Do Informed Consumers Pay Less?

Evidence from a Survey with Linked Grocery Purchase Data*

Richard Friberg† Emil M. S. Halseth‡

Frode Steen§ Simen A. Ulsaker¶

July 17, 2026

Abstract

We examine how consumer price knowledge affects prices paid in grocery markets, combining survey-based price recall data from over 2000 Norwegian households with 18 months of linked transaction histories. Better-informed consumers pay lower prices by timing purchases to coincide with sales: a 10 percentage-point increase in price knowledge is associated with a 1.3 percentage reduction in prices paid. This provides direct empirical support for the central mechanism in Varian's (1980) model of sales. Consumers who actively seek price information have higher price knowledge, suggesting that policies improving price accessibility may improve consumers' ability to identify and act on temporary discounts.

Keywords: Price Knowledge, Price Learning Effort, Price Competition, Discount Utilization, Search Costs

JEL Codes: D83, L10, L66

*We would like to thank Richard Gilbert, Christos Genakos, Ben Handel, Aljoscha Janssen, Eszter Kabos, Aviv Nevo, Patrick Rey, Philipp Schmidt-Dengler, Carl Shapiro, Lars Sørgard and Otto Toivanen for their helpful comments. We also thank seminar participants at Berkeley NHH Industrial Organization Conference, CRESSE, MACCI summer institute, NORIO and Sveriges Riksbank. We are grateful to NorgesGruppen for data. Financial support from the FOOD research project (https://www.nhh.no/en/research-centres/food/) is gratefully acknowledged.

†Stockholm School of Economics, Norwegian School of Economics and CEPR; richard.friberg@hhs.se

‡Oslo Economics and Norwegian School of Economics; emil.halseth@nhh.no

§Norwegian School of Economics and CEPR; frode.steen@nhh.no

¶OsloMet and Norwegian School of Economics; simen.ulsaker@oslomet.no

1 Introduction

Consumer knowledge of prices across retailers is crucial for effective market competition. As emphasized in Stigler (1961), acquiring price information requires effort, as consumers must allocate attention and often engage in various forms of search. The prevailing intuition, both among economists and policymakers, is that reducing search costs leads to more informed consumers, thereby intensifying competition and driving prices down.

A central modeling approach in the literature on costly search is to distinguish between two types of consumers—one with zero search costs ("informed") and one with some positive search cost ("uninformed")—see, for example, Salop and Stiglitz (1977), Varian (1980), and Stahl (1989).[1] These models predict that a higher share of informed consumers intensifies competition and drives prices down—a result that has become central to economic research on search costs and market outcomes. Such models have, for instance, been used to generate predictions for the effects of mandatory price disclosure in grocery markets (Ater and Rigbi, 2023), price dispersion in gasoline retailing (Pennerstorfer et al., 2020), the role of temporary sales in grocery markets (Villas-Boas, 1995), and the impact of reduced search costs as consumers gain access to the internet (Brown and Goolsbee, 2002). They have also been tested in laboratory settings to better understand price dispersion and consumer behavior (Morgan et al., 2006).[2]

While theory predicts that lower search costs lead to more informed consumers and stronger price competition, empirical evidence on how price knowledge affects actual consumer outcomes remains limited. This paper provides direct evidence on how consumer price knowledge affects prices paid in the grocery market by asking: do more informed consumers systematically pay lower prices? Using an online survey, we measure consumers' knowledge of grocery prices and link their responses to 18 months of individual-level purchasing history from Norway's largest grocery retail group. This unique combination of survey data and purchasing histories allows us to test the prediction that informed consumers pay lower prices and investigate whether price knowledge is associated with particular shopping behaviors, attitudes, or demographics. The final sample includes more than 2000 households and over 70 000 recalled prices across two grocery chains and 24 products.

Theoretical models such as Varian (1980) predict that when consumers become more

[1]While the intuition that lower search costs intensify competition is attractive, models that feature this outcome are surprisingly intricate. Indeed, the formal literature on the interaction between search costs and competition started with a paradox—the result in Diamond (1971) that introducing even minuscule search costs into a model of Bertrand competition with homogeneous products leads all firms to charge monopoly prices instead of pricing at marginal cost.

[2]See also empirical work documenting that patterns of price dispersion match predictions from search models, including Sorensen (2000) and, for a survey, Honka et al. (2019). Relatedly, structural models have been used to recover the distribution of search costs from observed prices and quantities (see, e.g., Hong and Shum, 2006).

informed, firms have stronger incentives to lower prices through temporary discounts and promotions to attract price-sensitive shoppers. A key assumption behind this prediction is that informed consumers exploit price variation by timing purchases to coincide with sales rather than paying regular prices. To our knowledge, we are the first to provide direct empirical support for this mechanism, namely that consumers with higher price knowledge systematically pay lower prices, not because they buy inherently cheaper products or shop at inherently cheaper stores, but because they are more likely to purchase products when the price is low. This suggests that informed consumers impose a competitive constraint on retailers by responding to price fluctuations, reinforcing the theoretical link between consumer information and price competition. The effects we find are economically significant: for instance, moving from relatively low price knowledge (an average percentage absolute deviation between recalled and true prices of 40%) to relatively high price knowledge (an average percentage absolute deviation between recalled and true prices of 10%) is associated with a 4 percentage-point decrease in the average price paid. This relationship holds even after controlling for demographic characteristics, shopping behavior, and stated price consciousness.

We further show that the returns to price knowledge are heterogeneous. The reduction in prices paid is significantly larger for consumers classified as price conscious, reflecting a higher motivation to act on price knowledge. Consistent with Stigler's (1961) prediction that the returns to search increase with price dispersion, we find that the effect is amplified for shopping baskets characterized by higher price volatility. We also observe stronger returns for non-perishable products, where the scope for intertemporal substitution is greater due to the possibility of stockpiling (Hendel and Nevo, 2006, 2013).

Finally, we assess the yearly monetary savings a nuclear family can expect on a representative shopping basket. These estimates capture only direct savings, disregarding the additional competitive pressure that a larger share of informed consumers would exert on equilibrium prices. We find that a 30-percentage-point increase in price knowledge is associated with meaningful but modest average savings of 2.1 percent (3629 NOK)—partly because high-spending products in our sample exhibit less price volatility—but the savings increase to 2.6 percent (4428 NOK) for consumers who purchase products with high price fluctuations.

Having established the relationship between price knowledge and prices paid, we go on to investigate the drivers of price knowledge. We find that, overall, product characteristics and observed consumer behaviors are more important than demographic characteristics and other self-reported variables in explaining consumers' price knowledge. For example, a high number of substitutes strongly predicts poorer price recall accuracy, while prior purchasing experience and spending on a product are associated with higher recall. Still, we also find that consumers who are more active in gathering information about prices—by comparing prices in stores, reading flyers, or consulting price comparison surveys in the media—have higher price knowledge. This association also holds after controlling for differences in demographics, shopping behavior, and consumers' reported price sensitivity, that is, how important prices are in purchasing decisions.

Our findings have implications for policy design. First, the relationship between price knowledge and the prices consumers pay suggests that policies that successfully raise consumers' price knowledge can lead to lower prices and stronger competition. Second, we find that consumers who actively engage in price-gathering activities tend to be more knowledgeable. This matters because if price knowledge were solely driven by traits such as cognitive ability or shopping patterns, reducing search costs through better access to price information might have limited effect. Instead, our results indicate that price knowledge is shaped by access to information and that interventions aimed at improving the accessibility and visibility of price information—such as price disclosure regulations or comparison tools—may be effective in raising consumer knowledge and affecting market outcomes.[3]

These policy implications relate closely to a large empirical literature studying the effects of price transparency regulations, which are often motivated by the prediction that better-informed consumers intensify competition and discipline prices.

For instance, Ater and Rigbi (2023) study a price transparency regulation in the Israeli food retail industry that required supermarkets to report daily prices to an online repository used by price-comparison websites, and find that average prices declined by 4 to 5 percent following the reform. In a different context, Brown (2019) analyze the introduction of a price transparency website for medical imaging procedures and find that prices fell due to both demand-side responses (patients switching to lower-cost providers) and supply-side responses (providers lowering prices). Studies of price transparency in gasoline markets find more mixed effects (Rossi and Chintagunta, 2016; Luco, 2019; Montag et al., 2024).

While this line of literature typically focuses on market-level outcomes—such as reduced average prices or lower price dispersion—our study complements it by providing direct micro-level evidence for the link between price knowledge and prices paid.

Clearly, the role of price knowledge may differ between product markets: compare, for instance, one-time, long-term purchases (e.g., pension savings plans), simple products with little physical differentiation (e.g., gasoline), and markets where consumers regularly purchase many differentiated products and must consider an entire basket of prices (e.g., grocery retailing). Given the large share of household budgets devoted to groceries, grocery retailing is an especially important setting. The focus on temporary sales as a source of price variation is highly relevant in retail grocery markets. Anderson et al. (2017), for instance, document that 95% of price changes in a large US supermarket chain are due

[3]Of course, such comparison tools may also have unintended effects, such as making it easier for competitors to learn about each other's prices.

to temporary sales. Clearly, we cannot be certain whether the magnitude of the effects translates to markets other than groceries, but we expect the fundamental mechanism at play—more informed consumers taking greater advantage of temporary sales—to manifest itself in other markets as well. Nakamura and Steinsson (2008), for instance, show the importance of temporary sales in a broader sample using product-level price data gathered by the Bureau of Labor Statistics to construct the US Consumer Price Index.[4]

While public policy is a key motivation for our interest in the links between price knowledge and discount utilization, we note that the interaction between price knowledge and behavior is also of interest from other perspectives. Firms may act to affect the amount of information that consumers have—for instance via informative advertising (Milyo and Waldfogel, 1999), obfuscation (Ellison and Ellison, 2009; Ellison and Wolitzky, 2012), by "educating" consumers about competitors' price obfuscation (Gabaix and Laibson, 2006), or via platform design (Dinerstein et al., 2018). Again, our results buttress the fundamental mechanism at play in such models.

We also contribute to the marketing literature on price knowledge and consumer behavior. Prior studies in this field have primarily measured price recall and examined its implications for pricing strategies and the marketing mix (Winer, 1986; Dickson and Sawyer, 1990; Loy et al., 2020), but have generally lacked data linking recall to actual purchase behavior. We address this gap by systematically combining survey-based price knowledge data with transaction histories from the same consumers. Our results show that much of the variation in price knowledge is explained by differences in shopping behavior, highlighting the value of behavioral data for understanding who is more or less informed.

The remainder of the paper is organized as follows. Section 2 introduces the data and outlines the empirical framework. Section 3 examines the relationship between price knowledge and discount utilization. Section 4 investigates the determinants of price knowledge. Section 5 concludes.

2 Data and Empirical Framework

2.1 Data Sources

We use four primary sources of data. The first is a digital survey, distributed with support from NorgesGruppen—Norway's largest grocery umbrella chain—to 50 000 members of the customer program Trumf, of whom 2505 responded.[5] After being asked about

[4] Bronnenberg et al. (2013) study a different margin through which informed consumers pay lower prices. They show that medical professionals are more likely than other consumers to purchase lower-priced generic pharmaceuticals rather than branded alternatives.

[5] Trumf membership provides a baseline discount of 1% on all purchases made in the umbrella chain in addition to other product-specific discounts. In contrast to many other membership programs, discounts

demographics, the respondents were asked to recall the current shelf price, accounting for sales but ignoring individual discounts, for 18 out of 24 representative grocery products at two grocery chains belonging to the umbrella chain: a leading discount chain and a leading supermarket chain in Norway.[6] The selected products cover the major categories of groceries and should be familiar to the vast majority of grocery shoppers in Norway. Finally, they were asked to state their degree of agreement with 10 different statements designed to elicit their subjective degree of price consciousness. The complete survey and more details can be found in Appendix C. Out of the 2505 respondents, 2285 made at least one price recall, resulting in 77 646 price recall responses. Note that there is one unique recall for every individual × chain × product combination, whereas the other survey questions only vary across individuals. The data are unbalanced primarily because each respondent was only asked about 18 out of the 24 products included in the survey, but also because of some missing responses.

The second data source comprises the daily purchasing history of each respondent at the grocery umbrella chain, derived from membership program data. Covering the period from 1 January 2023 to 12 June 2024, this dataset captures spending and quantities for the products in the survey and their close substitutes—a total of more than 700 items. The dimensions of the daily transaction data are individual × chain × product × date, and the data are highly unbalanced since most people do not buy every product, every day, at both chains.

The third data source is weekly purchasing frequency, spending, and the number of unique products purchased at both grocery chains, including products not included in the daily data. The weekly purchasing data have the following dimensions: individual × chain × week.

