Working Papers

FEDERAL RESERVE BANK OF BOSTON™

No. 18-4

A Concise Test of Rational Consumer Search

Chloe Lee, María José Luengo-Prado, and Bent E. Sørensen

Abstract:

A simple model of time allocation between work and price-search predicts that consumers spend relatively more time searching for better prices for goods of which they consume relatively more. Using scanner data, we confirm empirically that consumers pay lower (higher) prices for goods that they buy more (less) of than other consumers. Our results are conservative, because we compare goods that are defined as narrowly as possible by UPC codes, and provide a lower bound for the savings obtained from bargain hunting.

JEL Classifications: E21, D12
Keywords: consumption, consumer search, time use

Chloe Lee is a senior research assistant in the research department at the Federal Reserve Bank of Boston. Her e-mail address is chloe.lee@bos.frb.org. María José Luengo-Prado is a senior economist and policy advisor in the Boston Fed research department. Her e-mail address is maria.luengo-prado@bos.frb.org. Bent E. Sørensen is a professor of economics at the University of Houston and a fellow of the Programme in International Macroeconomics at the Center for Economic and Policy Research. His email address is besorensen@uh.edu.

This paper originated from joint work with Wen Long, and the authors thank him for sharing his knowledge of the IRI dataset with them.

This version: June 18, 2018

1 Introduction

Consumers trade off the value of time spent finding lower prices with the value of time spent on other pursuits. We conduct a simple novel test of rational price-search: A consumer who frequently drinks soda but seldom drinks beer should rationally spend more time searching for low-priced Coke than for low-priced Budweiser. Using detailed shopping information from the IRI academic dataset, we confirm this prediction. The dataset records the purchases that a panel of households made over an 11-year period at a selection of stores in Eau Claire, Wisconsin, and Pittsfield, Massachusetts. The data derive from UPC (scanner-code) transactions, the narrowest possible measurement. For each transaction, both the price of the item and the quantity purchased are reported.

Stigler (1961), in a pioneering paper, suggested that information is scarce and consumers invest time in finding favorable prices—an activity that he labeled "search." As summarized in Kaplan and Menzio (2013), many recent papers examine price-search using scanner data under the heading of "bargain hunting." Aguiar, Hurst, and Karabarbounis (2013) use the American Time Use Survey to show that households in states with higher unemployment spend relatively more time on home production and shopping, and Coibion, Gorodnichenko, and Hong (2015) use scanner data from IRI to show that consumers obtain better prices during recessions by switching to different stores when they shop.[1] Nevo and Wong (forthcoming), using Nielsen Homescan data, show that during the Great Recession consumers obtained lower prices by, among other practices, using more coupons, purchasing more sales items, and shopping more often at "big box" stores.[2] Nevo and Wong (forthcoming) also find that the return to shopping declined during the Great Recession, so the

[1] Aguiar, Hurst, and Karabarbounis (2013) show that about 30 percent of lost labor hours were reallocated toward non-market work, including shopping, during the Great Recession.

[2] Griffith et al. (2009) summarize four channels of saving: purchasing items when they are on sale, buying in bulk (at lower per-unit prices), buying generic brands, and shopping at outlets.

increased amount of search is consistent with a lower shadow value of time.

The literature has found intuitively reasonable differences in shopping behavior across individuals. Aguiar and Hurst (2007) show that retirees spend relatively more time shopping, and Stroebel and Vavra (2014), using changes in house prices to isolate exogenous changes in wealth, find that wealthier households spend relatively less time shopping. Chevalier and Kashyap (2011) examine purchases using the IRI data and posit a model with two types of consumers: (1) "shoppers," who pay the best price possible because they chase discounts, substitute across products, and/or store goods they purchase during sale weeks, and (2) "loyals," who buy only one brand and do not time purchases to coincide with sales. In this setting, it is optimal for firms to maintain a combination of constant regular prices and frequent short-lived sales.

Kaplan et al. (2016) use the Nielsen scanner dataset and find that most variation in prices happens within stores rather than across them. They construct a model with two groups of consumers: (1) "busy," who make all purchases in one store, and (2) "cool," who shop at several stores. Under their assumptions, in equilibrium stores will charge different prices for the same goods—intuitively, busy consumers will buy expensive and less-expensive goods in the same store, while consumers with more time for shopping will buy the cheaper goods at each store. Our paper is the first to document that individuals display different patterns of price-search across goods—becoming "cool shoppers" when buying diapers, for example, and "busy loyals" when buying beer.

Macroeconomists have paid attention to rational inattention since Reis (2006) found that consumption patterns are consistent with the notion that consumers update information infrequently due to difficulty acquiring, absorbing, and processing information. A large literature (see Sims 2010) —somewhat disconnected from the literature on bargain hunting—focuses on rational inat-tention and its implications for monetary policy. Our work is related to this literature in that we posit that consumers pay less attention to less important (for them) goods.

Following Aguiar and Hurst (2007), we construct for each consumer a bargain hunting index (BHI) which measures the price he or she paid for a consumption bundle relative to the cost of the same bundle based on average UPC prices in the same town and week. We refine the BHI to the category level, where a category (as defined by IRI) is a group of UPCs for similar goods, such as "carbonated drinks." Studying behavior at the category level, rather than the UPC level, reduces noise significantly, because most consumers purchase only a tiny fraction of UPCs. Figure 1 lists the categories we include.

Using the category BHI, we ask whether consumers who purchase more units than other consumers in a particular category pay relatively lower prices in that category. We find that they do, and the finding is very robust: If we regress prices on quantities category by category, the pattern holds for all categories. This result also holds in regressions with individual-specific fixed effects, which implies that consumers who increase their consumption in a category will find lower prices for the goods in that category. We also find that, consistent with other studies, retirees pay less, high-spending consumers ("wealthy consumers") pay more, and there are, in general, large differences in the prices consumers pay on average.

Building on earlier time-allocation models, such as those in Becker (1965), Benhabib, Rogerson, and Wright (1991), and Greenwood and Hercowitz (1991), we outline a simple static search model that we use to interpret our empirical results. The model predicts that consumers pay relatively less for goods of which they purchase relatively more. It also predicts that individuals with high wages pay relatively more and that retired individuals on fixed income pay relatively less.

We do not model the price setting of stores. Using scanner data from Nielsen, Kaplan, et al.(2016) find that price differences between stores are quite persistent, with more variation in prices across goods within a store than across stores. The logic of the model likely can be applied to consumers who search more or less intensively for different goods.

The rest of the paper is organized as follows: Section 2 derives a simple model of time use; Section 3 describes the data; Section 4 presents our results; and Section 5 concludes.

2 A Stylized Model

Consider a consumer who derives utility as summarized by the objective function

$$\max_{C_1,C_2,T_Y,T_1,T_2} \quad \alpha_1 \ln(C_1) + \alpha_2 \ln(C_2) + \mu \ln(T - T_Y - T_1 - T_2), \tag{1}$$

$$\text{subject to:} \qquad P_1 C_1 + P_2 C_2 = Y(T_Y), \tag{2}$$

where T_Y is the time devoted to income-generating activities (work). T_i $(i = 1, 2)$ is time spent searching good i. $T - T_Y - T_1 - T_2$ is leisure time, and T is the total time endowment. Time spent searching results in lower prices according to the function $P_i = T_i^{-\beta}$, with $\beta > 0$, so the larger that β is, the more that searching lowers the price paid. The marginal effect of an additional unit of search is $\frac{dP_i}{dT_i} = -\beta T_i^{-\beta-1} = -\beta P_i T_i^{-1}$. Income is a linear function of time spent working, with wage rate W_1, and nonwage income, W_0: $Y(T_Y) = W_0 + W_1 T_Y$. C_1 and C_2 are the purchased quantities of good 1 and good 2, and α_1/α_2 is the preference for good 1 over good 2, with $\alpha_1 + \alpha_2 = 1$.

