price

- The same commodities are traded at very different prices depending on stores and time.
- Occasionally we can observe temporally bargain sales.
- Commodities in the same category have different inflation rates.

How are the heterogeneity of the price level related with household level characteristics and behaviors?

Aguiar and Hurst (2007) -1

- Using AC Nielsen Homescan in Denver, 1993-1995.



Elderly have lower opportunity costs for shopping. Famous saving-retirement puzzle can be resolved by considering the price reduction for elderly people.

Aguiar and Hurst (2007) -2

Dependent Variable: Log (Relative Price Level)			
Estimated Elasticity	−0.010	−0.189	−0.074
	(0.006)	(0.035)	(0.032)
Measure of Shopping Intensity	Shopping trips per month	Shopping trips per month	Shopping trips per month
Regression Type	OLS	IV	IV
Instrument Set	None	Age Dummies	Income Dummies
Note: From Table3 in Aguiar and Hurst (2007)			

This paper

- Use Japanese Homescan.
- Construct household level price index. How much price dispersion exist? Its relationship with age and income.
- Is Aguiar and Hurst's story consistent with Japanese data? Is shopping intensity important?
- What is the main determinant of household level price index?

Findings

- People pay very different prices for the same commodity (consistent with A&H).
- Elderly people face higher prices (opposite to A&H!). But household level characteristics such as income and age have little (significant though) effects on the price level. Life cycle model of price level does not seem to function in Japan.
- Several variables of shopping strategies have some impact on the price level.
- Shopping frequency, the main variable of A&H, has limited power in explaining the heterogeneity in price dispersion (inconsistent with A&H).
- Purchasing during bargain sales is the most important determinants.
- Differences in store selection partly explain the price difference across age.

Homescan

- In our Research Project, we are able to access the raw data of INTAGE homescan for three years, 2004-2006.
- Data on household-commodity level transaction based on scanner installed in each household. All the purchased items with commodity code (JAN) are recorded!
- Conducted by several marketing companies all over the world (most developed countries, and some other, India, Indonesia, China).
- Traditionally, people in marketing science have used the data for forecasting future demands.
- Japanese homescan, 2004-2006 cover 12000 households all over Japan except Okinawa. Total number of shopping records is 30 million, 31,119,796.
- Drawbacks: No information on fresh foods, eat out, durable goods, car, rent, etc. Basically data on processed foods and daily use products such as detergent. (Rice and coffee beans are included.)

AH Price Index

- We can utilize actual transaction prices.
- Laspeyres or Paasche index are difficult to construct because there are so many zeros! Abe and Shiotani (2011).
- Aguiar and Hurst (2007) proposes a price index that captures the difference in commodity price.
- This compares how much each consumer pays relative to the regional average price for the same commodity.
  - Whether he/she spending 100 yen or 120 yen for cupnoodle?
  - Drawback: Cross-sectional variation only. We cannot calculate inflation.

Aguiar and Hurst's Price Index

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 $y_{i,t}^{j,c}$.

Actual Total Expenditure:

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

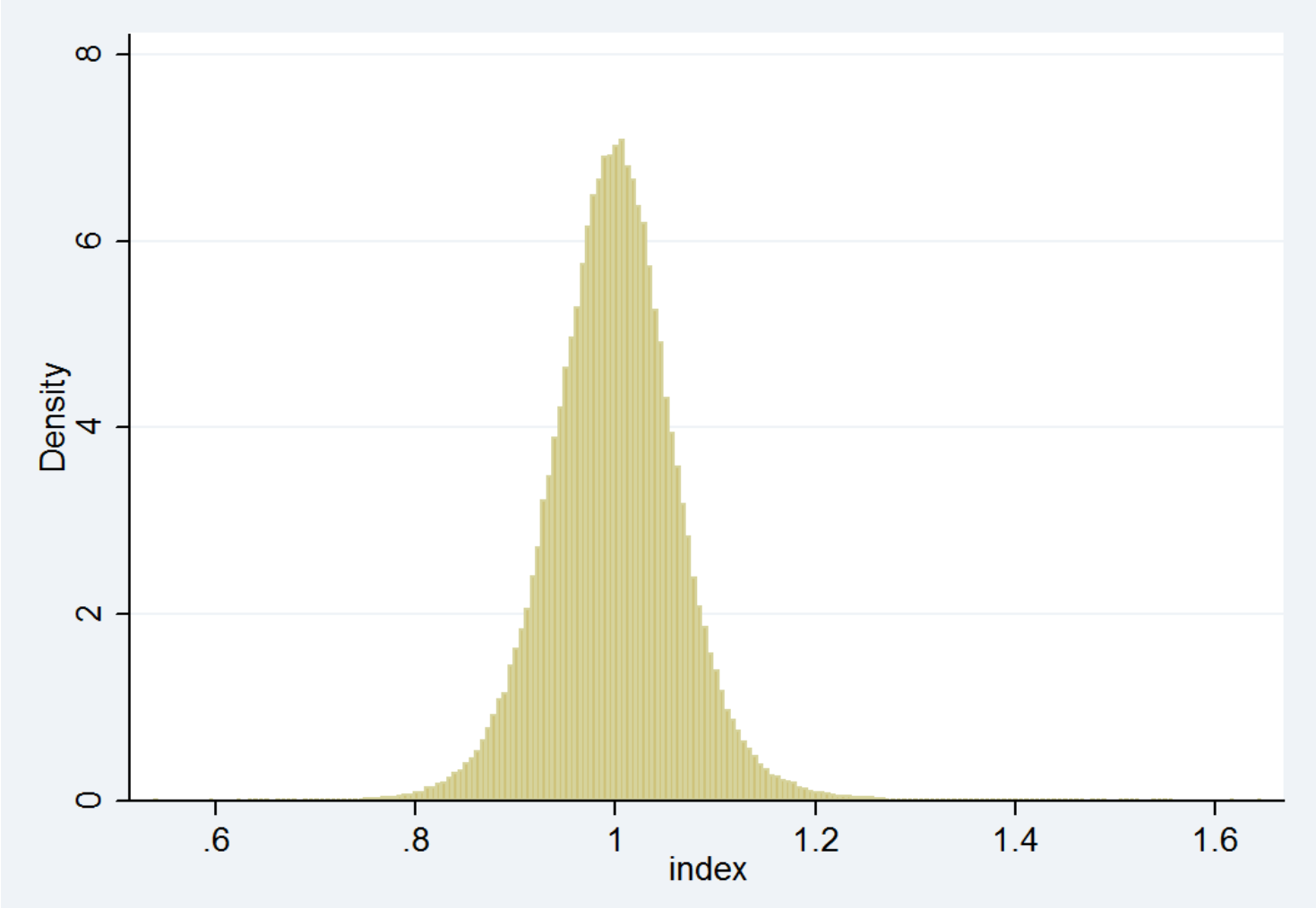
Hypothetical expenditure assuming household purchase goods at regional average price:

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

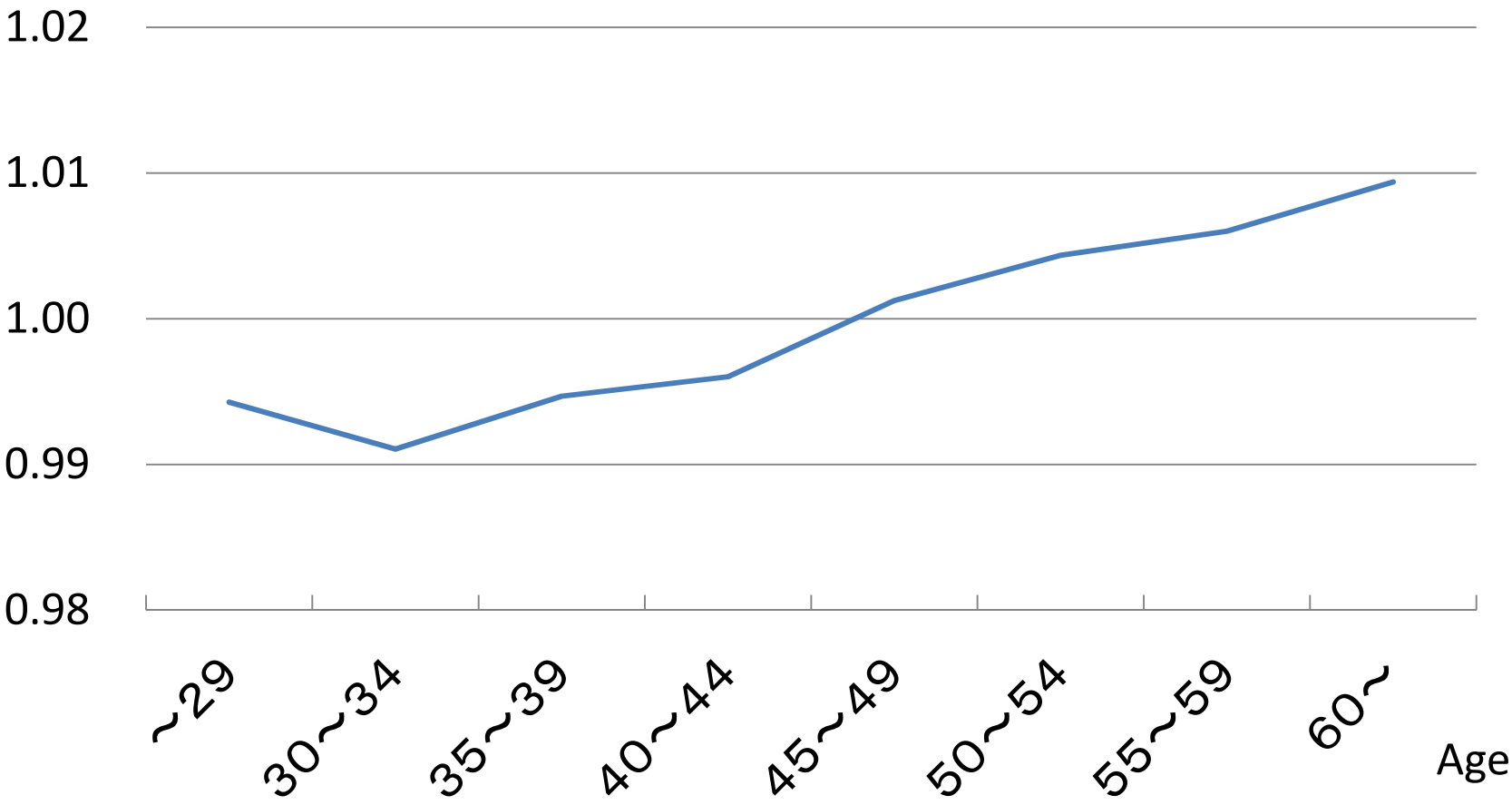
$$\tilde{p}_m^j \equiv \frac{X_m^j}{\overline{X}_m^j}.$$

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

Distribution of A&H Price Index (Log)