The fourth data source consists of the modal (most frequently paid) price for each chain × product × municipality × date combination for the products included in the daily transaction data, which almost always corresponds to the shelf price (hereafter denoted as P^{Shelf}).

In order to relate the sales and price data to the survey data, we aggregate various variables based on the purchasing histories, such as spending, purchasing frequencies, and prices paid, across time and product categories before merging. When appropriate, we further aggregate the data to the individual level, for instance by computing the average error rates for every chain × product combination for which the individual recalled prices, or by calculating the individual's overall discount utilization. Similarly, we calculate average price recall error rates for the different chains and products to see how the

in Trumf are not progressive, *i.e.*, the rebate level is not dependent on earlier purchases. The program therefore has no direct loyalty effect through consumer lock-in, which is also the case for the membership programs of the competing chains.

[6] By asking for specific numbers, we situate ourselves in the part of the price-knowledge literature focusing on explicit (rather than implicit) price knowledge (see Evanschitzky et al., 2004).

respondents' price knowledge varies across them, and whether they tend to over- or underestimate the prices of certain chains or products.

For comprehensive descriptive statistics and details on data cleaning, we refer the reader to Appendix A.[7] Here, to assess the external validity of our results, we focus on the composition of the survey sample. The survey is based on a random sample of Trumf members. Trumf is the largest membership program in the Norwegian grocery industry, with nearly 3 million members from an adult population of 4.5 million (about 80% of Norway's total population of 5.6 million is over 18 years old; see Statistics Norway, 2024a), suggesting that Trumf members are broadly representative of the Norwegian population.[8] Our survey sample of 2505 respondents is also broadly representative in terms of demographics, albeit with some noteworthy deviations. Compared to the general adult population in 2024, respondents are more likely to have higher education (71% vs. 38%), to be over 40 years of age (77% vs. 64%), and be female (62% vs. 50%).[9] The overrepresentation of highly educated respondents aligns with documented patterns in European survey data, where participation rates typically correlate positively with educational attainment (Spitzer, 2020). The older age profile may reflect that retirees have more time to respond to surveys, and the overrepresentation of women may reflect, on average, greater interest in grocery shopping and prices, consistent with women spending more time on household work—including purchasing goods and services (Statistics Norway, 2024c)—and with women more frequently being the household's main grocery shopper.[10] Including demographic controls in the regressions helps address these imbalances.

2.2 Discussion of Hypotheses and Empirical Framework

Conceptually, we can think of two sources of price dispersion: across retailers and within retailers over time. The latter is the focus of Varian (1980), who shows how a mixed-strategy equilibrium with random occurrences of low-price periods (temporary sales) allows retailers to increase profits by attracting informed shoppers, in essence, trading off incentives to attract informed consumers against incentives to extract surplus from uninformed shoppers. Models similar to Varian's remain central to the modeling of information frictions and search behavior, and recent examples in this class include Fabra

[7]The cleaning process involves removing extreme outliers in recall accuracy. In Appendices B.6 and B.7, we replicate the main analyses on the full sample using indicators for correct price recalls within fixed error bounds, instead of the continuous recall accuracy measure introduced below. These checks confirm that our results regarding both discount utilization and the drivers of price knowledge are robust to the cleaning process.

[8]In fact, the umbrella chain reports that it stopped tracking non-members' bank card activity after finding that their purchasing patterns closely mirrored those of Trumf members.

[9]Population benchmarks from Statistics Norway (2024a) for age and gender and Statistics Norway (2024b) for education.

[10]In our sample, there is a positive correlation between being the household's main grocery shopper and being a woman (Spearman's $\rho = 0.24$).

and Reguant (2020), which examines consumer heterogeneity beyond search costs; Shelegia and Wilson (2021), which modifies how ties are resolved by consumers when firms post identical prices; and Myatt and Ronayne (2026), which allows for asymmetric costs across firms.[11]

The central mechanism in Varian (1980), and the search literature that has followed, is that informed consumers act upon their price knowledge and take advantage of temporary sales. While, as discussed in the introduction, the previous literature is broadly consistent with the predictions from search models, it does not necessarily follow that variation in consumer price knowledge is key for driving consumer responses to temporary sales. Observed responses to temporary sales might instead reflect that some consumers are more willing to act upon observed price differences than others, rather than reflecting differences in knowledge per se. For instance, Hendel and Nevo (2013) show how temporary sales of storable products are profitable as they allow retailers to sell at low prices to more price-sensitive consumers who are willing to stockpile, while at the same time keeping high non-sale prices, allowing them to charge high prices to non-stockpiling customers.

The key hypothesis we want to evaluate is that more informed consumers pay lower prices by timing their purchases to coincide with temporary sales. Note that we focus on the demand-side mechanism: the effect of an individual's price knowledge on their discount utilization. A key result in the class of models that motivate our study is that informed consumers also provide a positive externality for other consumers in the market—strengthening competition. Our setting does not allow us to examine such equilibrium effects. Hence, we examine the quantitative importance of the key demand-side mechanism in this class of models, rather than trying to quantify equilibrium pricing responses. Figure 2.1 provides an overview of our strategy to identify the effect of consumer price knowledge on discount utilization.

[11]Related work also examines the "rockets and feathers" phenomenon, e.g., Tappata (2009), which introduces uncertainty over input costs in a search model similar to that of Varian (1980). In this model, expected gains from searching are lower when costs are expected to be high, giving rise to asymmetric price responses to cost shocks. We expect such mechanisms to be particularly relevant in, for example, gasoline markets with large fluctuations in input costs and no temporary sales.

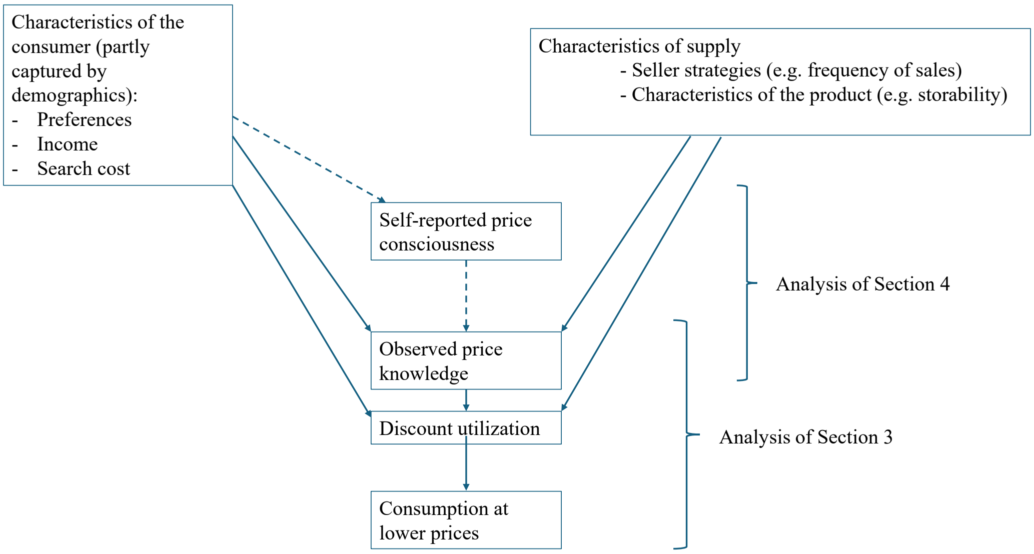


Figure 2.1: An overview of the links between price knowledge and discount utilization.

In Section 3, we use OLS to estimate variants of the following specification:[12]

$$DU_i = \alpha + \beta\, PK_i + \mathbf{X}_i'\boldsymbol{\gamma} + \varepsilon_i \tag{2.1}$$

where DU_i denotes observed discount utilization for individual i. We detail the construction of the variables further below, but note that DU_i is calculated using each individual's purchasing history over the preceding 18 months for more than 700 products: the products included in the questionnaire as well as their close substitutes. PK_i denotes the individual's observed price knowledge, based on their estimates of the shelf prices of 18 products from our survey. Our hypothesis is that the estimated β is positive, implying that consumers with greater price knowledge respond more strongly to temporary sales. We further include a rich set of demographic controls $\mathbf{X}_i$.

The estimated β captures the conditional association between price knowledge and discount utilization. A natural question is how much of the estimated association between price knowledge and discount utilization could plausibly be driven by confounding factors, as opposed to reflecting a direct relationship between price knowledge and behavior.

First, note the timing of the survey and the purchase history—the survey was unanticipated by respondents, and the price knowledge that they demonstrate on the day of the survey is set in relation to their use of temporary sales over a period of 18 months prior to the survey. One potential threat to identification would be reverse causality: discount utilization may increase price knowledge, and individuals might happen to know prices because they recently shopped at that price. Both the much larger set of products

[12]We assess the validity of the linearity assumption using a binned scatterplot in Section 3 and by replacing PK_i with quintile indicators in Appendix B.4.

used to calculate discount utilization (more than 700 vs. 24) and the long time period of 18 months suggest that this is unlikely to be a prominent concern[13]

Second, we do not expect measurement error in the dependent variable to be a material concern. Discount utilization is based on recorded sales, where the membership is linked to payment cards.[14]

The extent to which consumers can purchase products on temporary sale clearly depends on product characteristics and the competitive environment in which they make their purchases. We ask about the prices of 24 products in total, which represent items that are arguably typical of a grocery retail store: relatively low-priced consumer goods such as dairy products (milk, cheese, butter), vegetables (avocado, cucumber), fish and meat, potato chips, and carbonated soft drinks. We examine differences in discount utilization using sales across all close substitutes for the products we ask about. Our interpretation is that we capture the extent of price knowledge and its effect on discount utilization for a set of typical grocery products, reporting an average effect rather than focusing on how discount utilization varies with product characteristics. We make two exceptions along the heterogeneity dimension: first, we examine whether discount utilization depends on price variability; second, we examine heterogeneity with respect to product storability. Other things equal, we expect the gains from price knowledge to be higher for storable products such as detergent (the focus of Hendel and Nevo, 2006) or soft drinks (the focus of Hendel and Nevo, 2013).

Another concern is omitted variable bias—specifically, underlying preferences for purchasing on temporary sales. We therefore report estimates from specifications both with and without a rich set of controls. First, we directly target these preferences by controlling for a survey-based measure of price sensitivity, constructed from respondents' stated importance of price in their choice of products and stores and their tendency to purchase on sales. We also include a rich set of demographic controls that may affect both price knowledge and discount utilization. For instance, a person with full-time employment may find it less attractive both to spend time gathering price information and to act upon temporary sales when they become aware of them. Since these variables may be imperfect proxies for the underlying preferences, we additionally control for past shopping behaviors (e.g., number of store visits) and a self-reported measure of price learning effort to further mitigate omitted variable bias in a separate specification. However, these measures also absorb what we consider valid variation in price knowledge—for example, consumers who exert greater effort in gathering price information tend to have higher knowledge. This may affect the precision of our estimates, as identification then relies on residual variation

[13] To test this formally, we also estimate Equation 2.1 using a measure of discount utilization constructed only from products not included in the survey. The results are reported in Appendix B.2.

[14] Payment cards are the dominant way to purchase in retail settings in Norway. In a survey commissioned by the Norwegian central bank, only 2% of payments by consumers in retail settings were made with cash (see Norges Bank, 2024, Chart 21).

in price knowledge, driven, for instance, by differences in attention, memory, or exposure to environmental cues and distractions.

Beyond the average effects, we extend the main analysis by examining heterogeneity in the returns to price knowledge to better understand the conditions under which it is most effective. Specifically, we investigate whether the link between knowledge and discount utilization is stronger for consumers whose shopping baskets exhibit high price volatility and for those with higher subjective price consciousness. We also explore the role of product characteristics by distinguishing between perishable and non-perishable goods. Finally, to assess the economic significance of our findings, we quantify the expected monetary savings associated with higher price knowledge.

In Section 4, we turn our attention to the determinants of price knowledge. If price knowledge affects the prices consumers pay, understanding what drives it is a prerequisite for assessing whether and how it can be improved. We estimate regressions of the following form:

$$PK_{ijc} = \alpha + \mathbf{X}_i'\boldsymbol{\beta} + \mathbf{Z}_j'\boldsymbol{\delta} + \mathbf{S}_{jc}'\boldsymbol{\phi} + \mathbf{W}_{ic}'\boldsymbol{\theta} + \mathbf{Q}_{ijc}'\boldsymbol{\kappa} + \eta_c + \nu_{ijc}. \tag{2.2}$$

We thus regress the price knowledge of consumer i for product j in chain c on individual characteristics ($\mathbf{X}_i$), product characteristics ($\mathbf{Z}_j$), chain-specific product characteristics ($\mathbf{S}_{jc}$), individual–chain shopping behavior ($\mathbf{W}_{ic}$), individual–product–chain shopping behavior ($\mathbf{Q}_{ijc}$), and a chain effect (η_c).