The Lagrangian is $L = \alpha_1 \ln(C_1) + \alpha_2 \ln(C_2) + \mu \ln(T - T_Y - T_1 - T_2) + \lambda[Y(T_Y) - P_1C_1 - P_2C_2]$, and the first order conditions (FOCs) with respect to (wrt) consumption are:

$$\frac{\alpha_i}{C_i} = \lambda P_i\,;\ i = 1, 2\,. \tag{3}$$

This implies that $\frac{C_1}{C_2} = \frac{\alpha_1 P_2}{\alpha_2 P_1}$; that is, a higher α_1 (higher weight on good 1) increases C_1 over C_2. A higher relative price of good 2 has the same effect. Substituting into the budget constraint, (2), we find that expenditure shares for the two goods are constant: $P_1C_1 = \alpha_1 Y(T_Y)$ and $P_2C_2 = \alpha_2 Y(T_Y)$.

The FOCs of the Lagrangian wrt T_i, $i = 1, 2$ are:

$$-\mu(T - T_Y - T_1 - T_2)^{-1} - \lambda C_i \frac{dP_i}{dT_i} = 0;\ i = 1, 2\,. \tag{4}$$

Combining them, we find that the marginal gain from search time is equalized across goods:

$$\frac{dP_2}{dT_2} C_2 = \frac{dP_1}{dT_1} C_1, \quad \text{or} \quad \frac{\frac{dP_2}{dT_2}}{\frac{dP_1}{dT_1}} = \frac{C_1}{C_2}.$$

Given that $C_i = \alpha_i/\lambda P_i$, from FOC (3), and that $dP_i/dT_i = \beta T_i^{-\beta-1}$, we can rewrite the previous expression as:

$$\frac{-\beta T_2^{-\beta-1}}{-\beta T_1^{-\beta-1}} = \frac{\alpha_1}{\alpha_2} \frac{T_2^{-\beta}}{T_1^{-\beta}} \qquad \text{or} \qquad \frac{T_1}{T_2} = \frac{\alpha_1}{\alpha_2}\,.$$

That is, relative time allocated to searching for goods is proportional to their relative preferability.

The FOC of the Lagrangian wrt T_Y is $-\mu(T - T_Y - T_1 - T_2)^{-1} + \lambda \frac{dY}{dT_Y} = 0$, and combining this FOC with FOC (4), we obtain $-C_i \frac{dP_i}{dT_i} = \frac{dY}{dT_Y}$. That is, the marginal gain from a unit increase in shopping time is equal to the marginal loss of income.

Substituting for the price derivative, we obtain $C_i \frac{dP_i}{dT_i} = C_i(-\beta P_i T_i^{-1}) = -\beta(P_i C_i) T_i^{-1}$, and since $C_i P_i = \alpha_i Y(T_Y)$, we find $\beta \alpha_i Y(T_Y) T_i^{-1} = \frac{dY}{dT_Y}$ or $T_i = \beta \alpha_i \frac{Y(T_Y)}{\frac{dY}{dT_Y}}$, implying that $T_i = \beta \alpha_i \left(\frac{W_0}{W_1} + T_Y \right)$ and $T_1 + T_2 = \beta \left(\frac{W_0}{W_1} + T_Y \right)$. FOC (3) and the fact that $C_i P_i = \alpha_i Y(T_Y)$ imply that $\lambda = 1/Y(T_Y)$. Given that $\frac{dY}{dT_Y} = W_1$ and substituting for λ, we can rewrite the FOC wrt T_y as:

$$\mu(T - T_Y - T_1 - T_2)^{-1} = \frac{W_1}{W_0 + W_1 T_y}.$$

Substituting for the value of $T_1 + T_2$, we can solve for T_y :

$$T_y = \frac{T - (\beta + \mu)\frac{W_0}{W_1}}{1 + \beta + \mu}\ . \tag{5}$$

Work time increases in total time available and in wages (W_1), and it decreases in nonwage income (W_0), price-search efficiency (β), and leisure preference (μ).[3] Plugging the value of T_y into the solution for T_i, we can solve for T_i :

$$T_i = \beta\alpha_i \frac{T + \frac{W_0}{W_1}}{1 + \beta + \mu}. \tag{6}$$

Relative time spent searching is proportional to the utility weights, so agents rationally allocate more time to preferred goods. Search time increases with total time available, nonwage income, and search efficiency. Search time decreases with wages and leisure preference. Leisure is $T - T_1 - T_2 - T_y = \frac{\mu}{1+\beta+\mu}\left(T + \frac{W_0}{W_1}\right)$.

For a fixed W_0, a larger wage, W_1, implies more work time, greater income, and less search time, so the "wealthy" (in terms of labor income) pay more. Work time decreases with the ratio W_0/W_1, and T_i increases, so retirees are predicted to pay lower prices.[4] In the empirical section, we interpret "time spent searching" more broadly. Consumers may have to expend effort by paying attention to prices (to determine, for example, whether there are regular sales), consistent with the references to consumers' limited mental capacity in the rational inattention literature, or they may more literally spend time by driving to a larger selection of stores.

3 Data Description

3.1 The IRI Academic Dataset

We use the IRI academic dataset which, as Bronnenberg, Kruger, and Mela (2008) describe in detail, contains weekly transaction information on the purchases of groceries in 31 item categories.

[3] We assume that the parameters and W_0 and W_1 are such that T_Y is non-negative.

[4] For simplicity, assume W_0/W_1 is such that work time is 0 for retirees.

Our dataset spans 2001 through 2012 and includes information about purchases at the store level and at the individual level. At both levels, weekly total dollar and unit sales are collected for each UPC item. A UPC is encoded in a bar code used for scanning at the point of sale, and it contains information on very specific product attributes, such as volume, product type, brand name, package size, and even flavor or scent for some products. Products that are essentially the same but differ in size or packaging have different UPCs; for example, a bottle of Budweiser beer intended for single sale has a different UPC code from a physically identical bottle of Budweiser beer sold in a six-pack.

The store-level data contain weekly total-dollar and unit-sales information for each UPC from grocery stores and drug stores in 50 IRI markets (metropolitan areas). Most stores belong to large chains (masked), and each store has a unique identifier. The individual-level panel dataset provides price and quantity information for all transactions (where a "transaction" is a UPC-specific purchase) made by a consumer panel in two small markets (cities): Eau Claire, Wisconsin, and Pittsfield, Massachusetts. The dataset includes some demographic information about the consumers, such as age, marital status, education, income, employment status, and family size. However, these variables are collected sporadically, reported only for discrete categories, and not consistently coded over time, so we include only a dummy for 65-plus years of age in our regressions. Our main results are not sensitive to inclusion of panelist fixed effects, which control for all non-time-varying consumer characteristics, so it is unlikely that including this information would alter our conclusions.

The IRI dataset also includes a supplemental "trips file" that provides information on when (week) and where (store number) each panelist went shopping, as well as the amount of money spent while shopping. We calculate the total number of trips each panelist made to stores in a given period and the number of stores visited. We mainly use the individual-level transaction data, but we use price information from the store-level dataset to calculate average market prices by UPC.