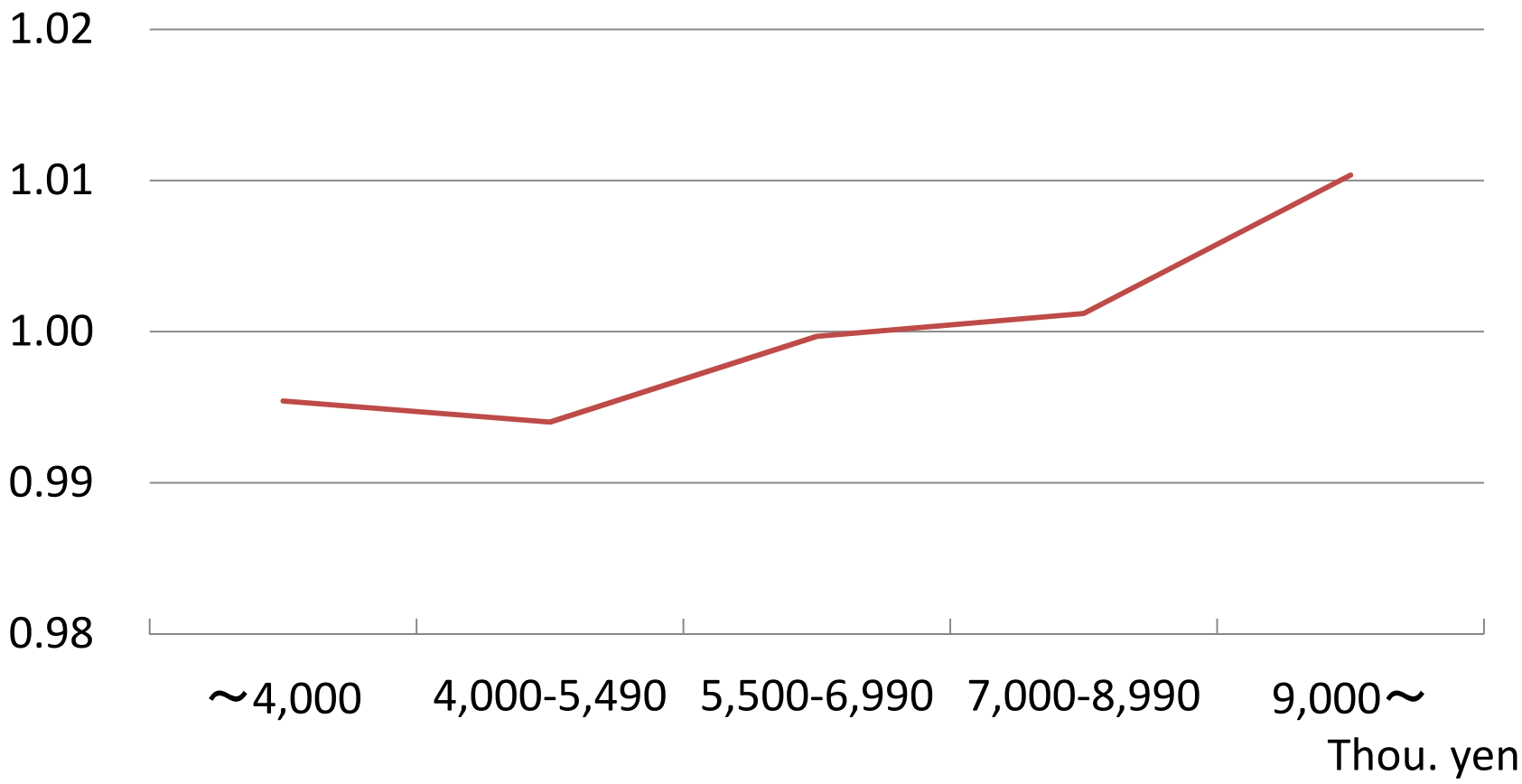


Age and Price Levels



Opposite to US, but similar to UK (Griffith et al (2009))