2.3 Defining Some Key Variables

2.3.1 Price Knowledge

The binary distinction between informed and uninformed consumers in many search models is clearly a simplification. To test whether price knowledge affects prices paid, we therefore construct a continuous measure of price knowledge based on consumers' ability to accurately recall current prices for a range of store × product combinations. This measure is then aggregated across store × product combinations to derive an overall price knowledge score for each consumer.

Our primary measure of price knowledge, PK, is based on the Percentage Absolute Deviation (PAD) of the surveyed price recall of individual i for product j at chain c relative to the true shelf price in the home municipality m on survey response date t:

$$\mathrm{PK}_{ijc} = -\underbrace{\left| \frac{P_{ijc}^{\text{Recall}} - P_{jcmt}^{\text{Shelf}}}{P_{jcmt}^{\text{Shelf}}} \right|}_{PAD_{ijc}} \tag{2.3}$$

The PAD measure has previously been used in the marketing literature on price knowledge (see, e.g., Dickson and Sawyer, 1990). By surveying respondents outside of purchasing situations, as opposed to studies that survey consumers post-purchase, in-store (see Conover, 1986; McGoldrick and Marks, 1987; Dickson and Sawyer, 1990; Wakefield and Inman, 1993; Le Boutillier et al., 1994), we aim to capture the long-term price memory of the respondents (see Evanschitzky et al., 2004 for a discussion of short- vs. long-term memory and survey timing). PK_{ijc} is non-positive, and an increase in PK_{ijc} represents an increase in price knowledge towards zero, i.e., perfect information.

To obtain a measure of the individual's overall price knowledge, we calculate the average percentage error the individual made in the survey:

$$\mathrm{PK}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \mathrm{PK}_{ijc} \tag{2.4}$$

where K_i is the number of product $\times$ chain combinations for which the individual recalled prices in the survey.

2.3.2 Price Sensitivity, Learning Effort, and Consciousness

The survey includes questions designed to capture different aspects of how consumers relate to grocery prices, using statements answered on a 5-point Likert scale. The exact statements are reported in Appendix C, and the distributions of responses are shown in Figure A.2. The first three questions measure the extent to which respondents incorporate prices into their purchasing decisions, including the importance of price for store and product choice and whether they try to buy on sales and promotions. We use these items to construct an index of *Price Sensitivity*, which throughout the paper is interpreted as consumers' underlying preferences for responding to price incentives, reflecting tastes rather than price beliefs or knowledge. Questions 4–7 capture behaviors related to monitoring and comparing grocery prices, such as checking shelf prices and reading grocery chain flyers, and are used to construct an index of *Price Learning Effort*. To summarize these two dimensions in a single measure, we additionally construct an aggregate index of subjective *Price Consciousness* based on questions 1–7. Following the broad use of the term in the marketing literature (see, e.g., Alford and Biswas, 2002; Loy et al., 2020), we interpret this aggregate as capturing a general tendency to care about prices.

Indices are constructed as weighted averages of the underlying Likert items. We obtain weights from a simple factor model, $\textit{Likert Response}_i = \tau_i + \lambda_i \textit{Factor} + \epsilon_i$, where i indexes the item, τ_i is an intercept, λ_i is the factor loading, ϵ_i is a residual, and *Factor* denotes either Price Sensitivity, Price Learning Effort, or Price Consciousness. Factor loadings are estimated using diagonally weighted least squares and reported in Table A.3 in Appendix A. Each item's weight is its factor loading divided by the sum of loadings for the corresponding factor.[15]

2.3.3 Discount Utilization

The products in Varian (1980) are assumed to be homogeneous, making it straightforward to compare prices paid by different consumer groups. However, in grocery retailing, products and stores are differentiated, and consumers may pay lower prices not only due to price knowledge but also due to preferences for lower-priced stores or products. Our detailed data on purchasing history allow us to calculate an individual-level price measure that strips out variation stemming from differences in the overall price level of the store and products.

To test the prediction that more informed consumers pay lower prices on average, and to quantify the gains from price knowledge, we calculate *Discount Utilization* (DU): the percentage difference between (i) the average of the daily shelf prices in the consumer's home municipality m over an event window e of length E ($\bar{P}^{\text{Shelf}}_{jcme}$), and (ii) the average price paid by consumer i on day t for product j at chain c (P^{Paid}_{ijct}):

$$\text{DU}(E)_{ijct} = \frac{\bar{P}^{\text{Shelf}}_{jcme} - P^{\text{Paid}}_{ijct}}{\bar{P}^{\text{Shelf}}_{jcme}} \tag{2.5}$$

Thus, we interpret "discount" broadly as any price paid below the average, and not only low prices due to promotional sales. A positive value indicates that the consumer paid less on that day than what someone who times their purchases uniformly over the event window would expect to pay (i.e., buying on any day with equal probability).

Of course, the scope for paying lower prices depends on the event window. Since traditional sales apply to everyone who buys a product on the same day, the only way to obtain intraday bargains is to use personal coupons or a quantity or expiration-date discount. Conversely, over longer periods like 1–3 months, one can pay less than the average shelf price by timing purchases with sales. We have 18 months of purchasing history, but Equation 2.5 could capture unbalanced spending over time combined with general inflation, so we focus on a three-month event window ($E = Q$). This limits the effects of inflation and leaves ample room for timing sales.[16]

In order to calculate the overall tendency of a consumer to buy specific products at lower prices, we average Equation 2.5 over time:

$$\text{DU}(E)_{ijc} = \frac{1}{T_{ijc}} \sum_{t=1}^{T_{ijc}} \text{DU}(E)_{ijct} \tag{2.6}$$

where T_{ijc} is the number of times that an individual i is observed purchasing product j at

[15] Using simple unweighted averages yields very similar results.

[16] The results are similar using 1, 3, or 18 months ($E \in \{M, Q, FP\}$), but are strongly muted using daily discount utilization ($E = D$), as expected.

chain c in the data. This measures how much less an individual tends to pay for specific products than someone buying the same product but timing their purchases uniformly over the event windows e that the individual's purchases fall into.

To obtain an overall measure of a consumer's ability to identify and act on bargains, we further average Equation 2.6 across product × chain combinations:

$$\text{DU}(E)_i^{\text{Simple}} = \frac{1}{N_i} \sum_{n=1}^{N_i} \text{DU}(E)_{ijc} \tag{2.7}$$

where N_i is the number of product × chain combinations that we observe individual i purchasing in the data. This gives us a simple measure that quantifies an individual's tendency to exploit price variation over time for all the different products that they buy across the sample of more than 700 products, and not only the 18 products for which they were later asked to provide price recalls.

Note that we do not average over every transaction, which would overweight frequently purchased products. The reason is that one could identify a sale once and then act on it several times on consecutive days. We are primarily interested in the lower prices one can achieve because of greater price knowledge. However, we also calculate a sales-weighted version of Equation 2.7,[17] which allows us to calculate the percentage savings that the consumer makes by buying their current basket of groceries on sale relative to buying the same basket uniformly over time. Since sales weights are also sensitive to purchasing frequency, we focus on simple discount utilization (Equation 2.7) when testing and discussing the theoretical predictions of models such as Varian's (1980) model of sales concerning price knowledge and average prices paid, and we instead focus on sales-weighted discount utilization (Equation 2.8) when calculating the monetary gains associated with price knowledge.

2.3.4 Price Variation and Scope for Discount Utilization

The scope for buying products on sale critically depends on the sales activity and overall price variation of the products. To capture and control for this feature of grocery prices, we calculate the coefficient of variation for each product j × chain c × municipality m combination over the sample period:

$$CV_{jcm} = \frac{\sigma_{jcm}}{\mu_{jcm}} \tag{2.9}$$

[17] Formally, we compute

$$\text{DU}(E)_i^{\text{Sales-Weighted}} = \sum_{n=1}^{N_i} \frac{S_{ijc}}{S_i} \text{DU}(E)_{ijc} \tag{2.8}$$

where S_{ijc} is the consumer's spending on each chain × product combination and S_i is the consumer's total spending, over the sample period.

where σ_{jcm} is the empirical standard deviation of price and μ_{jcm} is the product's mean price over the sample period. There is more room for discount utilization in products with high coefficients of variation, but the relation between knowledge of a product's price and its variation is less clear.

At the individual level, it is useful to consider the overall price variation in the consumer's shopping basket. Consumers who primarily buy products with little price variation over time are more limited in their opportunity to find bargains, and vice versa. We calculate the average price variation of the products that individual i buys in the following way:

$$CV_{i[\text{Simple}]} = \frac{1}{N_i} \sum_{n=1}^{N_i} CV_{jcm} \tag{2.10}$$

where N_i is the number of product $\times$ chain combinations that we observe individual i purchasing in the data.[18]

2.4 A Closer Look at Price Variation

Prices vary over time for many reasons, including changes in wholesale costs and and markdowns on products approaching their expiry dates. A particularly important source of price variation, however, is temporary sales.[19] Figure 2.2 shows daily shelf prices for two products included in our price survey, Barilla spaghetti and a large bottle of Pepsi Max for the two chains. The solid blue line represent the 90-day rolling median price which represents the regular price of the product. We categorize a discount spell as one where price is at least 3% lower than this regular price. For both products and chains we see several discount spells of varying depth as well as variation in the regular price, providing opportunities for informed customers to take advantage of low prices.

Table 2.1 provides systematic evidence of discounts in the data for the 18 month period preceding the survey, both only for the 24 survey products and for all the products included in the study. Broadly speaking, products are on sale 11 to 19% of days, with sale periods ranging from 3 days to two weeks and the mean sale depth at a substantial 14 to 18%. Column (4) shows that there is significant variation in the regular price not related to temporary sales, by reporting its coefficient of variation. In Column (5), we also

[18]Similarly, the price variation of the consumer's average shopping basket is given by:

$$CV_{i[\text{Sales-Weighted}]} = \sum_{n=1}^{N_i} \frac{S_{ijc}}{S_i} CV_{jcm} \tag{2.11}$$

where S_{ijc} is the consumer's spending on each chain $\times$ product combination and S_i is the consumer's total spending over the sample period.

[19]We use the term temporary sales broadly to refer to temporary deviations from the regular shelf price. These need not all correspond to advertised campaigns or marked promotional discounts as some may instead reflect short-run price responses among competing chains.

take account of temporary sales and find that the coefficient of variation approximately doubles. These results confirm that there is substantial scope for consumers to reduce expenditure through timing of purchases, both from variation in regular prices and from additional variation due to sales.