We exclude the years 2001 and 2002 due to incomplete information and inconsistencies with later years, and we exclude "soup" purchases due to unrealistic price variation that indicates low data quality for this category—the exclusion of these years and this product category does not significantly affect our results. We omit some observations with Florida ZIP codes. For regressions on overall expenditure, we drop panelist $\times$ quarter observations if the panelist's expenditure in the quarter is less than \$10. For regressions on category expenditure, we drop the panelist $\times$ quarter $\times$ category cell if the panelist's expenditure in the quarter is less than \$5 for that category.[5]

The appendix gives more details about the consumer panel, including the brackets in which income, age, and education are reported. Table A.1 displays summary statistics for the panelists in 2007. Average education is 13.8 years, and average age is 55.4 years. Individuals in our sample are between 21 and 70 years old. Average income is \$52,302 with a standard deviation of \$36,606 (the standard deviation is likely lower than the actual standard deviation because income is reported in brackets). About a third of the sample is over 65. Average expenditure is about \$200 per quarter.

Compared with the Panel Study of Income Dynamics (PSID), a representative sample for the United States (for which we do not tabulate the numbers), the IRI panelists in 2007 are somewhat older (the average age of a PSID household head is 50), poorer (average income in the PSID is \$67,000), and similarly educated (the average number of years of education completed in the PSID is 13.1). In the PSID, the average food-at-home expenditure in 2007 is roughly \$4,400. Using that number as an approximation of average food consumption for our sample, it implies that spending on categories and stores in the IRI dataset constitutes 17 to 22 percent of food-at-home

[5]IRI includes only respondents who make at least one transaction in each of the 13 four-week periods in each year. (The documentation does not make this more precise.)

expenditure.[6]

3.2 Data Construction

Similarly to Aguiar and Hurst (2007), we define the average price of a UPC item u in a given market m and week w as a quantity-weighted average of individual transaction (k) prices for that specific product in that week and market. A transaction in our analysis is the purchase of a given UPC/good during a visit to a store; in other words, one visit to a store usually comprises many transactions. The average price is

$$\overline{p}^{u,m,w} = \frac{\sum_{k \in u,m,w} q_k^{u.m,w} p_k^{u,m,w}}{\sum_{k \in u,m,w} q_k^{u,m,w}} \,,$$

where q_k is the quantity purchased in transaction k (involving UPC u), and p_k is the unit price. To compute this average price, we use the store-level dataset, which includes all transactions in all stores in a given market. We refer to this price as the average store-posted price.

We define a bargain hunting index for consumer i in period t ($\text{BHI}_{i,t}$) as the amount a consumer saves for the products he or she buys relative to the cost of the exact same products at average prices in the same week and market. Specifically, the bargain hunter index is computed as follows:

$$\text{BHI}_{i,t} = \left(1 - \frac{\text{Actual Exp}_{i,t}}{\text{Hypo Exp}_{i,t}}\right) \times 100 = \left(1 - \frac{\sum_{w \in t} \sum_{k=1}^{N_i^w} p_{i,k}^{u,m,w} \times q_{i,k}^{u,m,w}}{\sum_{w \in t} \sum_{k=1}^{N_i^w} \overline{p}^{u,m,w} \times q_{i,k}^{u,m,w}}\right) \times 100 \,, \tag{7}$$

where i is a consumer who purchases products in market m, and we aggregate expenditure to the quarterly frequency t. u is a UPC, and k indices all transactions of consumer i in week w, with N_i^w being the total number of transactions—a consumer purchases many products and can purchase a particular product more than once a week, so the number of transactions is at least as large as the number of different goods purchased. For each purchase of a good by a household (transaction)

[6] The lower number does not adjust for income differences in the two samples, whereas the higher number does.

in a week, we use the exact price of the good (identified by its UPC, $p_{i,k}^{u,m,w}$) to calculate actual expenditure. Given the consumer's own consumption bundle, hypothetical expenditure is measured using the average store-posted price ($\overline{p}^{u,m,w}$) of the good in the same week and city. Expenditure is aggregated over a quarter in order to avoid a large number of zero observations at the weekly frequency. A higher BHI means saving more (paying less) relative to the average store-posted prices given the household's consumption bundle. We also calculate the number of shopping trips for each individual in a quarter, and the average number of stores visited per week during the quarter.

Table 1 displays the mean and standard deviation of the BHI (along with summary statistics for other variables used in the regressions). The average BHI is 5.6 percent, which means that panelists save 5.6 percent on average by finding better-than-average prices. The average price for each UPC is calculated outside the panelist sample and includes transactions by all shoppers in these markets; a positive average likely reflects that panelists in our sample are, on average, older than the typical population. Going forward, we demean the BHI to 0 each period, as is standard in the literature.[7]

Our main focus is on selective bargain hunting (that is, whether consumers devote relatively more time to searching for lower prices for goods they prefer). To test for such a pattern, we construct (1) a bargain hunting index by individual and category in each quarter, and (2) a quantity index by category that measures whether a consumer buys relatively more or less of that category to proxy for his or her preferences. In principle, we could use the data by, say, week and UPC, but this would complicate the statistical analysis, because almost all observations would be zeros. We could also construct groups of UPCs ourselves, but because there is no obvious way of doing this, and an arbitrary choice of categories would open up a scope for data mining, we utilize the categories as defined by IRI.

[7]This is not strictly necessary for the regression analysis, because we include period (year×quarter) fixed effects.

Let c denote a category. A BHI by category for a given consumer i in period t, $\mathrm{BHI}^c_{i,t}$, is computed similarly to the overall BHI, except that only transactions involving products in a given category are added up:

$$\mathrm{BHI}^c_{i,t} = \left(1 - \frac{\sum_{w \in t} \sum_{k \in c} p^{u,m,w}_{i,k} \times q^{u,m,w}_{i,k}}{\sum_{w \in t} \sum_{k \in c} \overline{p}^{u,m,w} \times q^{u,m,w}_{i,k}}\right) \times 100\,. \tag{8}$$

Figure 1, Panel A, presents a box plot of the BHI by category, illustrating the range of prices paid, and thus consumers' potential for saving by searching, which varies by category. In the graph, IRI's categories are ordered by the interquartile range of the category-specific BHIs. For example, the interquartile range for beer is 1/11th of that for laundry detergent (1.66 percent versus 18.2 percent). This significant difference is likely due to very disparate pricing strategies employed by retailers and/or producers of the two products—in our pooled analysis, we include category fixed effects, when relevant, to take this into account. Nevertheless, there is price variation for identical UPCs within all product categories, implying potential gains from price-search.

A category-level quantity index for a consumer i in period t, $\mathrm{QI}^c_{i,t}$, is computed as the value of his or her transactions in a given category relative to the average value across consumers of transactions in the same category. Both values are computed at average prices so that the resulting ratio reflects differences in quantities and not prices. Specifically:

$$\mathrm{QI}^c_{i,t} = \frac{\sum_{w \in t} \sum_{k \in c} \overline{p}^{u,m,w} \times q^{u,m,w}_{i,k}}{\left(\sum_{j \in J^m_t} \sum_{w \in t} \sum_{k \in c} \overline{p}^{u,m,w} \times q^{u,m,w}_{j,k}\right) / J^m_t}\,, \tag{9}$$

where J^m_t is the number of consumers in the panel in market m in period t. The fixed price weights reflect differences in quantity and quality (broadly defined), so purchases of larger amounts of more expensive UPCs have greater weights than purchases of larger amounts of less expensive UPCs. This calculation of quantities purchased aligns with the model, because consumers have a stronger incentive to search for savings on goods that are, on average, more expensive.

Panel B of Figure 1 illustrates the variation in the quantity index (winsorized at the top and bottom 1 percent) by category (ordered by interquartile range). By construction, the mean for the quantity index is (roughly) 1 for each category, and there is substantial variation within each category.

4 Empirical Results

In Table 2, we first show regressions of the form

$$\text{BHI}_{i,t} = \mu_i + \gamma_{m,t} + X_{i,t}\alpha + \epsilon_{i,t} \; ,$$

where $\text{BHI}_{i,t}$ is the bargain hunting index for individual i in quarter t, μ_i is an individual fixed effect (we show results with and without this), $\gamma_{m,t}$ is a market×quarter fixed effect, and X is a vector of regressors: a dummy for age 65 and older, the logarithm of expenditure, the number of shopping trips, and the average weekly number of (different) stores visited. Aguiar and Hurst (2007) use this type of regression, although they do not include the average number of stores visited.