Income and Price Levels



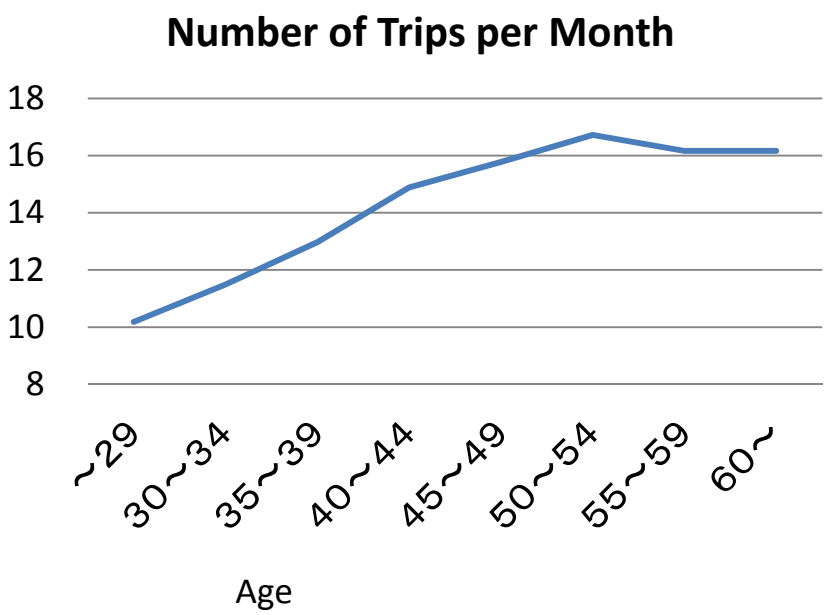
Basic Results

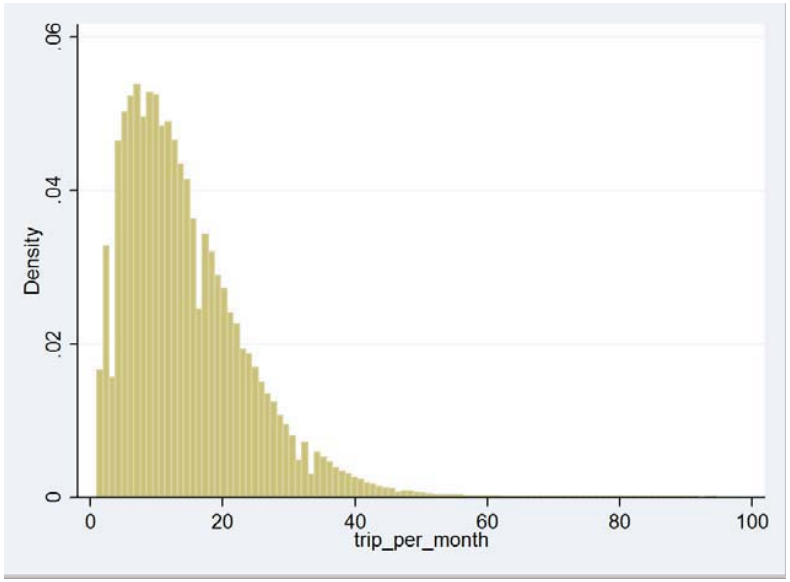
		(1)	(2)
Dummy for Income (1)	4,000-5,490	0.0019	0.0016
		(6.503)	(5.262)
	5,500-6,990	0.0059	0.0059
		(18.546)	(18.676)
	7,000-8,990	.0073	.0072
		(22.234)	(22.389)
	9,000-	0.0132	0.0133
		(38.275)	(40.347)
Dummy for Age (2)	30-34	-0.0015	-0.0021
		(-3.509)	(-4.859)
	35-39	0.0018	0.0020
		(3.931)	(4.833)
	40-44	0.0013	0.0036
		(2.721)	(8.522)
	45-49	0.0035	0.0078
		(6.571)	(17.705)
	50-54	0.0039	0.0085
		(7.051)	(19.368)
	55-59	0.0056	0.0101
		(10.121)	(23.035)
	60-	0.0107	0.0152
		(19.154)	(33.644)
	Constant	0.0326	0.0342
		(16.107)	(18.527)
	Observations	371,367	371,367
	R-squared	0.048	0.035
	Number of monitor_coc	14,442	14,442

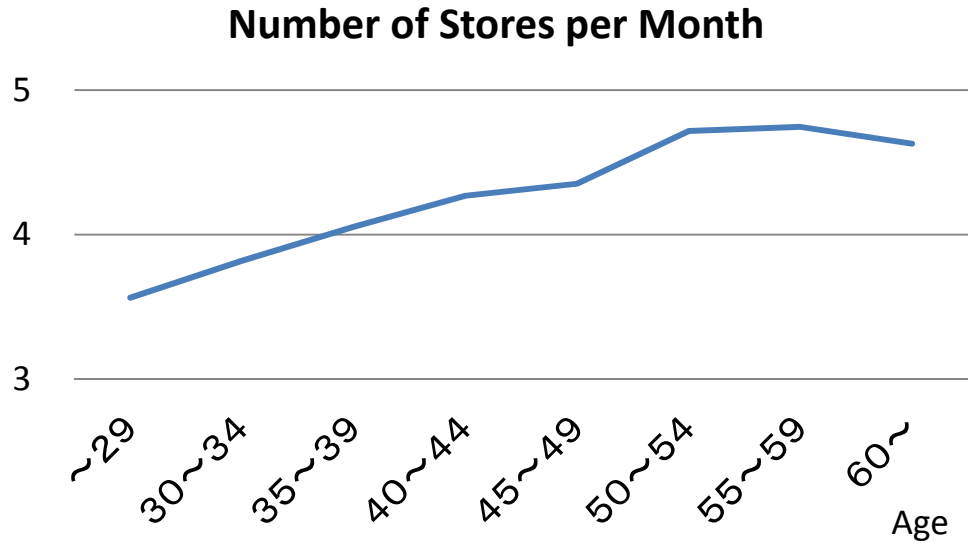
Spec (1) controlled for time dummies, locational dummies, and household characteristic

Seven Indicators for Shopping Strategy

- Shopping Frequency (-)
- The number of stores (-)
- Concentration of shopping (HHI) (+ or -)
- Quantity (-)
- Purchase at bargain (-)
- Quality of Goods (high quality goods) (+)
- Quality of Stores (cheap discount store or other stores) (+)


Number of Trips per Month
18
16
14
12
10
8
~29
30~34
35~39
40~44
45~49
50~54
55~59
60~
Age