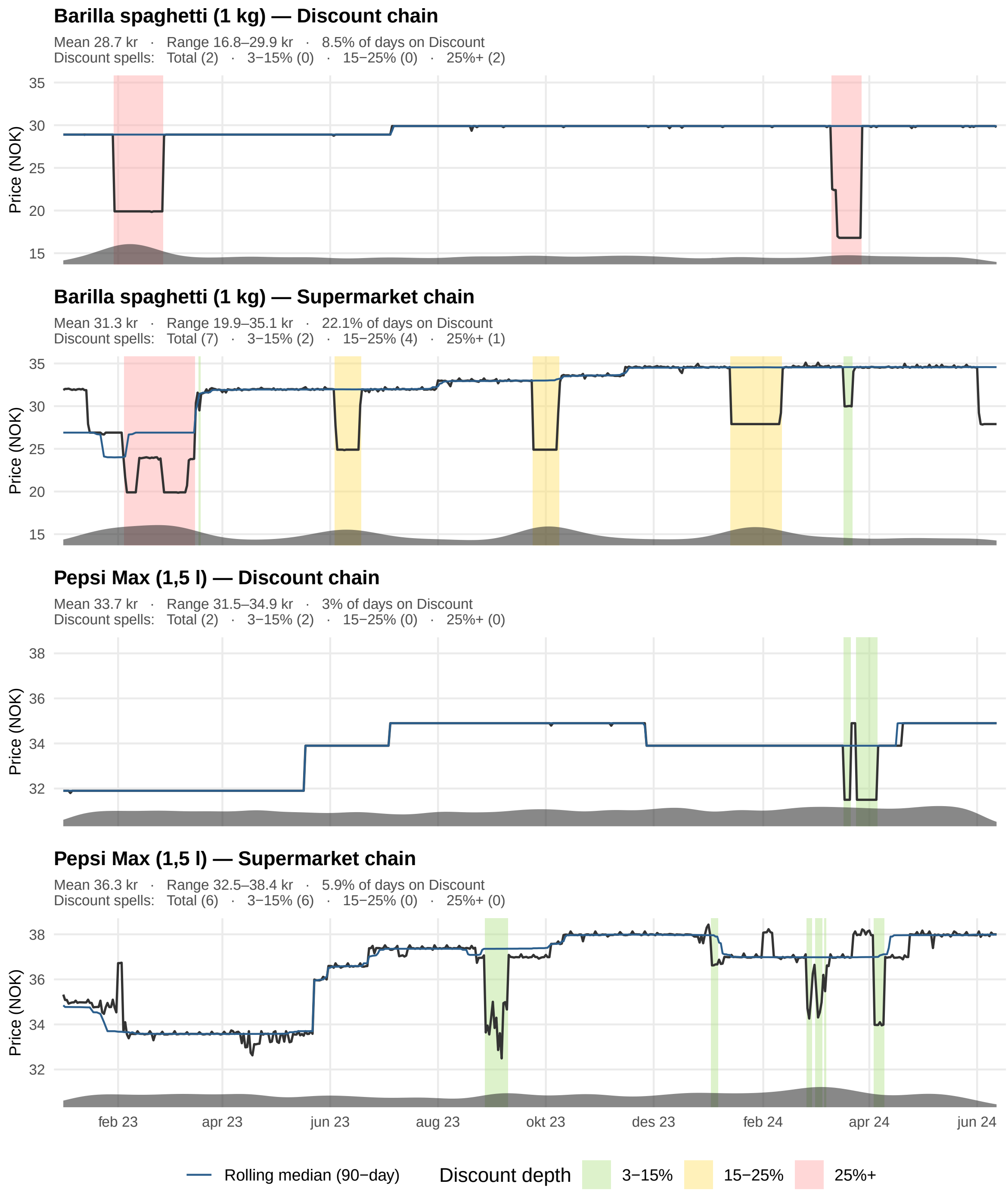


Figure 2.2: Daily shelf prices for Barilla spaghetti (1 kg) and Pepsi Max (1.5 l) at the discount and supermarket chains over the 18-month panel. Shaded rectangles mark discount spells classified by depth (3–15%, 15–25%, 25%+) relative to a 90-day rolling-median reference (the solid blue line). The dark band along the floor of each panel is the kernel-density estimate of the purchase distribution for the (product × chain) cell.

Table 2.1: Price dynamics: sale frequency, depth and dispersion

	Share days on sale	Mean sale length (days)	Mean sale depth	Regular price CV	Total price CV
	(1)	(2)	(3)	(4)	(5)
Survey products, Both chains	0.122	11.444	0.164	0.070	0.107
Survey products, Discount chain	0.108	14.498	0.185	0.066	0.103
Survey products Supermarket chain	0.136	8.518	0.144	0.074	0.111
All products, Both chains	0.160	4.271	0.172	0.052	0.108
All products, Discount chain	0.108	6.290	0.173	0.059	0.100
All products Supermarket chain	0.189	3.296	0.171	0.048	0.113

Notes: The reference (regular) price is the 90-day centered rolling median of the daily shelf price; the daily price is first averaged across municipalities to a single national series per product and chain. A day is on sale when the posted price is at least 3% below that reference; Share days on sale is the fraction of observed days meeting this cut. A sale (spell) is an unbroken stretch of consecutive days on sale: Mean sale length is its average duration in days, and Mean sale depth is the average across sales of each sale's deepest discount (largest % below the reference within the spell). Regular-price CV and Total price CV are the coefficients of variation (standard deviation over mean) of the reference price and the posted price, respectively. All statistics are computed per product-by-chain series (series with fewer than 30 observed days are excluded, to keep the rolling-median reference based on multiple observations); cross-product means by chain, and Both chains pools the two chains. Survey products: the 24 surveyed items; all products: these 24 survey products together with the more than 700 close substitutes for which we obtained shelf prices from the grocery group.

3 Price Knowledge and Discount Utilization

Before turning to regression analysis of the relation between price knowledge, discount utilization, and other variables, let us describe the variation in these variables. The empirical distribution of individual price knowledge, summarized by PK, is shown in the left panel of Figure 3.1. The mean PK is -0.27, indicating that, on average, respondents over- or underestimate the true shelf price by 27%. The interquartile range spans from -0.32 to -0.22. We find that 17% of recalled prices are within 5% of the true shelf price, and 50% are within 20%. The distribution shows no clear mass points that would suggest discrete consumer information types. Moreover, no respondent recalls prices perfectly—the maximum PK is -0.04.[20]

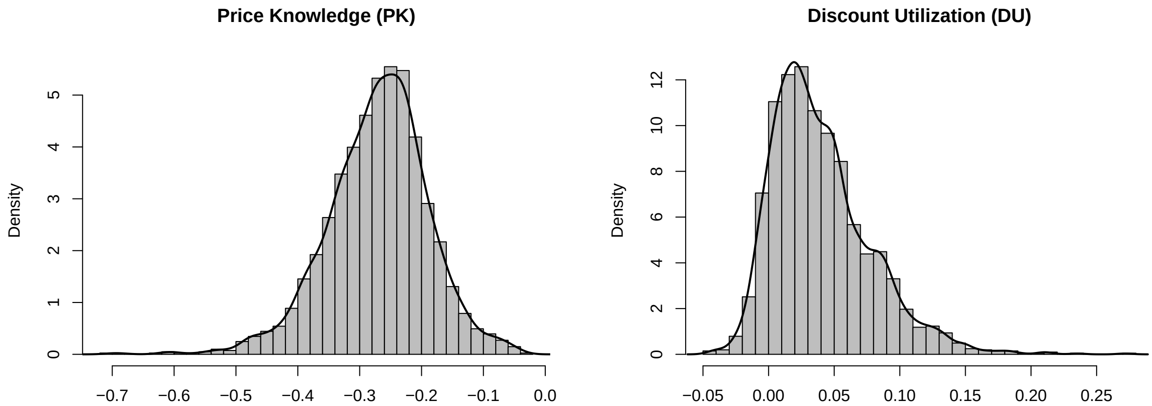


Figure 3.1: Histograms and density estimates of price knowledge (PK) and discount utilization (DU), computed as simple averages across the 24 survey products and across the 700+ products in the full purchase sample, respectively. Histogram bin sizes follow Scott's rule and density bandwidths the Sheather–Jones plug-in estimator.

The distribution of discount utilization, computed as the simple average across products, is shown in the right panel of Figure 3.1. A value of zero implies behavior analogous to the uninformed consumers in Varian's (1980) model—those who time purchases randomly with respect to sales. Most consumers, however, exhibit positive discount utilization: the interquartile range is 0.01 to 0.06, and the mean is 0.04, indicating that consumers tend to concentrate purchases on days when prices are below average.

The Spearman rank correlation between PK and DU is 0.31. To obtain a nonparametric estimate of the expected value of discount utilization conditional on price knowledge, we present a binned scatterplot in Figure 3.2.[21] The figure provides clear

[20] We also note that among all recalled prices included in the analysis, 31% were provided by respondents who had purchased the product at the relevant chain during the sample period. For these cases, the average time since purchase is 4.7 months, and the median is 2.7 months.

[21] The implementation follows Cattaneo et al. (2019). Each point is the average discount utilization among the respondents with price knowledge within the corresponding bin. The bins are chosen to minimize the integrated mean squared error, and the 95% confidence bands are calculated using piecewise polynomials.

evidence that respondents with greater price knowledge purchase more on sale and pay lower prices on average, consistent with Varian's (1980) model of sales.

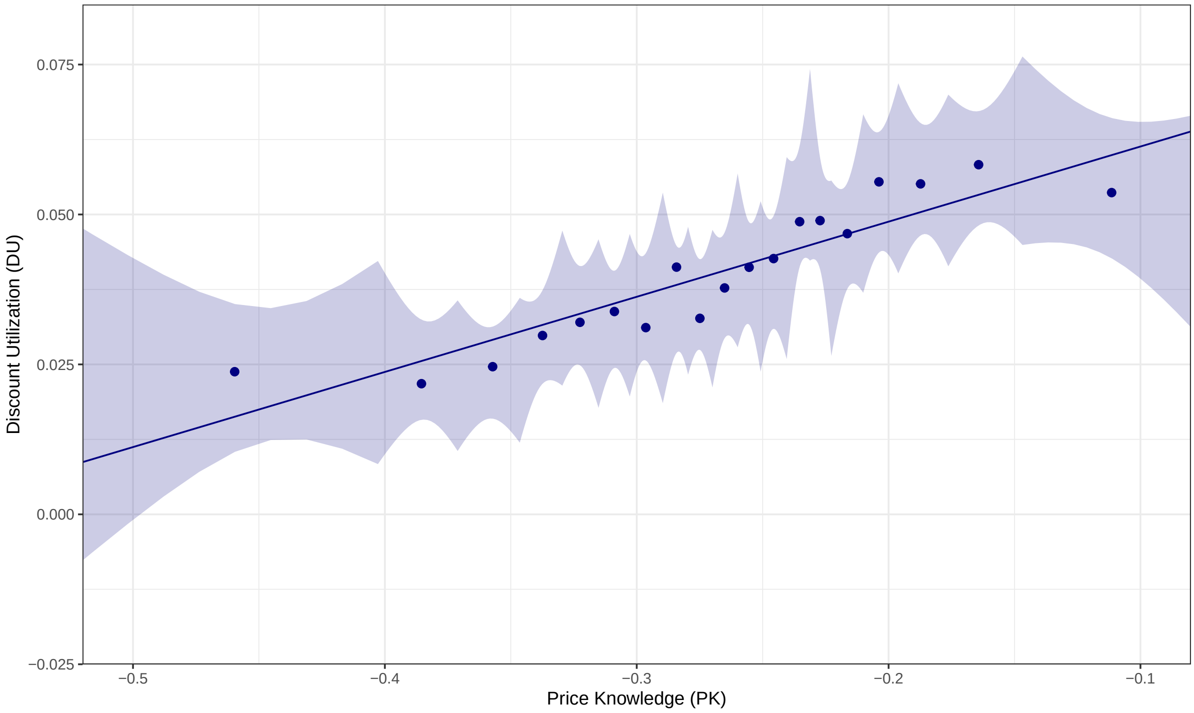


Figure 3.2: Estimate of the expected value of discount utilization (DU) conditional on price knowledge (PK) from a binned scatterplot with a linear fit. PK and DU are computed as simple averages across the 24 survey products and across the 700+ products in the full purchase sample, respectively. The number of bins is chosen to minimize the integrated mean squared error, and 95% confidence bands are calculated using piecewise polynomials.

3.1 Linear Regression Analysis

To parameterize the relationship between discount utilization and price knowledge, and to control for confounding variables, we present OLS estimates of Equation 2.1 with different sets of controls in Table 3.1.[22]

[22]We report heteroskedasticity-robust standard errors. Following Abadie et al. (2022), we do not cluster by municipality because the sampling design was not clustered.

Table 3.1: Discount utilization regressions

	Discount Utilization		
	(1)	(2)	(3)
Price Knowledge	0.125*** (0.011)	0.121*** (0.010)	0.087*** (0.009)
Male		0.000 (0.002)	0.001 (0.002)
Single Person Household		0.004 (0.003)	0.004 (0.003)
Has Children		0.004* (0.002)	0.005** (0.002)
Full-Time Employment		-0.005* (0.002)	-0.003 (0.002)
Higher Education		0.001 (0.002)	0.002 (0.002)
Rural		-0.006*** (0.002)	-0.005** (0.002)
Household Income > 950' NOK		-0.002 (0.002)	-0.001 (0.002)
Age: 40-59 Years		0.007*** (0.002)	0.007*** (0.002)
Age: Above 60 Years		0.008** (0.003)	0.009*** (0.002)
Price Sensitivity		0.007*** (0.001)	0.003* (0.001)
Price Learning Effort			0.008*** (0.001)
# Frequented Grocery Chains			0.000 (0.001)
Mainly Uses Discount Stores			-0.003 (0.002)
Average Weekly Store Visits			0.003*** (0.001)
Main Grocery Shopper			0.002 (0.002)
Average Weekly Variety (per 10 products)			-0.003*** (0.001)
Three Customer Program Memberships			0.005*** (0.001)
Basket Coefficient of Variation			1.216*** (0.074)
Constant	0.074*** (0.003)	0.041*** (0.006)	-0.116*** (0.010)
Observations	2,028	1,935	1,928
Dependent variable mean	0.04	0.04	0.04
R^2	0.07	0.12	0.35

Notes: This table reports results from estimation of Equation 2.1. The dependent variable, individual-level discount utilization, is the percentage difference between the quarterly average of daily modal (shelf) prices and the average price paid on the purchase day, averaged across purchase days and then across product × chain combinations (simple weights). Price knowledge is calculated as the negative average percentage absolute deviation between the respondent's recalled prices and true shelf prices. The average weekly variety is calculated as the average number of unique products purchased per week over the sample period. Price sensitivity and learning effort are calculated as the factor-weighted averages of Likert-scale survey items 1–3 and 4–7, respectively. Basket CV is calculated as the coefficient of variation of shelf prices over the sample period, averaged across all products in the consumer's purchase history, using simple weights. Column 1 reports the simple regression on price knowledge (PK), Column 2 adds controls for demographics and price sensitivity, and Column 3 adds controls for shopping habits and price learning effort. The standard errors shown in parentheses are robust to heteroskedasticity (HC1). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

The regression of discount utilization on price knowledge, presented in Column 1, shows that a 10 percentage-point increase in price knowledge is associated with a 1.3 percentage-point increase in discount utilization. This change in price knowledge is well within the variation in the data; the average PK in the sample is −27%, the standard deviation is 8%, and 50% of respondents have a PK between −32% and −22%. An individual with high price knowledge (e.g., $PK_i = -10\%$) pays, on average, 4% less for groceries than someone with low price knowledge (e.g., $PK_i = -40\%$).