The left panel of Table 2 shows results for regressions without individual fixed effects. The results for expenditure and age confirm previous results. They also confirm the model's predictions that consumers 65 and older find lower prices and that higher-spending consumers pay relatively more. This is consistent with older individuals' having more time to search and working high-wage individuals searching less.

As a direct measure of search effort, we include the average number of stores visited in a week and the number of shopping trips. The former predicts lower prices paid robustly and with high statistical significance. The economic interpretation of the coefficient is that consumers who visit one more store each week (compared with the average) pay almost 2 percent less for identical goods. The inclusion of the average number of stores visited (in a given week) increases the R-square from 0.01 to 0.07, so this variable has much greater explanatory power than does age or expenditure (although this likely reflects that the age dummy is somewhat imprecisely correlated with retirement). Including the number of shopping trips, while omitting the average number of stores visited, gives a highly significant coefficient and an R-square of 0.04. However, including the number of shopping trips with the average number of stores visited turns the coefficient for the number of shopping trips negative without increasing the R-square. One interpretation of this pattern is that some consumers are not able to buy large quantities of non-perishable foods during sales, so they need to take many trips to the same stores. Clearly, it is the average number of stores visited rather than the number or trips that correlates with lower prices.[8]

In the right panel of Table 2, we include individual fixed effects. The R-square jumps to 0.38, so it appears that some consumers are consistently "shoppers" while others are "loyals" (in the parlance of Chevalier and Kashyap 2011). Expenditure and age are still significant. The coefficient on age is now identified only from consumers who turn 65 during the sample period, consistent with a clear effect of retirement on time available for shopping. The number of stores visited is still significant, but the coefficient is half as large as when individual fixed effects are not included, which indicates that some consumers consistently shop in many stores. The number of trips now has a small, positive, and statistically significant coefficient. However, the number of trips and the number of stores visited do not add much to the explanatory power of the regressors.

To further explore how consumers save money, we compute an alternative store BHI, $\mathrm{BHI}^s_{i,t}$, that computes the value of consumer i's basket using the average price of each UPC in a given week *in the store*, s, where the item was purchased, $\bar{p}_s^{u,m,w}$. To compute this average price, we use the store-level dataset.[9]

[8]This is consistent with the findings of Kaplan et al. (2016) that some stores are cheaper for some goods but not for others.

[9]This exercise is performed using data from 2003 through 2007, because store identifiers in the store-level dataset

$$\mathrm{BHI}^{s}_{i,t} = \left(1 - \frac{\sum_{w \in t} \sum_{k=1}^{N_i^w} \overline{p}_s^{u,m,w} \times q_{i,k}^{u,m,w}}{\sum_{w \in t} \sum_{k=1}^{N_i^w} \overline{p}^{u,m,w} \times q_{i,k}^{u,m,w}}\right) \times 100. \tag{10}$$

If the BHI for a consumer is lower than the store BHI, the consumer has, on average, purchased goods at times of the week when the prices of the goods were lower than store-specific weekly average prices. Figure A.2 in the appendix compares the original BHI to the store BHI using histograms. The correlation between the two indices is 0.88, and the histogram in Panel A shows similar distributions of the two bargain hunting indices, suggesting that most gains come from store selection, not from the timing of purchases. In Panel B, we display the histogram of savings by individuals by time of purchase; that is, we show the percentage saved by paying the actual price for each transaction rather than paying the store-weekly average for the relevant UPC. The distribution in Panel B includes many more observations of positive savings than of negative savings, indicating gains from the timing of shopping; however, the mode of the histogram is 1 percent, again indicating small gains from timing of purchases.

The main innovation of this paper is that it examines the relation between quantities by category and prices. We estimate the regression

$$\mathrm{BHI}^{c}_{i,t} = \mu_{i,c} + \gamma_{m,t} + \beta \mathrm{QI}^{c}_{i,t} + X_{i,t}, \alpha + \epsilon_{i,t}\ ,$$

where $\mu_{i,c}$ denotes individual $\times$ category fixed effects, and $\mathrm{QI}^{c}_{i,t}$ the quantity that consumer i consumes from category c compared with the average consumer. These data form a panel indexed by individual $\times$ category and by time, and the coefficient β captures whether consumers find lower prices for the goods in the categories from which they buy relatively more. In some specifications, we exclude individual fixed effects so that $\mu_{i,c} = \mu + \delta_c$, where δ_c represents category dummies.

The results presented in Table 3 show that consumers do indeed pay less for goods of which

are not consistent with identifiers in the panelist dataset after 2007.

they purchase relatively larger quantities, which is consistent with the rational allocation of time across goods categories. The coefficient to the quantity index of 1.17 in column (1) implies that a one standard deviation increase in the quantity index (about 0.74) results in approximately 1 percent savings in the category. The coefficient is estimated with very high precision, with a t-statistic greater than 60. As shown in the left-most columns, without individual fixed effects, age is not significant, whereas expenditure retains the magnitude of the previous set of regressions. The number of stores visited is highly significant, while the number of total trips has a very small coefficient.

Including individual × category fixed effects—see columns (3) and (4) of Table 3—we find coefficients for the quantity index that are nearly the same, implying that consumers search for better prices when they increase the quantities they purchase, because the coefficient now is identified from deviations from the consumer specific mean.[10] The coefficient to expenditure is barely affected by the fixed effect, while the age dummy becomes significant, as it does in Table 2. The effects of the average number of stores visited and the number of store visits are estimated similarly to what was found in Table 2, which is not surprising because the variables are not varying by category. However, the largest boost to the R-square comes from the fixed effects, indicating that the number of trips is not the main avenue to saving; paying better attention to prices and shopping accordingly is.

Columns (5) and (6) of Table 3 revisit the comparison between the original BHI and the store BHI, using the store BHI as the dependent variable in column (6) and the original BHI in column (5). Column (5) repeats the regression reported in column (4) on a smaller sample, because

[10]Strictly speaking, the quantity index for a consumer can change, even if the consumer's purchases are unchanged, if the denominator of average purchases in the category changes. However, consumer-level quantities are much more variable than aggregated quantities.

the comparison of results with the different BHI indices can only be done with pre-2008 data. The restriction is not very important, because the results in column (5) are very similar to those of column (4). The coefficient on the quantity index variable is lower for the store BHI, at 0.9—see column (6)—than for the regular BHI, at 1.3—see column (5)—but the magnitude of the coefficient falls by only 30 percent. Therefore, the interpretation is that 70 percent of savings obtained result from choosing stores with relatively lower prices (by category), and 30 percent come from within-store timing.

The regression results reported in Table 3 are pooled across categories, but pooling may mask differences across categories. As shown in Figure 1, Panel A, the BHI is significantly more compressed for some goods than for other goods, with almost no variation in prices paid for identical categories of beer and little variation for cigarettes (followed by milk and sugar substitutes). Laundry detergent displays the largest variation, followed by hot dogs and mayonnaise. This is not a simple reflection of relative quantities consumed: For the quantity index, blades and ketchup/mustard show the least variation across consumers and carbonated beverages and cigarettes the most.

To test whether our results are robust across categories, we estimate the regression

$$\mathrm{BHI}^c_{i,t} = \mu^c_i + \gamma^c_{m,t} + \beta^c \mathrm{QI}^c_{i,t} + X_{i,t}\alpha^c + \epsilon_{i,t}$$

separately for each category c. The data in each regression form an individual $\times$ time panel, and all coefficients, including dummies, can take different values for the different categories.