.06
.04
.02
0
Density
0
20
40
60
80
100
trip_per_month


Number of Stores per Month
5
4
3
~29
30~34
35~39
40~44
45~49
50~54
55~59
60~
Age

Shopping Strategies

- Total number of Goods

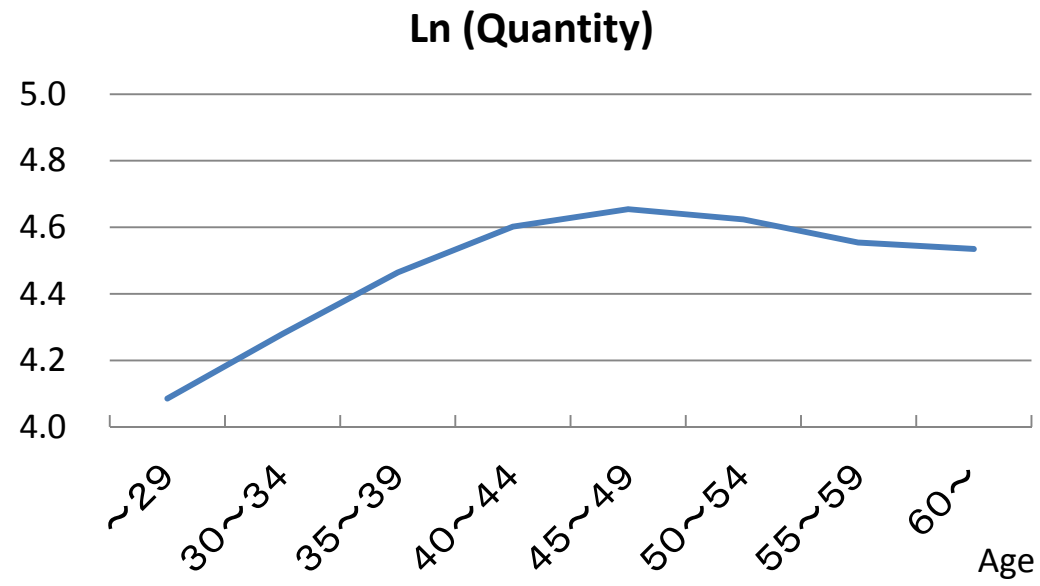
  Monthly total number of goods a household buys

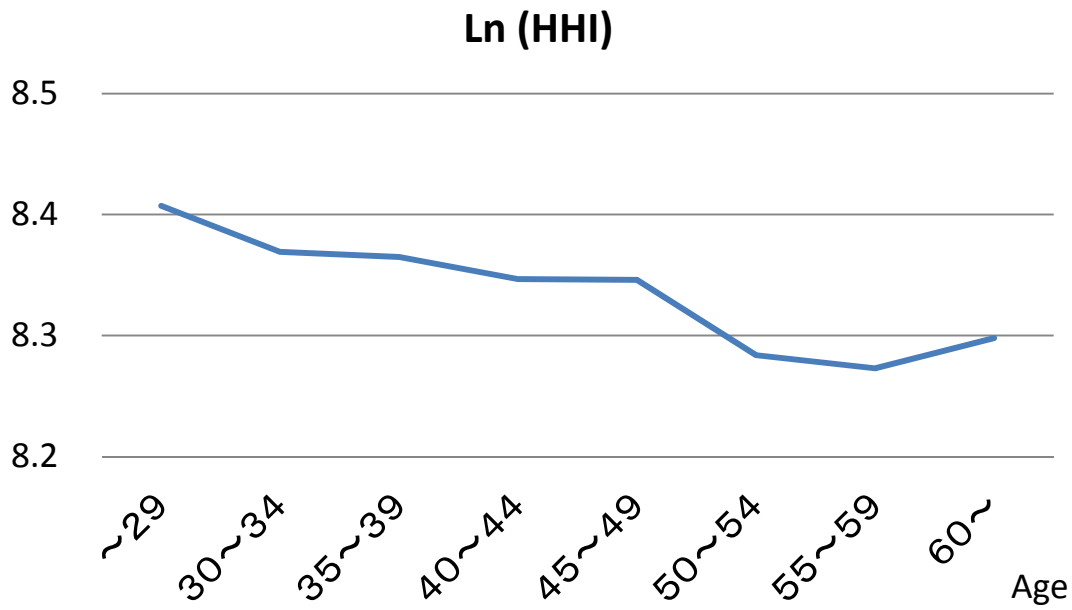
$$Quantity_m^j = \sum_{c \in C, i \in I_c, t \in m} y_{i,t}^{j,c}$$

- Herfindahl-Hirschman Index

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

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


Ln (Quantity)
5.0
4.8
4.6
4.4
4.2
4.0
~29
30~34
35~39
40~44
45~49
50~54
55~59
60~
Age


Ln (HHI)
8.5
8.4
8.3
8.2
~29
30~34
35~39
40~44
45~49
50~54
55~59
60~
Age

Shopping Strategies

- Bargain Ratio

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

where

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

$min\ P_{i,t}^{c}$:the store-level monthly minimum price for each good

$max\ P_{i,t}^{c}$:the store-level monthly maximum price for each good

Shopping Strategies

- Quality Index

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

where

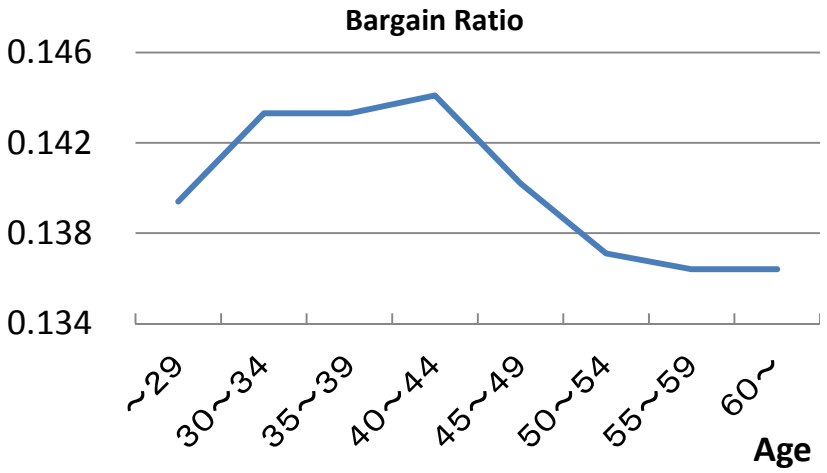
$$Z_m^j = \sum_{c\in C, i\in I_c, t\in m} \bar{p}_{i,m}^c y_{i,t}^{j,c},$$