The regression constant is also of theoretical interest. Assuming a linear extrapolation to $PK_i = 0$ is reasonable, it yields the expected discount utilization of a perfectly informed consumer. At 7.4%, the discount utilization of a perfectly informed consumer exceeds both the sample average of 4% and the reference level associated with "uninformed" consumers in Varian's (1980) model, which is 0%.

Some consumers may have a greater propensity to exploit temporary sales, independent of price knowledge. If consumers with higher conditional price sensitivity are also more inclined to acquire price information, this would confound the relationship between price knowledge and discount utilization—more price-sensitive consumers might have paid lower prices even in the absence of superior price knowledge. To address this concern, Column 2 in Table 3.1 adds controls for demographics and the price sensitivity index, which proxy for stable traits and preferences that may jointly affect price knowledge and discount utilization. Reassuringly, the coefficient on PK is not materially altered, suggesting that price knowledge is an independent prerequisite for identifying and acting on temporary sales.

In terms of explanatory power, the R^2 increases modestly by 5 percentage points as we move from Column 1 to Column 2, and the coefficients generally have the expected signs: individuals with children and those without full-time employment pay lower prices, as do older respondents and those living in urban areas. The price sensitivity index ranges from 1 to 5, and a one-unit increase (equal to the interquartile range) is associated with a 0.7% reduction in prices paid. Income, conditional on price sensitivity and other demographics, shows no ceteris paribus effect on discount utilization.

The price sensitivity index is constructed to capture consumers' underlying preference for responding to price incentives. However, the index relies on subjective interpretations of three survey questions and may only imperfectly capture the individual's preferences. Even if the price sensitivity index is the best single proxy at our disposal, a broader set of control variables may better capture the underlying trait.

Column 3 adds variables related to consumers' shopping habits and price learning effort, further increasing R^2 by 23 percentage points. Frequent store visits are associated with higher discount utilization, consistent with having more opportunities to act on sales. Buying many different products per week is associated with lower discount utilization, perhaps because regularly purchasing a wide variety of products implies that some are bought outside of sales.[23] Membership in all major customer programs is associated with paying 0.5% lower prices. A substantial share of the increase in R^2 is driven by the inclusion of the average coefficient of variation of the products purchased by the respondent: a 0.01 standard deviation increase in basket price variation is associated with 1.2% lower prices paid. This reflects that the scope for finding bargains depends on price variation over time.[24] Finally, a one-unit increase in the price learning effort index is associated with 0.8% lower prices paid, while price sensitivity continues to independently predict higher discount utilization.

The ceteris paribus association between price knowledge and discount utilization is now somewhat muted but remains economically and statistically significant. One interpretation is that we have further mitigated omitted variable bias by better capturing underlying preferences for exploiting temporary sales. For example, shopping frequency, program membership, and engagement in price information gathering may all reflect a general interest in exploiting temporary sales that the price sensitivity index does not capture. On the other hand, as we now rely on variation in price knowledge unrelated to learning effort—for instance, variation driven by attention, memory, or environmental cues and distractions—we also absorb what we consider valid variation in price knowledge. As this may affect the precision of our estimates, we view the estimates in Columns 1 and 2 as our primary results.[25]

3.2 Heterogeneity and Robustness

We now consider several extensions of the main specification, presented in Table 3.2, to examine heterogeneity across products and respondents and thereby assess the conditions under which price knowledge yields the highest returns. Column 1 reproduces Column 1 of Table 3.1 for ease of comparison.

[23]If we instead count the number of different products purchased over the entire sample period, the sign flips. A consumer might buy many different products over time precisely because they take advantage of discounts when they arise.

[24]The coefficient on PK is minimally affected by the inclusion of Basket CV, despite its substantial contribution to R^2. More generally, to the extent that variables affecting the scope for discount utilization (e.g., basket price variation, shopping frequency, and product variety) are driven by underlying preferences unrelated to price, we do not expect their inclusion to alter the ceteris paribus relationship between PK and DU.

[25]As discussed in Section 2.2, one potential concern is that discount utilization may itself increase price knowledge, and that individuals may know the prices of the surveyed products simply because they recently purchased them at a discounted price. In Appendix B.2, we therefore report results from estimating Equation 2.1 using a measure of discount utilization constructed only from products not included in the survey. The association between price knowledge and discount utilization remains robust and highly statistically significant when the survey products are excluded. If anything, the estimated coefficient on price knowledge is slightly larger in this specification than in the baseline results.

Table 3.2: Discount utilization regressions: effect heterogeneity

Discount Utilization:	Simple Weights			Perishable	Non-Perishable	Sales Weights
	(1)	(2)	(3)	(4)	(5)	(6)
PK	0.125***	0.087***	-0.003	0.117***	0.133***	0.070***
	(0.011)	(0.010)	(0.044)	(0.012)	(0.013)	(0.010)
High Basket CV		0.041***				
		(0.006)				
PK × High Basket CV		0.053**				
		(0.020)				
Price Consciousness			0.021***			
			(0.004)			
PK × Price Consciousness			0.028*			
			(0.012)			
Constant	0.074***	0.050***	-0.011	0.067***	0.081***	0.040***
	(0.003)	(0.003)	(0.014)	(0.003)	(0.004)	(0.003)
Observations	2,028	2,028	1,961	2,028	2,028	2,028
Dependent variable mean	0.04	0.04	0.04	0.04	0.05	0.02
R^2	0.07	0.20	0.14	0.05	0.06	0.04

Notes: This table reports discount utilization regressions examining heterogeneity in the relationship between price knowledge and discount utilization across products and respondents. The dependent variable, individual-level discount utilization, is the percentage difference between the quarterly average of daily modal (shelf) prices and the average price paid on the purchase day, averaged across purchase days and then across product × chain combinations. Price knowledge (PK) is calculated as the negative average percentage absolute deviation between the respondent's recalled prices and true shelf prices. Price consciousness is calculated as the factor-weighted averages of Likert-scale survey items 1–7. High basket CV is an indicator for above-median basket price variation, calculated as the coefficient of variation of shelf prices over the sample period, averaged across all products in the consumer's purchase history, using simple weights. Column 1 reports the baseline regression of DU, calculated using simple weights, on PK. Column 2 interacts PK with an indicator for above-median basket price variation. Column 3 interacts PK with the price consciousness index. Columns 4–5 calculates discount utilization separately for perishable and non-perishable products, respectively. Column 6 calculates discount utilization using sales weights. The standard errors shown in parentheses are robust to heteroskedasticity (HC1). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Basket Price Variation If price knowledge drives discount utilization, its impact should be most pronounced for consumers purchasing products with high price volatility—precisely where the opportunity for identifying bargains is greatest. Column 2 of Table 3.2 presents the results of interacting price knowledge with an indicator for a shopping basket with above-median price variation. The interaction-term coefficient is positive and statistically significant, confirming that the link between price knowledge and discount utilization is stronger for individuals with higher basket volatility—a 10 percentage-point increase in price knowledge is associated with 0.9 and 1.4 percentage-point increases in discount utilization for the low- and high-volatility consumers, respectively.

This finding aligns with the fundamental prediction of Stigler (1961): that the returns to search are increasing in price dispersion. Interpreting dispersion here as variation over time, this result confirms that the marginal value of acquiring price knowledge is greatest when the underlying distribution of prices is widest.[26]

Price Consciousness While price knowledge provides the ability to identify bargains, acting on this information requires motivation. We therefore examine whether the link between price knowledge and discount utilization is stronger for consumers who are fundamentally more concerned with prices.

For this analysis, we utilize the aggregate *Price Consciousness* index constructed in Section 2.3.2. While our previous specifications used distinct components (such as price sensitivity and learning effort), here we employ the composite index drawing on a broader set of survey items to capture a consumer's general tendency to care about prices. Using a single proxy allows for a tractable and easily interpretable interaction model that tests whether price-conscious consumers benefit more from being informed.

Column 3 of Table 3.2 presents the regression estimates including the interaction between price knowledge and price consciousness.[27] We find that the interaction term is positive and statistically significant at the 5% level, confirming that the effect of price knowledge on discount utilization is increasing in price consciousness. Our interpretation is that knowledge translates into lower prices paid more effectively when the consumer is highly motivated to use it.

To quantify this heterogeneity, we compute the marginal effect of price knowledge at different levels of price consciousness, reported in Table 3.3. The results show that the return to price knowledge is substantially higher for highly price-conscious consumers. For a consumer at the first quartile of price consciousness, a 10 percentage-point increase in price knowledge is associated with a 0.92 percentage-point increase in discount utilization. For a consumer at the third quartile, approximately one unit higher on the price consciousness index, this effect rises to 1.21 percentage points—corresponding to an increase in efficacy of roughly 30%.[28]

These findings have important implications for policy. First, they suggest that policies aimed at improving price availability and knowledge may have progressive distributional effects. In our sample, high price consciousness is negatively correlated with income and positively correlated with being young, being a woman, and having children. To the extent that disadvantaged groups are more price conscious, population-wide gains in price knowledge (e.g., through clearer labeling or price comparison tools) may disproportionately

[26] The expected minimum price paid for a product after sampling n sellers when prices are uniformly distributed on $p \in [0, b]$ is equal to $E[p^{\text{Min}}] = \frac{b}{n+1}$ (see Stigler, 1961, p. 215). The gain from more search n is then $\frac{\partial |E[p^{\text{Min}}]|}{\partial n} = \frac{b}{(n+1)^2}$, which is increasing in price dispersion b: $\frac{\partial^2 |E[p^{\text{Min}}]|}{\partial n \partial b} = \frac{1}{(n+1)^2} > 0$.

[27] The estimating equation is $DU_i = \alpha + \delta \cdot PK_i + \beta_1 \cdot PC_i + \beta_2 \cdot PK_i \cdot PC_i + \epsilon_i$

[28] The marginal effect is $\frac{\partial DU_i}{\partial PK_i} = \delta + \beta_2 \cdot PC_i$, with standard error $SE = \sqrt{\text{Var}(\delta) + PC_i^2 \cdot \text{Var}(\beta_2) + 2PC_i \cdot \text{Cov}(\delta, \beta_2)}$, and t-statistic $t = \frac{\delta + \beta_2 \cdot PC_i}{SE}$. The only non-significant marginal effect occurs at the minimum value of the price consciousness index (PC = 1). However, this estimate relies on sparse data, as only *one* respondent in our sample reports a price consciousness score of one, and fewer than 40 respondents (<2%) report a score between 1 and 2.

Table 3.3: Marginal effects of PK on DU conditional on price consciousness

	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Price Consciousness	1	3.36	3.91	3.81	4.38	5
Effect of Price Knowledge	0.025	0.092***	0.107***	0.105***	0.121***	0.138***
	(0.032)	(0.012)	(0.012)	(0.012)	(0.015)	(0.020)

Notes: This table reports average and quantile marginal effects of price knowledge on discount utilization. The standard errors shown in parentheses are robust to heteroskedasticity (HC1) and calculated using the delta method. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

benefit those who need the savings the most.

Second, the dependence of the effect on consciousness suggests that information provision alone has limits. For consumers with low price consciousness, simply making prices easier to learn may not lead to behavioral changes, as they lack the intrinsic motivation to act on the signal. This suggests that the effect of price-transparency measures is likely to depend not only on the availability of information, but also on consumers' willingness to use that information when making purchase decisions.