Table 4 shows that our main qualitative result is remarkably robust—the coefficient to the quantity index is positive and highly significant in every single category. The sizes of the coefficients to quantity vary (even though the quantity index in these regressions has been standardized to have a mean 0 and a standard deviation of 1 by category for an easier comparison of the coefficients across the 30 regressions). For some categories, including beer, the coefficient is small, which reflects that there is little variation in the prices of those products (for the exact same UPC, that is). In fact, all the categories with a coefficient to quantity that is less than unity are among the categories with the lowest variation in prices paid. The largest coefficient is for photography, one of the categories with the highest price variation. So the variation in coefficients is intuitive, but because our focus is on the qualitative result, we do not model this variation in more detail.

5 Discussion and Conclusion

We find that, consistent with a model of rational price search, consumers pay lower prices for goods of which they consume more, and they pay more for goods of which they consume less. The empirical results provide robust support for the notion that consumers rationally search for best prices. Because we compare prices at the UPC level, our findings provide a lower bound on the savings obtained, but buying different brands and even package sizes to obtain savings brings a potential loss in utility that can be evaluated only by using functional forms, which we avoid in this paper. Our results may be important for determining optimal pricing for stores and producers, but we do not examine this issue.

Our results are consistent with those from models of “shoppers” versus “loyals,” or “cool” versus “busy,” in that we document significant variation in the prices consumers pay on average. In Table A.2 in the appendix, we illustrate the magnitudes of savings. The top quarter of consumers in the BHI distribution (the cool shoppers) pay, on average, 7.63 percent less than the average consumer, whereas consumers in the bottom quarter (the busy loyals) pay 6.67 percent more for the exact same goods. For each consumer, we also rank his or her purchased categories according to the quantity index, and then we divide the goods into top-half and bottom-half categories (for this exercise, we include only consumers who purchase at least two categories).[11] We then compute separate BHIs for top-half and bottom-half categories (in terms of the quantity index). On average, consumers save 0.44 percent on the goods of which they buy more (relative to other consumers) and pay 1.18 percent more for the goods of which they buys less. Inattentive consumers, in terms of the overall BHI, pay more for all goods, and attentive consumers realize the majority of their savings on goods of which they purchase a lot.

Overall, there is substantial heterogeneity across consumers, as previously documented. Our contribution is to document how rational consumers shop across goods and conduct more price-searching for the most desired goods.

[11] If the number of categories is not even, the top group has one more category.

References

Aguiar, Mark, and Erik Hurst. 2007. "Life-Cycle Prices and Production." *American Economic Review* 97(5): 1533–1559.

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis. 2013. "Time Use during the Great Recession." *American Economic Review* 103(5): 1664–1696.

Becker, Gary. 1965. "A Theory of the Allocation of Time." *Economic Journal* 75(299): 493–508.

Benhabib, Jess, Richard Rogerson, and Randall Wright. 1991. "Homework in Macroeconomics: Household Production and Aggregate Fluctuations." *Journal of Political Economy* 99(6): 1166–1187.

Bronnenberg, Bart J., Michael W. Kruger, and Carl F. Mela. 2008. "The IRI Marketing Data Set." *Marketing Science* 27(4): 745–748.

Chevalier, Judith A., and Anil K Kashyap. 2011. "Best Prices." NBER Working Paper No. 16680. Cambridge, MA: National Bureau of Economic Research, Inc.

Coibion, Olivier, Yuriy Gorodnichenko, and Gee Hee Hong. 2015. "The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications." *American Economic Review* 105(3): 993–1029.

Greenwood, Jeremy, and Zvi Hercowitz. 1991. "The Allocation of Capital and Time over the Business Cycle." *Journal of Political Economy* 99(6): 1188–1214.

Griffith, Rachel, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo. 2009. "Consumer Shopping Behavior: How Much Do Consumers Save?" *Journal of Economic Perspectives* 23(2): 99–120.

Kaplan, Greg, and Guido Menzio. 2013. "The Research Agenda: Greg Kaplan and Guido Menzio on the Macroeconomics of Bargain Hunting." *Economic Dynamics Newsletter* 14(2).

Kaplan, Greg, Guido Menzio, Leena Rudanko, and Nicholas Trachter. 2016. "Relative Price Dispersion: Evidence and Theory." NBER Working Paper No. 21931. Cambridge, MA: National Bureau of Economic Research, Inc.

Nevo, Aviv, and Arlene Wong. Forthcoming. "The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession." *International Economic Review.*

Reis, Ricardo. 2006. "Inattentive Consumers." *Journal of Monetary Economics* 53(8): 1761–1800.

Sims, Christopher A. 2010. "Rational Inattention and Monetary Economics." In *Handbook of Monetary Economics Vol. 3*, eds. Benjamin M. Friedman and Michael Woodford. Amsterdam, The Netherlands: Elsevier.

Stigler, George J. 1961. "The Economics of Information." *Journal of Political Economy* 69(3): 213–225.

Stroebel, Johannes, and Joseph Vavra. 2014. "House Prices, Local Demand, and Retail Prices." NBER Working Paper No. 20710. Cambridge, MA: National Bureau of Economic Research, Inc.

Table 1. Summary Statistics for Regressions

	Count	Mean	SD	Min	Max
Bargain Hunting Index (BHI)	192,708	5.60	6.11	–10	23
BHI (demeaned)	192,708	–0.00	5.62	–21	21
BHI, age 65+	67,289	0.51	5.94	–21	21
BHI, age<65	125,419	–0.27	5.42	–21	21
BHI, exp. $<$ median exp.	96,335	0.42	6.15	–21	21
BHI, exp. $\geq$ median exp.	96,373	–0.42	5.00	–21	21
Category-Specific BHI	1,627,149	0.00	10.61	–64	47
Category-Specific Quantity Index	1,627,149	0.98	0.74	0	4
Expenditure (quarterly)	192,708	201	145	10	3,972
Old (65+)	192,708	0.35	0.48	0	1
No. trips to store (quarterly)	192,708	26	18	1	252
No. $\neq$ stores visited (weekly average)	192,708	1.65	0.73	1	11

Notes: Authors' calculations using all IRI panelist data from 2003 through 2012. The BHI computation is described in equation (7). The index measures how much a consumer saves (positive values), in percent, or overpays (negative values) relative to buying his or her consumption bundle at average prices. The BHI is broken up by age group and expenditure group. The category-specific BHI is described in equation (7) and focuses on savings in a specific category. The category-specific quantity index, which measures whether a consumer purchases more or less of that category than does the average consumer, is computed according to equation (9). The other variables are used in our regressions: (1) Expenditure is total dollars spent in a given quarter by a panelist in IRI transactions; (2) Old (65+) is a dummy variable for whether consumers are 65 or older; (3) No. trips to store (quarterly) is the total number of trips to stores by a panelist in a given quarter; (4) No. $\neq$ stores visited (weekly average) is the weekly average number of different stores that a consumer visits in a given quarter.