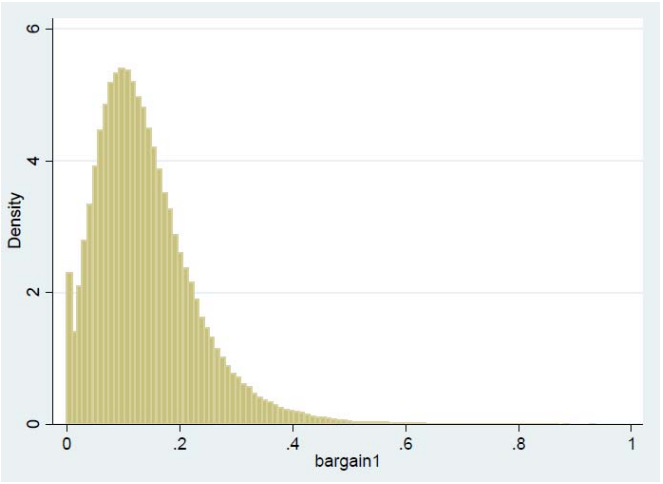
$\bar{P}_{i,m}^c$: comoditiy (i) level average price

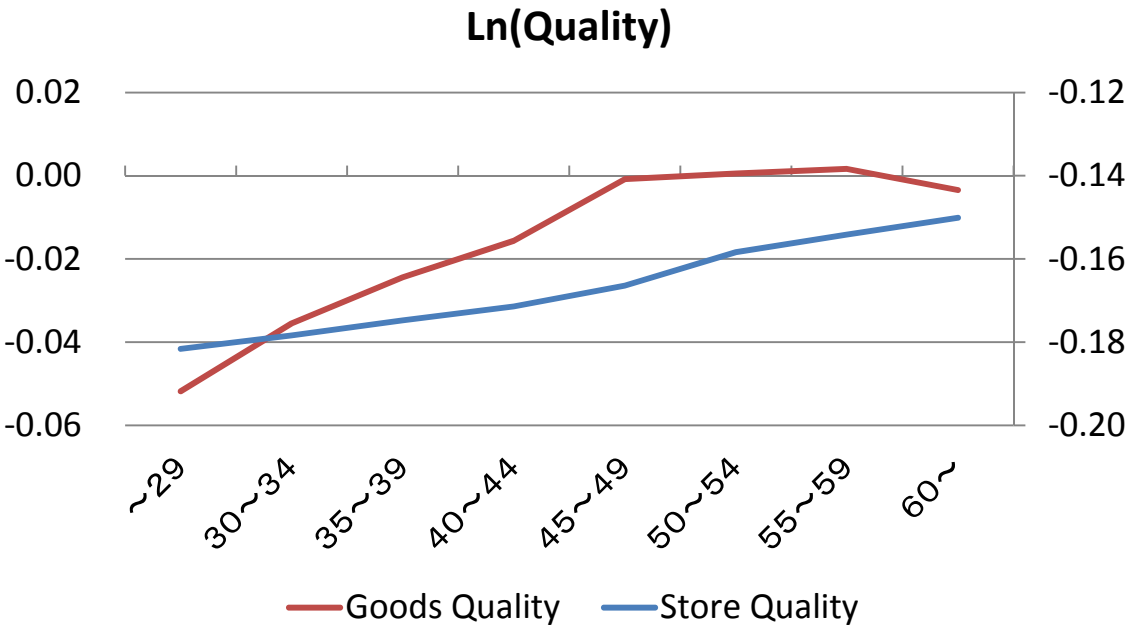
$$\bar{Z}_m^j = \sum_{c\in C, i\in I_c, t\in m} \bar{p}_m^c y_{i,t}^{j,c},$$

$\bar{P}_m^c$: categorical average price

Capture the difference arising from the difference in commodities.


Bargain Ratio
0.146
0.142
0.138
0.134
~29
30~34
35~39
40~44
45~49
50~54
55~59
60~
Age


Density
0
2
4
6
0
.2
.4
.6
.8
1
bargain1


Ln(Quality)
0.02
0.00
-0.02
-0.04
-0.06
-0.12
-0.14
-0.16
-0.18
-0.20
~29
30~34
35~39
40~44
45~49
50~54
55~59
60~
Goods Quality
Store Quality

Age

Shopping Strategies

- Store Quality Index

$$Store\ choice_m^j \equiv \sum_{k \in K} S_{k,m}^j q_m^k$$

where

$$\tilde{q}_m^k \equiv \frac{\bar{z}_m^k}{z_m^k}.$$

Is the index for the quality of goods sold at store *k* as one for each households, and

$S_{k,m}^j$: share of each store in monthly total purchases of a household *j*.