Perishable vs. Non-Perishable Products Discount utilization may naturally vary with product characteristics. Our analysis of discount utilization includes consumers' purchases across 709 products that differ substantially in storability. As emphasized by Hendel and Nevo (2013), intertemporal price discrimination in storable goods markets can arise because firms use temporary sales to price discriminate between consumers who stockpile and those who do not. Their analysis focuses on highly storable goods that can be easily stored for long periods.

To examine whether product storability matters in our setting, we divide all products into two groups: perishable items such as milk, sausages, minced meat, vegetables, and fresh smoothies that typically must be consumed within a month, and non-perishable items such as soda, crispbread, jam, and frozen fish that can be stored for much longer. This classification results in 301 perishable products (42%) and 408 non-perishable products (58%).

Columns 4 and 5 of Table 3.2 present the baseline estimates for perishable and non-perishable products, respectively. We find, as anticipated, that the effect of price knowledge is somewhat stronger for non-perishable products, with an increase of about 0.8 percentage points relative to the aggregate effect reported in Column 1. For perishable products, the estimated effect is 0.8 percentage points lower but remains highly statistically significant.

We report the full specifications including control variables in Table B.1 in Appendix B.1, along with histograms and binned scatterplots for both groups. Neither the histograms nor the scatterplots differ qualitatively across the two categories. Furthermore, the control variables in the full specification (mirroring Column 3 in Table 3.1) retain their expected signs and significance levels. Taken together, these results suggest that while storability amplifies the effect, the relationship between price knowledge and discount utilization extends beyond storable

Discount Utilization for Non-Survey Products As discussed in Section 2.2, a potential concern is reverse causality. Specifically, if recently purchasing products on sale causes consumers to recall those prices more accurately, the positive association between price knowledge and discount utilization would be upward biased, as it would partly reflect the reverse effect of discount utilization on price knowledge.

To address this, we re-estimate the relationship between price knowledge and discount utilization, but we calculate the dependent variable using only the products in the respondents' purchase history that were not included in the survey. Thus, we separate the measurement of price knowledge (based on the 18 surveyed items) from the behavior it predicts (discount utilization across the remaining ~700 items in the purchase history).

The results are presented in Table B.2 in Online Appendix B.2. We find that the association between price knowledge and discount utilization remains robust and highly statistically significant when excluding the survey products. In fact, the estimated coefficient on price knowledge is slightly larger in this specification than in the baseline results. This provides evidence against the reverse causality hypothesis; if price knowledge were partly a function of recalling prices recently paid on discount, we would expect a weaker association between price knowledge of the surveyed products and discount utilization on a different set of products.

Event Window Length In our baseline specification, we use an event window of 3-months when calculating discount utilization. Since the scope for paying lower prices depends on the length of the event window, we estimate our model using different event window lengths. The results are reported in Table B.3 in Online Appendix B.3, and illustrated in Figure 3.3. As expected, the estimated effect of price knowledge increases with the length of the event window. The effect is however positive and statistically significant also for a daily window. We also note that the effect stabilizes when we reach our benchmark event window of 3-months.

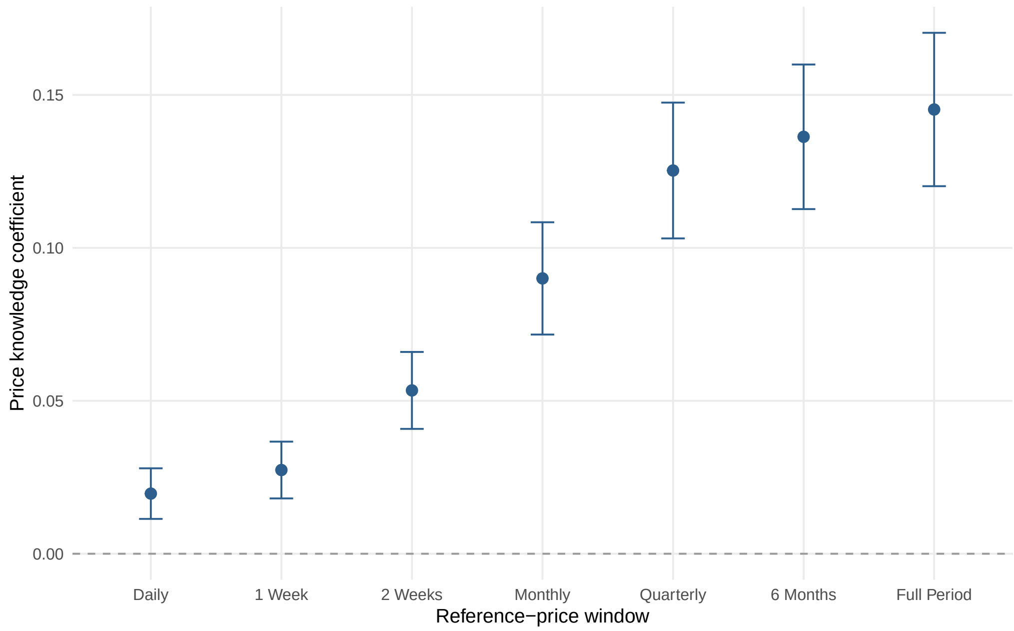


Figure 3.3: Price-knowledge coefficient (with 95% confidence interval) from bivariate regressions of discount utilization (DU) on price knowledge (PK), one per reference-price window. DU is computed as a simple average across the 700+ products in the full purchase sample, re-averaging the reference (regular) price over the window on the horizontal axis; PK is the simple average of recall accuracy across the 24 survey products. Heteroskedasticity-robust (HC1) standard errors.

Sale Timing versus Personal Discounts Discount utilization combines two ways of paying below the average shelf price, which Table B.3 in the Online Appendix separates. Column (2) isolates sale timing by re-valuing each purchase at the same-day shelf price relative to the quarterly average; Column (3), the daily measure, captures the residual gain from paying below that day's posted price through coupons or quantity and expiration-date discounts. The baseline, Column (1), is approximately their sum. Price knowledge raises both components: informed consumers both time sales and secure idiosyncratic discounts.

The Share of Purchases Bought on Sales Above we define discount utilization by the average price paid by a consumer. As a robustness check, we may also examine the effect of price knowledge on the share of purchases made during temporary sales, where the latter is as defined in Section 2.4. We expect a greater share of purchases to occur during temporary sales for more informed consumers. As seen in Table 3.4 this prediction is indeed borne out in the data. Across Columns (1)-(4) we examine the share of purchases over the 18-month preperiod that are purchased when the product is on temporary sale, using successively deepening thresholds for what constitutes a temporary sale. As seen, the coefficient on price knowledge is positive and statistically significant at conventional levels across all specifications. As expected, the estimated effect of price knowledge diminishes as we apply a stricter definition of what counts as a temporary sale. Price knowledge is a continuous measure, so to interpret the magnitude, we need to evaluate it at different points in the distribution. Making a similar thought experiment as in Section 3.1, and using results from Column (1), we estimate that someone with relatively poor price knowledge at $PK = -0.4$ would make around 13.8% of their purchases on sale whereas someone with high price knowledge at $PK = -0.1$ would make around 18.5% of their purchases on sale[29] The corresponding numbers for really deep sales, Column (4), is that an informed customer with $PK = -0.1$ would make around 8.8% of purchases on such deep sales whereas a less informed customer with $PK = -0.4$ would make around 6.7% of purchases on such deep sales.

Table 3.4: Share of purchases on sale regressed on price knowledge

	Share on sale			
	≥3%	≥5%	≥15%	≥25%
	(1)	(2)	(3)	(4)
Price Knowledge (PK)	0.157***	0.160***	0.136***	0.118***
	(0.031)	(0.030)	(0.029)	(0.023)
Constant	0.201***	0.188***	0.147***	0.114***
	(0.008)	(0.008)	(0.008)	(0.006)
Observations	2,027	2,027	2,027	2,027
Dependent variable mean	0.16	0.14	0.11	0.08
R^2	0.03	0.03	0.02	0.03

Notes: The dependent variable is the consumer-level share of purchases made at or below a municipality-specific 90-day rolling-median reference by at least the stated threshold. Price knowledge (PK) is calculated as the negative average percentage absolute deviation between the respondent's recalled prices and true shelf prices. Heteroskedasticity-robust (HC1) standard errors in parentheses. $*$ p $<$ 0.05, $**$ p $<$ 0.01, $***$ p $<$ 0.001.

Price Knowledge Quintiles To assess the functional form of the relationship between price knowledge and discount utilization, and to examine the empirical relevance of the binary distinction between "informed" and "uninformed" consumers often used in search-theoretic models (e.g., Varian, 1980), we re-estimate the model using price knowledge quintile indicators. The results are reported in Table B.4 in Appendix B.4.

The results show a clear monotonic positive relationship between price knowledge and discount utilization: discount utilization increases steadily from one price knowledge quintile to the next. This suggests that the relationship is indeed more gradual than a

[29]These predicted purchase shares may appear low relative to the fact that, across product-chain combinations, products are on sale on 16.0% of days (see Table 2.1). The comparison is not direct, however, since the predicted shares are purchase-based, whereas the sale incidence is measured across available product-chain-day observations. Two features of the data are particularly relevant. First, as reported in Table A.4 in Appendix A, respondents make a large share of their purchases in the discount chain, where products are on sale on only 10.8% of days. Second, frequently purchased products are less likely to be on sale: at the product level, purchase frequency is negatively correlated with sale incidence (see Table A.5 in Appendix A).

simple binary distinction between informed and uninformed consumers would imply, and supports the linear specification used in our main analysis as a reasonable and robust approximation. Another insight from this exercise is that we also confirm that our main results are not driven by potential large outliers.

3.3 The Gains Associated with Price Knowledge

We now turn to a quantification of the direct gains associated with price knowledge. We take current prices and quantities as given and thereby disregard additional equilibrium effects resulting from the fact that a greater share of informed individuals would be expected to strengthen competition.

To do this, we rely on the sales-weighted version of discount utilization (see Equation 2.8), which allows us to quantify the savings a consumer is expected to make by buying their current basket of goods at more opportune times. The simple regression results for this specification are shown in Column 6 of Table 3.2. The estimated coefficient on price knowledge is smaller in magnitude than the coefficient in the simple-average specification by 5.5 percentage points. This difference likely arises because consumers tend to spend more on products with lower price variation, reducing the scope for finding bargains in a sales-weighted basket.[30]

We report the full results, including control variables, in Table B.6 in Appendix B.5. The results are qualitatively similar to Column 3 in Table 3.1, except that primarily shopping at discount stores now exhibits a detectable negative effect. The coefficients are generally smaller in magnitude across covariates.

If we consider a nuclear family consisting of a mother (31–60 years), father (31–60 years), daughter (10–13 years), and son (2–5 years), their reference budget for groceries in 2024 according to SIFO (2024) amounts to 173 640 NOK per year. With fixed quantities, a 30 percentage-point increase in price knowledge—a substantial but feasible change given its empirical distribution (see Figure 3.1)—is associated with yearly household savings of 3629 NOK (2.1%).

It is also instructive to consider the savings that a perfectly informed consumer is expected to make relative to a consumer who makes their purchases randomly throughout the quarter, akin to the uninformed consumers in Varian (1980). Assuming that a linear extrapolation of the relation between price knowledge and discount utilization to the point of $PK = 0$ is reasonable, the discount utilization of a perfectly informed consumer is given by the regression constant in Column 6 in Table 3.2, which is 4%. This suggests that a perfectly informed consumer shopping on behalf of the nuclear family pays 4% less for its groceries than someone who does not time their purchases with temporary price

[30]Table A.1 in Appendix A.2 shows that the average basket coefficient of variation is lower when we weigh products according to spending as opposed to using a simple average.

discounts, amounting to yearly household savings of 6925 NOK.

The potential household-level savings related to improvements in price knowledge are quantitatively noteworthy but modest. However, from a policy perspective, improvements in price knowledge add up across households and over time. Low-cost initiatives that persistently improve the price knowledge of the broader population could lead to large welfare gains. Furthermore, a core search-theoretic prediction—whose foundations we argue this article provides empirical support for—is that an increase in the share of informed consumers leads to lower equilibrium prices. The direct savings from better price knowledge that we calculate here should therefore be seen as a lower bound on the total savings to consumers.[31]

Finally, it is important to note that these potential savings are not uniformly distributed; they depend heavily on the price volatility of the consumer's typical basket. When we re-estimate the interaction model using sales-weighted discount utilization (see Table B.7 in Appendix B.5), we find substantial heterogeneity in the economic value of being informed. For consumers with below-median basket price variation, a 30 percentage-point increase in price knowledge is associated with yearly household savings of 2396 NOK (1.4%). However, for consumers with above-median variation, the savings associated with the same increase in knowledge nearly double to 4428 NOK (2.6%). The gap is even wider when comparing a perfectly informed consumer to one who disregards sales entirely: the theoretical savings potential rises from 4167 NOK (2.4%) for low-variance baskets to 9203 NOK (5.3%) for high-variance baskets. This underscores that the monetary gains from price knowledge are highest for households that consume products subject to frequent or deep price fluctuations.