Table 2. The Bargain Hunting Index for Overall Expenditure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log. Expenditure	–0.61***	–0.92***	–1.03***	–0.84***	–0.58***	–0.75***	–0.83***	–0.79***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
Old (65+)	0.54***	0.11	0.21**	0.14*	0.35***	0.29**	0.32**	0.29**
	(0.09)	(0.08)	(0.09)	(0.08)	(0.13)	(0.13)	(0.13)	(0.13)
# stores visited (weekly avg.)		1.89***		2.16***		0.83***		0.73***
		(0.06)		(0.09)		(0.05)		(0.06)
# trips (quarterly)			0.06***	–0.02***			0.03***	0.01***
			(0.00)	(0.00)			(0.00)	(0.00)
Quarter-Year × Market FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	No	No	No	Yes	Yes	Yes	Yes
Observations	192708	192708	192708	192708	192705	192705	192705	192705
Adj. R-squared.	0.01	0.07	0.04	0.07	0.38	0.38	0.38	0.38

Notes: Regression: $\text{BHI}_{i,t} = \mu_i + \gamma_{m,t} + X_{i,t}\,\alpha + \epsilon_{i,t}$, where $\text{BHI}_{i,t}$ is the bargain hunting index for individual i in quarter t, μ_i is an individual fixed effect (FE) ($\mu_i = \mu$ in the first four columns), $\gamma_{m,t}$ is a market×quarter FE, and X is a vector of regressors: a dummy for age 65 and older, the logarithm of total expenditure, the average weekly number of different stores visited, and the total number of shopping trips in the quarter. Standard errors clustered by individual. *** (**) [*] significant at the 1 (5) [10] percent level.

Table 3. Rational Inattention. Pooled Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Quantity Index	1.17***	1.18***	1.08***	1.27***	1.30***	0.90***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Log. Expenditure	–1.01***	–0.96***	–0.93***	–1.15***	–1.36***	–1.01***
	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)
Old (65+)	–0.03	–0.01	0.20*	0.22*	–0.14	0.14
	(0.08)	(0.08)	(0.11)	(0.11)	(0.34)	(0.30)
# stores visited (weekly avg.)	1.66***	1.86***	0.74***	0.64***	1.04***	0.83***
	(0.06)	(0.08)	(0.04)	(0.05)	(0.07)	(0.07)
# trips (quarterly)		–0.01***		0.00**	0.00	0.00
		(0.00)		(0.00)	(0.00)	(0.00)
Quarter-Year × Market FE	Yes	Yes	Yes	Yes	Yes	Yes
Category FE	Yes	Yes	No	No	No	No
Category × Individual FE	No	No	Yes	Yes	Yes	Yes
Observations	1627149	1627149	1627148	1597418	877744	877744
Adj. R-squared.	0.03	0.03	0.10	0.18	0.20	0.25

Notes: Regression for columns (1) and (2): $\mathrm{BHI}^c_{i,t} = \mu + \delta_c + \gamma_{m,t} + \beta \mathrm{QI}^c_{i,t} + X_{i,t}, \alpha + \epsilon_{it}$, where $\mathrm{BHI}^c_{i,t}$ is the category-specific bargain hunting index for individual i in quarter t, δ_c is a category FE, $\gamma_{m,t}$ is a market×quarter FE, X is a vector of regressors, and $QI^c_{i,t}$ is the quantity index described in equation (9). Regression for columna (3)–(6): $\mathrm{BHI}^c_{i,t} = \mu_{i,c} + \gamma_{m,t} + \beta \mathrm{QI}^c_{i,t} + X_{i,t}, \alpha + \epsilon_{it}$, where $\mu_{i,c}$ denotes individual× category fixed effects. The quantity index, which measures whether a consumer purchases more or less of a category than the average consumer, is standardized (mean 0, sd 1) for easier interpretation. Standard errors clustered by individual. *** (**) [*] significant at the 1 (5) [10] percent level.

Table 4. The BHI and the QI by Category. Separate Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Quantity Index	0.07***	1.35***	1.25***	0.33***	1.08***	1.09***	2.65***	0.53***	1.58***	0.84***
	(0.02)	(0.15)	(0.06)	(0.08)	(0.06)	(0.05)	(0.17)	(0.12)	(0.08)	(0.05)
Log. Expenditure	–0.06	–1.24***	–1.76***	–0.29*	–1.18***	–1.30***	–1.02***	–0.67**	–0.93***	–1.35***
	(0.04)	(0.36)	(0.09)	(0.15)	(0.13)	(0.10)	(0.38)	(0.31)	(0.16)	(0.12)
Old (65+)	–0.01	0.09	–0.05	0.24	0.31	0.34	0.51	0.01	–0.05	0.71**
	(0.11)	(0.76)	(0.26)	(0.56)	(0.31)	(0.28)	(0.91)	(0.66)	(0.37)	(0.33)
# stores (weekly avg.)	0.06	–0.02	0.80***	0.47***	0.56***	0.75***	0.61	–0.65*	0.51***	0.68***
	(0.05)	(0.37)	(0.11)	(0.16)	(0.16)	(0.13)	(0.42)	(0.36)	(0.18)	(0.14)
# trips (quarterly)	–0.00	0.01	–0.00	0.01	0.01**	0.00	0.01	0.03**	0.00	0.01***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Observations	49979	10718	145385	11985	75971	125202	13973	7450	45606	77531
Adj. R-squared.	0.07	0.12	0.17	0.28	0.18	0.22	0.17	0.13	0.21	0.17

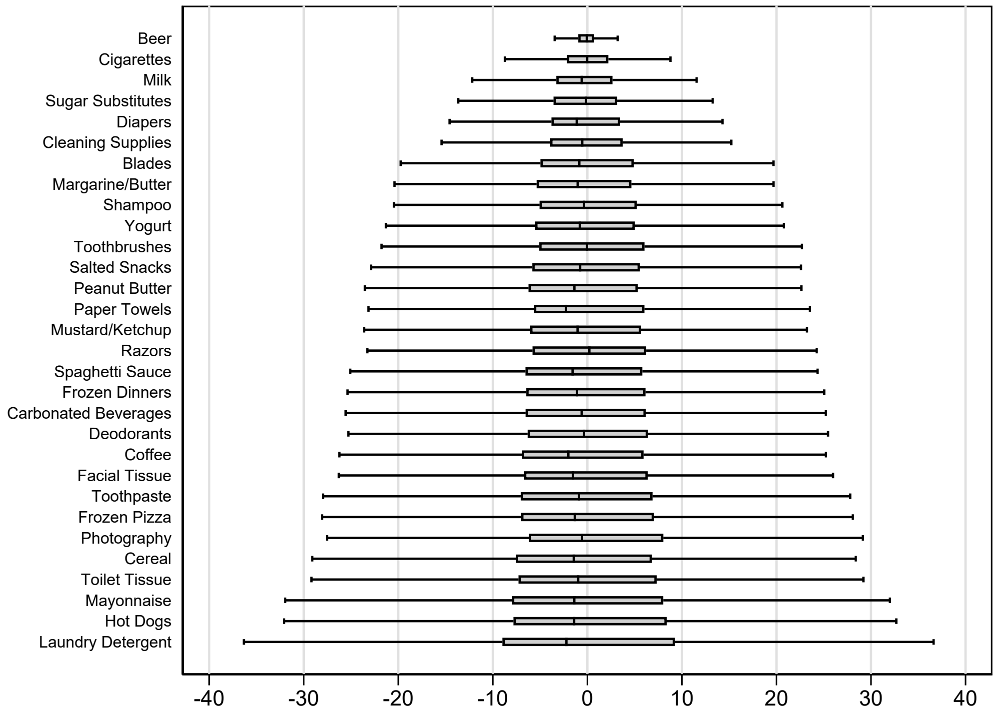
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Quantity Index	1.17***	0.64***	2.58***	1.71***	1.13***	1.94***	0.37***	2.19***	1.43***	1.81***
	(0.06)	(0.08)	(0.10)	(0.08)	(0.07)	(0.09)	(0.03)	(0.11)	(0.08)	(0.09)
Log. Expenditure	–1.03***	–0.57***	–1.39***	–1.92***	–0.79***	–1.02***	–0.39***	–0.81***	–1.48***	–0.55***
	(0.14)	(0.18)	(0.22)	(0.17)	(0.13)	(0.20)	(0.05)	(0.24)	(0.15)	(0.19)
Old (65+)	0.13	0.34	0.06	–0.13	0.08	0.41	0.28*	–0.47	0.39	0.51
	(0.37)	(0.40)	(0.57)	(0.44)	(0.36)	(0.49)	(0.15)	(0.56)	(0.34)	(0.50)
# stores (weekly avg.)	0.57***	0.44*	0.75***	1.17***	0.45***	0.36	0.34***	0.61**	0.64***	0.34
	(0.16)	(0.24)	(0.26)	(0.21)	(0.16)	(0.25)	(0.07)	(0.30)	(0.19)	(0.22)
# trips (quarterly)	0.01	0.01	0.01	0.02**	0.00	0.02*	0.00	–0.01	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
Observations	78471	17312	48682	76342	53258	42804	158984	21727	58140	35139
Adj. R-squared.	0.13	0.19	0.15	0.18	0.22	0.16	0.21	0.21	0.14	0.26