Estimation Model

$$\ln P_{it} = \alpha Income_{it} + \beta Age_{it} + \gamma ShoppingStrategy_{it} + \delta X_{ij}$$

$$+ time_t + c_{it} + \varepsilon_{it},$$

P : Natural log of Household Level Price Level Index (following AH)
Income: Categorical Dummies for household level income
Age: Categorical Dummies for Age of house wives
Shopping Strategy: Household level Strategies

X: “Shopping Needs” the number of product Categories, log of the number of commodities, and the total expenditure per month (following AH)

Time: time dummy,
C: household level information including fixed effects

OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln_trip	−0.0095	−0.0135						
	(−38.224)	(−61.168)						
ln_store	−0.0017		−0.0047					
	(−5.582)		(−23.898)					
ln_HHI	0.0035			0.0033				
	(11.712)			(15.732)				
ln_quantity	−0.0019				−0.0341			
	(−3.254)				(−69.317)			
bargain	−0.2401					−0.2489		
	(−209.735)					(−217.255)		
ln_store_choice	0.0305						0.065	
	−(23.159)						(45.236)	
ln_quality	0.0388							0.0579
	(54.630)							(100.066)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln trip	−0.0095	−0.0135						
		(−38.224)	(−61.168)						
	ln_store	−0.0017		−0.0047					
		(−5.582)		(−23.898)					
	ln_HHI	0.0035			0.0033				
		(11.712)			(15.732)				
	ln_quantity	−0.0019				−0.0341			
		(−3.254)				(−69.317)			
	bargain	−0.2401					−0.2489		
		(−209.735)					(−217.255)		
	ln_store_choice	0.0305						0.065	
		−(23.159)						(45.236)	
	ln_quality	0.0388							0.0579
		(54.630)							(100.066)
Dummy for Income (1)									
	4,000−5,490	0.0007	0.0016	0.0019	0.0019	0.0013	0.0015	0.0016	0.001
		(2.472)	(5.368)	(6.568)	(6.484)	(4.487)	(5.421)	(5.519)	(3.483)
	5,500−6,990	0.0029	0.0051	0.0058	0.0057	0.0049	0.0044	0.0053	0.0042
		(10.086)	(16.128)	(18.119)	(18.017)	(15.405)	(14.845)	(16.713)	(13.478)
	7,000−8,990	0.0039	0.0065	0.0072	0.0071	0.0061	0.0057	0.0064	0.0052
		(12.938)	(19.650)	(21.667)	(21.567)	(18.662)	(18.561)	(19.452)	(15.919)
	9,000−	0.0076	0.0121	0.0129	0.0129	0.0119	0.0098	0.0121	0.0105
		(23.881)	(34.824)	(36.916)	(36.901)	(34.171)	(30.196)	(34.575)	(30.215)
Dummy for Age (2)									
	30−34	−0.0015	−0.0015	−0.0016	−0.0016	−0.0011	−0.0012	−0.0018	−0.0021
		(−3.820)	(−3.492)	(−3.671)	(−3.562)	(−2.647)	(−2.980)	(−4.020)	(−4.923)
	35−39	0.0014	0.0019	0.0017	0.0017	0.0025	0.0017	0.0015	0.0011
		(3.356)	(4.267)	(3.824)	(3.784)	(5.492)	(4.082)	(3.311)	(2.455)
	40−44	0.0012	0.0017	0.0012	0.0011	0.0018	0.0015	0.0008	0.0004
		(2.590)	(3.372)	(2.383)	(2.307)	(3.625)	(3.211)	(1.657)	(0.761)
	45−49	0.003	0.0039	0.0032	0.0032	0.004	0.0034	0.003	0.0022
		(6.100)	(7.229)	(5.969)	(5.886)	(7.543)	(6.952)	(5.601)	(4.215)
	50−54	0.0047	0.0052	0.004	0.0039	0.0051	0.004	0.0036	0.0034
		(9.366)	(9.576)	(7.370)	(7.048)	(9.453)	(8.004)	(6.603)	(6.300)
	55−59	0.0073	0.0075	0.0059	0.0057	0.0075	0.0062	0.0054	0.0053
		(14.360)	(13.660)	(10.767)	(10.354)	(13.697)	(12.157)	(9.866)	(9.686)
	60−	0.012	0.0131	0.011	0.0108	0.0135	0.0109	0.0103	0.0103
		(23.030)	(23.197)	(19.464)	(19.122)	(24.009)	(20.894)	(18.324)	(18.462)
	Constant	0.0238	−0.0076	0.0203	−0.0023	−0.0607	0.0471	0.0734	0.0842
		(6.135)	(−3.524)	(9.514)	(−0.756)	(−24.839)	(24.422)	(32.775)	(39.653)
	Observations	371,367	371,367	371,367	371,367	371,367	371,367	371,367	371,367
	R-squared	0.215	0.059	0.05	0.049	0.062	0.19	0.056	0.074
	Number of monitor_code	14,442	14,442	14,442	14,442	14,442	14,442	14,442	14,442

Fixed Effects

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln_trip	0.0042	0.0036						
		(9.590)	(9.533)						
	ln_store	0.0006		0.0025					
		(1.647)		(9.034)					
	ln_HHI	0.004			-0.0011				
		(10.571)			(-3.329)				
	ln_quantity	0.0003				-0.0175			
		(0.257)				(-21.362)			
	bargain	-0.1288					-0.1311		
		(-102.358)					(-103.887)		
	ln_store_choice	0.021						0.0349	
		(10.046)						(16.063)	
	ln_quality	0.0264							0.0325
		(27.956)							(41.500)