4 Drivers of Price Knowledge

We now turn to an examination of the drivers of price knowledge itself. The level of consumer price knowledge of groceries has been documented in many studies (see, e.g., Gabor and Granger, 1961; Conover, 1986; Vanhuele and Drèze, 2002). However, previous studies (e.g., Aalto-Setälä and Raijas, 2003; Loy et al., 2020) have focused on how price recall relates to measures that can be obtained from consumer surveys, such as demographics, reported shopping habits, and subjective price consciousness, or that are easily observed in stores by researchers, such as current product prices and their dispersion. We extend this literature by relating price knowledge to variables that are derived from transaction data. These include historical purchasing frequencies and spending on specific

[31] Most people in our sample buy from one of the leading discount chains in Norway, which has temporary sales less often than other grocery chains. The gains to price knowledge can only be expected to be larger for the general population, some of whom buy groceries from chains with more sales activity, further supporting the lower bound interpretation.

products, the number of store and chain visits, and other features of the respondents' shopping baskets, such as product variety.

Results from estimating Equation 2.2 with OLS are shown in Table 4.1. Standard errors are clustered at the individual level (2028 clusters) to account for within-respondent dependence across price recalls. This choice reflects the sampling design: respondents are the sampled units, while products and chains are fixed by design. Accordingly, inference is directed at the consumer population rather than a broader population of products or chains. Column 1 includes only variables derived from the survey, while Column 2 adds variables constructed from the transaction data, which invites some caution in the interpretation of these coefficients. For instance, having purchased a product increases the knowledge of the price of that product, but on the other hand increased price knowledge might also affect the purchasing decision.

The transaction-derived measures contribute substantially to explaining price knowledge ($R^2 = 8\%$ vs. $R^2 = 1\%$), underscoring the importance of revealed shopping behavior for understanding price knowledge. We therefore focus the following discussion on the results in Column 2, which represents our preferred specification.

Table 4.1: Price knowledge regressions

	Price Knowledge	
	(1)	(2)
Male	-0.015*** (0.004)	-0.015*** (0.004)
Single Person Household	-0.002 (0.007)	-0.001 (0.007)
Has Children	0.003 (0.004)	-0.001 (0.004)
Full-Time Employment	-0.005 (0.005)	-0.004 (0.004)
Higher Education	-0.013** (0.004)	-0.012** (0.004)
Rural	0.000 (0.004)	-0.002 (0.004)
Household Income > 950' NOK	0.010* (0.004)	0.008* (0.004)
Age: 40-59 Years	0.008 (0.005)	0.006 (0.004)
Age: Above 60 Years	0.004 (0.006)	0.002 (0.006)
# Frequented Grocery Chains	0.002 (0.001)	0.002 (0.001)
Mainly uses discount stores	0.006 (0.006)	0.004 (0.006)
Main Grocery Shopper	0.013** (0.004)	0.014** (0.004)
Supermarket Chain	-0.017*** (0.001)	0.033*** (0.002)
Private Label	0.044*** (0.004)	0.018*** (0.004)
Three Customer Program Memberships	0.011** (0.004)	0.010** (0.004)
Price Sensitivity	0.001 (0.003)	0.000 (0.003)
Price Learning Effort	0.016*** (0.002)	0.015*** (0.002)
Average Weekly Chain Visits		0.006*** (0.001)
Average Weekly Chain Variety (per 10 products)		-0.010*** (0.002)
Category Size (per 10 products)		-0.029*** (0.001)
Price Coefficient of Variation		-0.120*** (0.019)
Purchased		0.068*** (0.003)
Purchased × Sales in 1000 NOK		0.043*** (0.003)
Purchased × Months Since Last Purchase		-0.001** (0.000)
Constant	-0.344*** (0.013)	-0.308*** (0.014)
Observations	69,408	69,408
Dependent variable mean	-0.27	-0.27
R^2	0.01	0.08

Notes: This table reports results from estimation of Equation 2.2. Price knowledge is calculated as the negative percentage absolute deviation between the respondents' recalled price and the true shelf price. Price sensitivity and learning effort are calculated as the factor-weighted averages of Likert-scale survey items 13 and 47, respectively. The average weekly chain variety is calculated as the average number of unique products purchased at each chain per week over the sample period. Category size is the chain-specific number of close substitutes for the product as defined by the grocery chain. The price coefficient of variation is calculated for each product × chain combination over the sample period. Purchased is an indicator of whether the respondent bought the product at a given chain during the sample period, and sales is the consumer's total product–chain spending. Column 1 includes individual characteristics and indicators for private labels and supermarket chain, and Column 2 adds variables derived from transaction data. The standard errors shown in parentheses are clustered by individual. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Demographics We see that men have poorer price knowledge than women. This replicates the findings of Loy et al. (2020) and Estelami et al. (2001), as opposed to Estelami (1998); Jensen and Grunert (2014); Kenning et al. (2011) and Wakefield and Inman (1993) who find no difference between men and women. We find no statistically precise relation between price knowledge and being single, having children, being in full-time employment, living in a rural area, or age. Consumers with higher education are less knowledgeable about grocery prices, but even though we might expect higher income to reduce price consciousness (Estelami et al., 2001; Estelami and Lehmann, 2001; Wakefield and Inman, 1993; Loy et al., 2020), high income is, in our sample, associated with slightly more price knowledge, which is also reflected in the raw correlation. Small and sometimes counterintuitive effects of age, income, and employment on price knowledge are not new in the literature. Gaston-Breton and Raghubir (2012) points out that these sociodemographic variables are all correlated with both the ability and motivation to memorize prices in ways that can cancel out, leaving their relation to price knowledge ambiguous. For example, older consumers might have more time and incentives to think about grocery prices but also poorer memory. Similarly, high income or education can be correlated with both good memory and lower motivation to focus on grocery prices.

Consumer statements about shopping habits Price knowledge does not appear to be related, all else equal, to the number of different grocery chains respondents regularly visit or whether they primarily shop at discount stores. However, direct and indirect markers of price consciousness do independently predict price knowledge; for example, being the sole and primary grocery shopper of the household (as opposed to sharing or not having the main responsibility), or being a member of all three main programs (as opposed to only one or two), are both associated with higher price knowledge.[32]

Gathering information about grocery prices is strongly related to price knowledge: a one-point increase in the price learning effort index (see questions 4–7 in Table A.3 in Appendix A), which is less than the interquartile range, has a stronger association with price knowledge than either primary shopping responsibility or program membership. Price sensitivity, represented by questions 1–3 in Table A.3 in Appendix A, is almost completely unrelated to price knowledge, conditional on price learning effort.

The finding that price learning effort has an independent effect on price knowledge suggests that improving access to price information could be an effective way to raise consumer awareness. If knowledge were primarily driven by price sensitivity, shopping patterns, or demographics, such interventions would likely have limited effect. In the Norwegian grocery market, where there is currently no formal price transparency regulation

[32] Note that these three programs cover more than 95% of the Norwegian grocery market. The primary tangible benefit of participating in all three main membership programs is access to personal price discounts, signaling price consciousness, consistent with the estimated positive effect.

or comprehensive comparison tools, collecting price information remains relatively costly for consumers. Policies that reduce these costs could both encourage more consumers to seek out prices and increase the overall level of price knowledge in the market.

Transaction data Turning now to the variables derived from the transaction data whose estimates are reported in Column 2 of Table 4.1, we find that more frequent visits to the stores of the relevant chain are related to better price recall. This is consistent with what Loy et al. (2020) find using self-reported shopping frequencies. The weekly average number of unique products purchased at each chain, a variable that, to our knowledge, has not been used in the literature before, is negatively associated with price knowledge. One explanation for the negative association between basket variety and price knowledge could be that it is difficult to keep track of the prices of any given product if one tends to buy a large variety of products.

Price knowledge is decreasing in category size, defined by the number of close substitutes according to the grocery chain: Increasing category size by 10 units is associated with 3 percentage-points lower price recall accuracy in our sample. The findings are consistent with Vanhuele and Drèze (2002), who refer to the number of SKUs in a category as "clutter" and argue that increased category complexity increases the risk of confusion and poor recall.

Similarly, more price variation over time makes it harder to stay informed and predicts poorer price recall. An interquartile-range increase in the price coefficient of variation (0.09) is associated with 1.1 percentage-points lower price recall accuracy. Loy et al. (2020) and Vanhuele and Drèze (2002) find a related result that more price variation within the product's category is associated with poorer recall. However, our long purchasing histories allow us to see how variation over time in the price of the product itself affects learning and recall, as opposed to the cross-sectional price variation within the product's category.

Spending 1000 NOK more on the product over the last 18 months is associated with 4.3 percentage-points more accurate price recall. This is conditional on having ever bought the product and how long ago that was, so the effect of spending on price recall accuracy is not driven by familiarity with the product, but rather how important it is for the consumer to keep track of its price.[33] In other words, consumers make sure to stay more informed about products for which the potential savings are the greatest. This is in line with the predictions of Stigler (1961) concerning the incentives for search, i.e., price knowledge acquisition.

The respondents are much more informed about products that they have experience buying: Recall accuracy is 7 percentage points higher on average for products that the consumer has purchased before. What is somewhat surprising is that this effect seems to be very persistent. Even after 12 months, the effect associated with buying the product

[33] Replacing time since last purchase with purchasing frequency produces similar results.

has only diminished by approximately $12 \times (-0.001) = -1.2$ percentage points. It seems that once a consumer buys a product, they continue to stay informed about the price of that product even if they buy it infrequently.

The marketing literature has long emphasized widespread consumer inattention and limited price knowledge but has provided less evidence on how price knowledge relates to objectively measured purchasing experiences. By combining survey and transaction data, we show that price knowledge increases with the frequency of store visits, recent purchase experience, and spending on the product in question, and decreases with basket variety and infrequent shopping—suggesting that consumers acquire price knowledge through repeated exposure and focused attention. In contrast, demographics and survey-based indicators are comparatively less important in explaining price knowledge—the R^2 increases from 1% to 8% when transaction-derived variables are added. This mirrors research in marketing showing that observed purchasing behavior is substantially more predictive than individual characteristics for price knowledge (Gupta and Chintagunta, 1994; Islam et al., 2022).

5 Concluding Comments

This paper provides the first direct evidence that consumers with better knowledge of grocery prices systematically pay lower prices. Using a unique dataset that links survey-based price recall measures to detailed, individual-level grocery purchase histories, we quantify the relationship between price knowledge and discount utilization. We show that consumers with more accurate knowledge of prices pay less, not by purchasing generally cheaper products or frequenting cheaper stores, but by timing their purchases to take advantage of price variation over time. Our results also show that the benefits of price knowledge are economically meaningful: a 30 percentage-point improvement in price recall is associated with an average price reduction of about 4 percent.

Our findings provide direct empirical support for a central mechanism in search-theoretic models of sales, such as Varian (1980), in which informed consumers impose a competitive constraint on firms by responding strategically to price variation. By demonstrating that price knowledge is associated with paying lower prices—also after controlling for potential confounders—we provide micro-level support for a key theoretical assumption that underlies a broad class of models in industrial organization and marketing.

We find stronger effect of price knowledge on discount utilization for consumers with high price consciousnesses. To the extent that disadvantaged groups are more price conscious, as our data suggests, population-wide gains in price knowledge may disproportionately benefit those who need the savings the most.

By linking survey-based recall data with transaction records, we provide a more detailed analysis of the determinants of price knowledge than survey data alone can offer. We find substantial heterogeneity between consumers, with greater price knowledge among those who visit stores frequently, have recently purchased the product, and spend more on it. In contrast, varied or infrequent shopping, larger product categories, and greater price variation are associated with lower knowledge.

The strong association we observe between price learning effort and recall accuracy demonstrates that price knowledge is not solely determined by fixed traits, but is responsive to information-gathering activities. Taken together, our results support the idea that price disclosure regulations and other policies that lower search costs (by making it easier for consumers to learn prices) have the potential to effectively raise price knowledge and improve market efficiency.

Bibliography

Aalto-Setälä, V. and Raijas, A. (2003). Actual market prices and consumer price knowledge. *Journal of Product & Brand Management*, 12(3):180–192.

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2022). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1):1–35.

Alford, B. L. and Biswas, A. (2002). The effects of discount level, price consciousness and sale proneness on consumers' price perception and behavioral intention. *Journal of Business Research*, 55(9):775–783.