	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
Quantity Index	2.88***	0.99***	1.33***	1.31***	1.12***	0.95***	1.46***	2.00***	2.02***	0.57***
	(0.38)	(0.35)	(0.05)	(0.12)	(0.07)	(0.13)	(0.05)	(0.22)	(0.11)	(0.05)
Log. Expenditure	–1.61*	–1.29	–0.79***	–0.90***	–0.56***	–0.46*	–1.64***	–0.03	–1.10***	–0.49***
	(0.86)	(1.00)	(0.08)	(0.27)	(0.14)	(0.28)	(0.11)	(0.51)	(0.26)	(0.10)
Old (65+)	–3.47*	–1.25	0.42*	0.40	0.09	0.61	0.28	–1.22	0.03	0.34
	(2.07)	(2.81)	(0.24)	(0.70)	(0.43)	(0.76)	(0.28)	(1.15)	(0.59)	(0.28)
# stores (weekly avg.)	–0.92	1.49*	0.85***	0.33	1.21***	–0.46	0.66***	0.25	0.68**	0.80***
	(1.01)	(0.85)	(0.11)	(0.31)	(0.19)	(0.34)	(0.13)	(0.50)	(0.29)	(0.13)
# trips (quarterly)	0.00	–0.00	–0.01***	0.01	–0.00	0.01	0.01*	0.03	0.02*	–0.00
	(0.04)	(0.04)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Observations	2715	1732	139666	16071	54645	7183	102647	7441	23042	87617
Adj. R-squared.	0.17	0.06	0.20	0.14	0.19	0.23	0.19	0.12	0.21	0.17

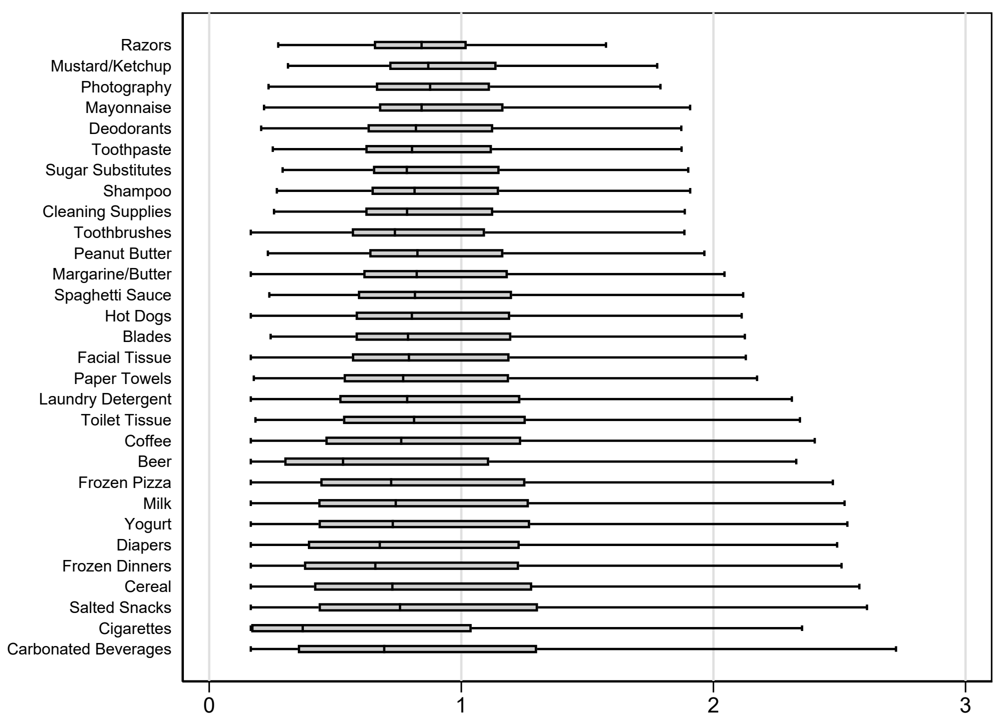
Notes: Regression: $\text{BHI}^c_{i,t} = \mu^c_i + \gamma^c_{m,t} + \beta^c \text{QI}^c_{i,t} + X_{i,t}\alpha^c + \epsilon_{i,t}$, estimated category by category. The quantity index, $QI^c_{i,t}$, is standardized by category (mean 0, sd 1) for easier interpretation. All regressions include quarter×market FE and individual FE. Categories as follows: (1) beer, (2) blades, (3) carbonated beverages, (4) cigarettes, (5) coffee, (6) cold cereal, (7) deodorants, (8) diapers, (9) facial tissue, (10) frozen dinners, (11) frozen pizza, (12) cleaning supplies, (13) hot dogs, (14) laundry detergent, (15) margarine/butter, (16) mayonnaise, (17) milk, (18) mustard/ketchup, (19) paper towels, (20) peanut butter, (21) photography, (22) razors, (23) salted snacks, (24) shampoo, (25) spaghetti sauce, (26) sugar substitutes, (27) toilette tissue, (28) toothbrushes, (29) toothpaste, (30) yogurt. Standard errors clustered by panelist. *** (**) [*] significant at the 1 (5) [10] percent level.

Figure 1. Variation by Category

Panel A: The Bargain Hunting Index



Panel B: The Quantity Index



Notes: Product categories are defined by IRI. Panel A shows variation in the bargain hunting index (BHI). The left and right borders of each category box depict the 75th and 25th percentiles of the BHI for that category, while the whiskers represent upper and lower adjacent values (outside values not plotted). The categories have been sorted by the interquartile range. Panel B depicts variation in the quantity index by category and is created analogously.

Online Appendix

Data

In our dataset, age is reported in categories, and the age distribution by category is as follows: 10 percent are younger than 45 years old; 25 percent are aged 45 to 54; 21.7 percent are aged 55 to 64; 22 percent are 65 or older; and 21 percent are unclassified. Household income is reported by category: 12.5 percent have income that is less than $20,000; 20.3 percent earn $20,000 to $35,000; 27.4 percent earn $35,000 to $55,000; 19.2 percent earn $55,000 to $75,000; 12.7 percent earn $75,000 to $100,000; and 7.8 percent have income that is more than $100,000. Education categories have the distribution: 35.7 percent of panelists have not completed high school; 18.9 percent have at least graduated from college; and the rest are high school graduates. Relative to the U.S. population, the IRI sample is somewhat older and poorer, and spending in the IRI categories represents roughly 20 percent of PSID food-at-home expenditure.

Table A.1. Summary Statistics for Panelists in 2007, quarter 1.

	Count	Mean	SD	Min	Max
Years of Education	4,555	13.76	2.01	6	18
Age	4,867	55.43	12.72	21	70
Household Income	4,865	52,302	36,608	5,000	150,000
Old (65+)	4,867	0.32	0.46	0	1
Retired	4,867	0.27	0.44	0	1
Expenditure (quarterly)	4,867	194	152	10	2,765

Notes: Authors' calculations using all IRI panelist data for the first quarter of 2007.