Comparisons with A & H

Dependent Variable: Log (Relative Price Level)						
Estimated Elasticity	−0.010	−0.189	−0.074		−0.0126	0.0036
	(0.006)	(0.035)	(0.032)		(0.00025)	(0.00039)
Measure of Shopping Intensity	Shopping trips per month	Shopping trips per month	Shopping trips per month		Shopping trips per month	Shopping trips per month
Regression Type	OLS	IV	IV		OLS	FE
Instrument Set	None	Age Dummies	Income Dummies		None	None
From Table3 in Aguiar and Hurst (2007)					Japanese Data	

Age and income dummies are very weak instruments in Japan

The effects of an increase by one standard deviation of each variable on ln (Prices)

	ln_trip	ln_store	ln_HHI
SD	0.73807	0.60218	0.48529
Coefficients	0.00360	0.00250	−0.00110
Effects on ln Prices	0.00266	0.00151	−0.00053
	bargain	ln_store_quality	ln_quality
SD	0.08840	0.08413	0.17129
Coefficients	−0.1311	0.03490	0.03250
Effects on ln Prices	−0.01159	0.00294	0.00557

Why Elderly Faces Higher Prices in Japan?

(1)Do Elderly go to stores that sell at higher prices more often than young?
-- we add the expenditure shares of each store type; convenient stores, discount store, supermarket, etc…

(2)Given the selection of stores, do elderly purchase at higher prices than young?

Table 8: 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	110.0	1.07%	99.19	6.66%	99.28	17.78%	97.85	6.19%
	30～34	109.0	0.88%	98.90	6.57%	98.87	14.48%	97.58	6.18%
	35～39	108.7	0.89%	99.27	5.26%	98.85	11.16%	97.58	5.68%
	40～44	108.6	1.01%	99.59	5.22%	98.61	10.05%	97.83	5.01%
	45～49	107.9	0.78%	99.68	5.98%	99.20	9.34%	98.35	5.60%
	50～54	109.2	1.06%	99.84	7.85%	99.16	9.22%	98.83	5.93%
	55～59	108.5	1.02%	100.67	9.20%	99.44	8.48%	98.89	5.47%
	60～	107.9	0.92%	100.84	10.61%	99.85	7.90%	99.29	4.53%
	Total	108.6	0.96%	99.91	7.02%	99.08	11.11%	98.27	5.58%

		Home delivery & Door to door sales		Supermarket		Others			
		price	share	price	share	price	share		
age of wife	～29	100.30	4.87%	99.67	62.5%	99.53	1.15%		
	30～34	100.42	6.78%	99.37	64.2%	99.81	1.09%		
	35～39	100.41	7.80%	97.26	68.6%	99.86	0.75%		
	40～44	100.41	8.43%	99.83	69.7%	99.52	0.81%		
	45～49	100.57	8.43%	100.19	69.3%	99.99	0.83%		
	50～54	100.85	7.98%	100.36	67.2%	101.15	1.00%		
	55～59	100.93	7.50%	100.60	67.2%	100.74	1.33%		
	60～	100.97	5.94%	100.89	68.2%	101.72	2.16%		
	Total	100.62	7.65%	100.11	67.7%	100.50	1.19%		

OLS Estimates after controlling for store selection

		(1)	(2)	(3)	(4)
				After Controlling for Store Selection	
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)
Spec (1) and (3) controlled for time dummies, locational dummies, and household characteristics.					
Spec (3) and (4) controlled for store selection.					

Summary

- People face very different price level.
- Price levels are increasing with age and income. But, the magnitudes of the effects are not large.
- People adopt very different shopping strategies. Bargain sales seems most important in determining the price level.
- Elderly shop more, but purchase smaller quantity, use less bargain.
- Bargain sales seems to be the key to understand the price level differential between households.
- Elderly go to specialized stores more often than young. Such difference in store choices can explain about 10% of the price differences across age.
- Within the same store choices, elderly purchases goods at higher prices than young! So, it is a remaining task.

Supplementary

Age and Income Profiles

		price	ln_trip	Number of Trips	ln Number of Stores	Bargain	ln quantity	ln HHI	ln store choice	ln quality
age of wife	～29	0.994	2.066	10.181	1.278	0.139	4.085	8.407	-0.182	-0.052
	30～34	0.991	2.197	11.491	1.345	0.143	4.280	8.369	-0.178	-0.035
	35～39	0.995	2.325	12.969	1.404	0.143	4.465	8.365	-0.175	-0.024
	40～44	0.996	2.459	14.886	1.456	0.144	4.602	8.347	-0.171	-0.016
	45～49	1.001	2.528	15.777	1.475	0.140	4.654	8.346	-0.166	-0.001
	50～54	1.004	2.611	16.719	1.557	0.137	4.624	8.284	-0.158	0.001
	55～59	1.006	2.598	16.162	1.563	0.136	4.554	8.273	-0.154	0.002
	60～	1.009	2.598	16.163	1.529	0.136	4.535	8.298	-0.150	-0.003
income	～4000	0.995	2.343	13.138	1.369	0.144	4.350	8.381	-0.174	-0.049
	4000-5490	0.994	2.389	13.680	1.425	0.144	4.444	8.351	-0.173	-0.031
	5500-6990	1.000	2.422	14.216	1.455	0.141	4.503	8.340	-0.168	-0.013
	7000-8990	1.001	2.492	15.250	1.500	0.141	4.577	8.313	-0.162	0.001
	9000～	1.010	2.543	15.768	1.543	0.133	4.612	8.282	-0.155	0.022
	Total	1.000	2.435	14.365	1.456	0.140	4.495	8.334	-0.167	-0.015