Table A.2. Average Values of BHI Within Quarters of its Distribution

	(1)	(2)	(3)	(4)	(5)
		BHI in Group			
	All	Q1	Q2	Q3	Q4
Mean BHI	0.00	7.63	1.28	–2.31	–6.67
Mean BHI top-half categories	0.44	8.93	1.69	–2.14	–6.81
Mean BHI bottom-half categories	–1.18	4.34	0.11	–2.66	–6.59

Notes: The table displays in the first row the overall average value of the bargain hunting index (BHI), normalized to be 0, and the average value for the quarter of consumers with the highest, second-highest, second-lowest, and lowest value of the BHI. For each consumer, we rank his or her purchased categories in terms of the quantity index and divide them into top-half and bottom-half (we include only consumers who purchase from at least two categories). We then compute two BHIs for top-half and bottom-half categories separately. The average values of these two BHIs are presented in the second and third rows of the table, first for all consumers in column (1), and then by quarter-group of the overall BHI in columns (2) through (5).

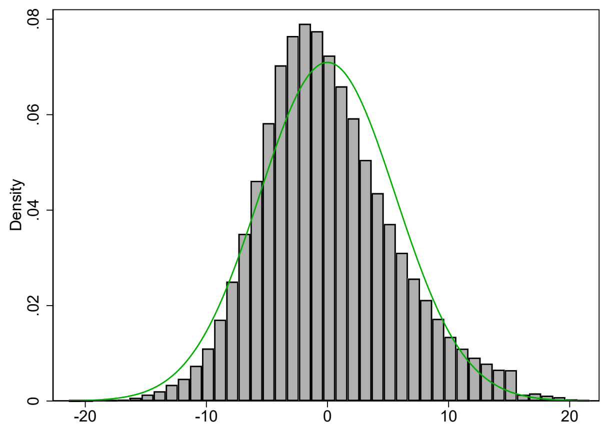
Additional Figures

In Figure A.1, we use a histogram to display the dispersion of the (demeaned) overall BHI. The BHI is slightly leptokurtic (kurtosis is 3.3) and skewed to the right (skewness is .43). The bottom two panels split the sample into 65-plus and younger panelists, and into panelists with below- and above-median expenditure in a given period. As our model predicts, the older individuals pay lower prices on average than do the younger panelists, and the poorer panelists (as proxied by expenditure) also pay relatively lower prices.

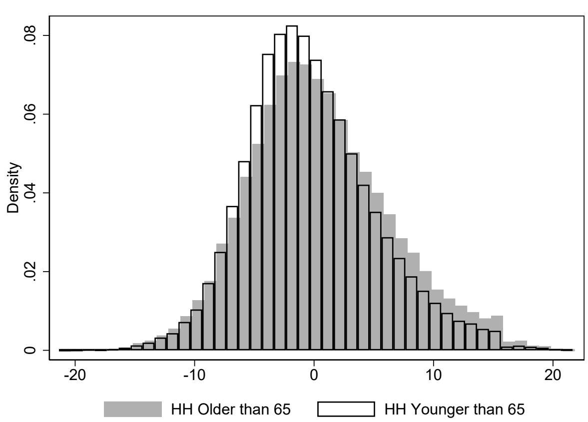
Figure A.2 depicts histograms for the overall BHI and the store BHI. The histograms are very similar, which indicates that savings from the timing of purchases within the same store are small: A large fraction of the savings (dis-savings) occurs from consumers' purchasing products in stores where they are relatively cheaper (more expensive).

Figure A.1. The Bargain Hunting Index

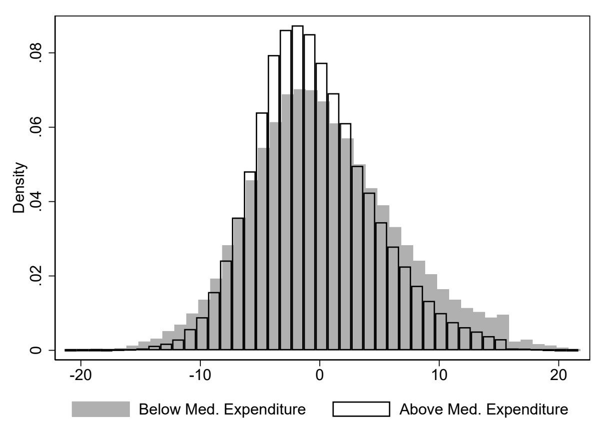
Panel A: Overall



Panel B: By Age Group



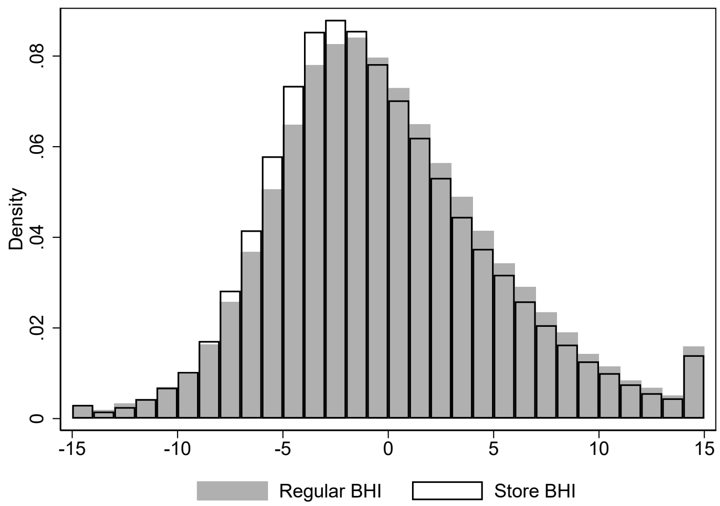
Panel C: By Expenditure Group



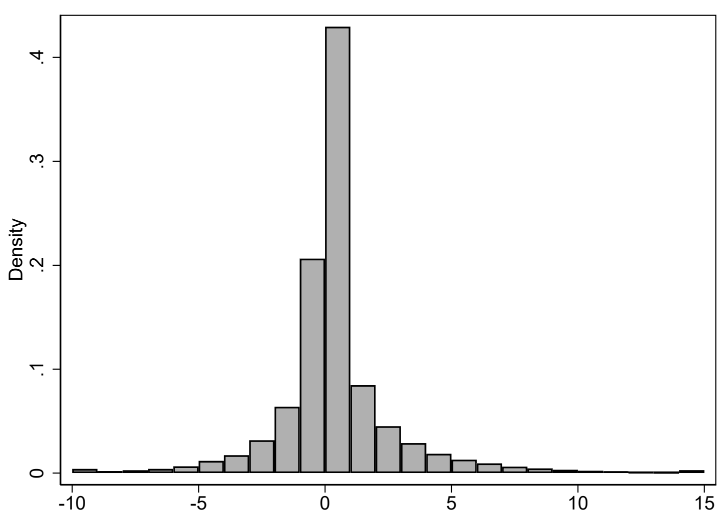
Notes: The BHI index shows how much a consumer saves in percentages compared with the counterfactual of buying his or her consumption bundle at average prices. The BHI index has been normalized to have a mean of 0 every quarter-year by market. *Source:* IRI, all panelist data from 2003 through 2012.

Figure A.2. The BHI vs. the Store BHI

Panel A: Comparing the Indices



Panel B: The Difference between the Indices



Notes: The regular BHI index is defined in equation 7 and represents how much a consumer saves relative to buying at average prices across stores. The store BHI is defined in equation 10 and measures how much a consumer would save if he or she paid average prices in the store relative to buying the consumption bundle at average prices across all stores. Panel B plots the distribution of the difference between the indices (individual by individual). *Source:* IRI, all panelist data from 2003 through 2012.