Anderson, E., Malin, B. A., Nakamura, E., Simester, D., and Steinsson, J. (2017). Informational rigidities and the stickiness of temporary sales. *Journal of Monetary Economics*, 90:64–83.

Ater, I. and Rigbi, O. (2023). Price transparency, media, and informative advertising. *American Economic Journal: Microeconomics*, 15(1):1–29.

Bronnenberg, B. J., Dube, J.-P., Gentzkow, M., and Shapiro, J. M. (2013). Do pharmacists buy bayer? sophisticated shoppers and the brand premium. *Kilts Center for Marketing at Chicago Booth–Nielsen Dataset Paper Series*, pages 1–017.

Brown, J. R. and Goolsbee, A. (2002). Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of Political Economy*, 110(3):481–507.

Brown, Z. Y. (2019). Equilibrium effects of health care price information. *The Review of Economics and Statistics*, 101(4):699–712.

Cattaneo, M. D., Crump, R. K., Farrell, M. H., and Feng, Y. (2019). Binscatter regressions.

Conover, J. N. (1986). The accuracy of price knowledge: Issues in research methodology. *ACR North American Advances*, NA-13.

Diamond, P. A. (1971). A model of price adjustment. *Journal of Economic Theory*, 3(2):156–168.

Dickson, P. R. and Sawyer, A. G. (1990). The price knowledge and search of supermarket shoppers. *Journal of Marketing*, 54(3):42–53.

Dinerstein, M., Einav, L., Levin, J., and Sundaresan, N. (2018). Consumer price search and platform design in internet commerce. *American Economic Review*, 108(7):1820–1859.

Ellison, G. and Ellison, S. F. (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica*, 77(2):427–452.

Ellison, G. and Wolitzky, A. (2012). A search cost model of obfuscation. *The RAND Journal of Economics*, 43(3):417–441.

Estelami, H. (1998). The price is right...or is it? Demographic and category effects on consumer price knowledge. *Journal of Product & Brand Management*, 7(3):254–266.

Estelami, H. and Lehmann, D. R. (2001). The impact of research design on consumer price recall accuracy: An integrative review. *Journal of the Academy of Marketing Science*, 29(1):36–49.

Estelami, H., Lehmann, D. R., and Holden, A. C. (2001). Macro-economic determinants of consumer price knowledge: A meta-analysis of four decades of research. *International Journal of Research in Marketing*, 18(4):341–355.

Evanschitzky, H., Kenning, P., and Vogel, V. (2004). Consumer price knowledge in the German retail market. *Journal of Product & Brand Management*, 13(6):390–405.

Fabra, N. and Reguant, M. (2020). A model of search with price discrimination. *European Economic Review*, 129:103571.

Gabaix, X. and Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics*, 121(2):505–540.

Gabor, A. and Granger, C. W. J. (1961). On the price consciousness of consumers. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 10(3):170–188.

Gaston-Breton, C. and Raghubir, P. (2012). Opposing effects of sociodemographic variables on price knowledge. *Marketing Letters*, 24(1):29–42.

Gupta, S. and Chintagunta, P. K. (1994). On using demographic variables to determine segment membership in logit mixture models. *Journal of Marketing Research*, 31(1):128–136.

Hendel, I. and Nevo, A. (2006). Measuring the implications of sales and consumer inventory behavior. *Econometrica*, 74(6):1637–1673.

Hendel, I. and Nevo, A. (2013). Intertemporal price discrimination in storable goods markets. *American Economic Review*, 103(7):2722–2751.

Hong, H. and Shum, M. (2006). Using price distributions to estimate search costs. *The RAND Journal of Economics*, 37(2):257–275.

Honka, E., Hortaçsu, A., and Wildenbeest, M. (2019). Empirical search and consideration sets. In *Handbook of the Economics of Marketing*, volume 1, pages 193–257. Elsevier.

Islam, T., Meade, N., Carson, R. T., Louviere, J. J., and Wang, J. (2022). The usefulness of socio-demographic variables in predicting purchase decisions: Evidence from machine learning procedures. *Journal of Business Research*, 151:324–338.

Jensen, B. B. and Grunert, K. G. (2014). Price knowledge during grocery shopping: What we learn and what we forget. *Journal of Retailing*, 90(3):332–346.

Kenning, P., Hartleb, V., and Schneider, H. (2011). An empirical multimethod investigation of price knowledge in food retailing. *International Journal of Retail & Distribution Management*, 39(5):363–382.

Le Boutillier, J., Le Boutillier, S. S., and Neslin, S. A. (1994). A replication and extension of the Dickson and Sawyer price-awareness study. *Marketing Letters*, 5(1):31–42.

Loy, J.-P., Ceynowa, C., and Kuhn, L. (2020). Price recall: Brand and store type differences. *Journal of Retailing and Consumer Services*, 53:101990.

Luco, F. (2019). Who benefits from information disclosure? The case of retail gasoline. *American Economic Journal: Microeconomics*, 11(2):277–305.

McGoldrick, P. J. and Marks, H. J. (1987). Shoppers awareness of retail grocery prices. *European Journal of Marketing*, 21(3):63–76.

Milyo, J. and Waldfogel, J. (1999). The effect of price advertising on prices: Evidence in the wake of 44 liquormart. *American Economic Review*, 89(5):1081–1096.

Montag, F., Sagimuldina, A., and Winter, C. (2024). When does mandatory price disclosure lower prices? Evidence from the German fuel market. *Journal of Political Economy Microeconomics*.

Morgan, J., Orzen, H., and Sefton, M. (2006). An experimental study of price dispersion. *Games and Economic Behavior*, 54(1):134–158.

Myatt, D. P. and Ronayne, D. (2026). Asymmetric models of sales. *American Economic Journal: Microeconomics*, 18(1):281–311.

Nakamura, E. and Steinsson, J. (2008). Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics*, 123(4):1415–1464.

Norges Bank (2024). Retail payment services 2023. Norges Bank Paper 1/2024, Norges Bank. https://www.norges-bank.no/en/news-events/publications/Norges-Bank-Papers/2024/memo-1-24-payment-services/.

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C., and Yontcheva, B. (2020). Information and price dispersion: Theory and evidence. *International Economic Review*, 61(2):871–899.

Rossi, F. and Chintagunta, P. K. (2016). Price transparency and retail prices: Evidence from fuel price signs in the italian highway system. *Journal of Marketing Research*, 53(3):407–423.

Salop, S. and Stiglitz, J. (1977). Bargains and ripoffs: A model of monopolistically competitive price dispersion. *The Review of Economic Studies*, 44(3):493–510.

Shelegia, S. and Wilson, C. M. (2021). A generalized model of advertised sales. *American Economic Journal: Microeconomics*, 13(1):195–223.

SIFO (2024). Reference Budget 2024. https://oda.oslomet.no/oda-xmlui/handle/11250/3128175?locale-attribute=en, Accessed: 2025-03-19.

Sorensen, A. T. (2000). Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy*, 108(4):833–850.

Spitzer, S. (2020). Biases in health expectancies due to educational differences in survey participation of older europeans: Its worth weighting for. *The European Journal of Health Economics*, 21:573–605.

Stahl, D. O. (1989). Oligopolistic pricing with sequential consumer search. *The American Economic Review*, pages 700–712.

Statistics Norway (2024a). 07459: Population, by sex and one-year age groups (m) 1986–2025. Accessed June 2025.

Statistics Norway (2024b). 08921: Educational attainment, by county, age and sex (c) 1980–2024. Accessed June 2025.

Statistics Norway (2024c). Hva bruker vi tiden vår på? https://www.ssb.no/kultur-og-fritid/tids-og-mediebruk/statistikk/tidsbruksundersokelsen/artikler/hva-bruker-vi-tiden-var-pa. [Accessed on: January 10, 2026].

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

Tappata, M. (2009). Rockets and feathers: Understanding asymmetric pricing. *The RAND Journal of Economics*, 40(4):673–687.

Vanhuele, M. and Drèze, X. (2002). Measuring the price knowledge shoppers bring to the store. *Journal of Marketing*, 66(4):72–85.

Varian, H. R. (1980). A Model of Sales. *The American Economic Review*, 70(4):651–659.

Villas-Boas, J. M. (1995). Models of competitive price promotions: Some empirical evidence from the coffee and saltine crackers markets. *Journal of Economics & Management Strategy*, 4(1):85–107.

Wakefield, K. L. and Inman, J. (1993). Who are the price vigilantes? An investigation of differentiating characteristics influencing price information processing. *Journal of Retailing*, 69(2):216–233.

Winer, R. S. (1986). A reference price model of brand choice for frequently purchased products. *Journal of Consumer Research*, 13(2):250–256.

Online Appendix to "Do informed consumers pay less"

Richard Friberg, Emil M.S. Halseth, Frode Steen, Simen A. Ulsaker

A Data Cleaning and Descriptive Statistics

A.1 Data Cleaning

A few price recalls are so extreme that they are likely erroneous entries or, at least, not based on a serious attempt at recalling the right price. We filter out all individuals who make a price recall that lies outside the 10–300 percent range of the true shelf price. Furthermore, we exclude respondents who do not make at least 18 out of the 36 possible product recalls on account of likely lack of focus and effort. This leaves us with 2028 respondents, 2 chains, and 24 products, and 72 392 observations, a loss of 11% of respondents and 7% of recalled prices.[34]

A.2 Descriptive Statistics

The primary outcome and explanatory variables that we use in the individual-level analysis are summarized in Table A.1, and the individual × chain × product-level, or recall-level, variables are summarized in Table A.2. We include the mean, standard deviation, minimum, maximum, and quartile values. The variables that overlap between the tables have identical means, largely because most respondents answer all 18 price questions, but the recall-level variables have different quantiles and higher standard deviations.

Table A.2 shows that the mean PK of the recalled prices is -0.27, meaning that the average price recall either over- or underestimates the true shelf price by 27%. In addition to the PK defined in Equation 2.3 above, we include the share of recalled prices within different error bounds. Only 1% of the product recalls are exactly correct at the two-decimal level, i.e., approximately 724 out of the 72 392 recalled prices.

Loy et al. (2020), who surveyed grocery shoppers in two stores in Northern Germany, found that the average PK in their sample is 35% and that around 14% of recalled prices

[34]Fewer observations end up being used in the regression analysis due to missing values for some of the survey variables. Finally, we also checked for respondents with very few transactions in their purchasing histories, but removing these respondents does not meaningfully alter the results.

Table A.1: Summary statistics: individual-level variables

	Mean	SD	Min	P25	Median	P75	Max
Panel A: Discount Utilization and Purchase-Based Outcomes							
Discount Utilization (Simple)	0.040	0.037	−0.044	0.013	0.033	0.059	0.272
Discount Utilization (Sales-Weighted)	0.021	0.027	−0.073	0.004	0.014	0.031	0.219
By reference-price window (simple):							
Daily	0.002	0.015	−0.136	−0.002	0.000	0.006	0.221
1 Week	0.007	0.017	−0.054	−0.001	0.002	0.011	0.229
2 Weeks	0.015	0.022	−0.050	0.001	0.010	0.023	0.232
Monthly	0.027	0.031	−0.059	0.006	0.020	0.041	0.262
6 Months	0.043	0.040	−0.048	0.015	0.036	0.064	0.265
Full Period	0.044	0.043	−0.115	0.014	0.037	0.067	0.262
Discount Utilization (Quarterly, Timing Only)	0.037	0.031	−0.037	0.014	0.032	0.055	0.186
Share of purchases on sale:							
≥3%	0.158	0.079	0	0.104	0.142	0.191	0.755
≥5%	0.144	0.077	0	0.094	0.128	0.174	0.735
≥15%	0.110	0.071	0	0.063	0.092	0.139	0.714
≥25%	0.082	0.058	0	0.042	0.067	0.106	0.551
Basket Coefficient of Variation (Simple)	0.11	0.01	0.07	0.10	0.11	0.12	0.17
Basket Coefficient of Variation (Sales-Weighted)	0.10	0.02	0.04	0.08	0.09	0.11	0.16
Panel B: Price Knowledge							
Price Knowledge (PK)	−0.27	0.08	−0.70	−0.32	−0.27	−0.22	−0.04
Price Bias (PD)	0.02	0.13	−0.44	−0.06	0.01	0.09	0.60
Share Over-estimated	0.46	0.15	0	0.36	0.44	0.56	0.92
Share Under-estimated	0.53	0.15	0.06	0.44	0.53	0.64	1
Share Correct	0.01	0.04	0	0	0	0	0.53
Share Within 1%	0.06	0.07	0	0.03	0.06	0.08	0.56
Share Within 5%	0.17	0.11	0	0.08	0.14	0.22	0.78
Share Within 10%	0.29	0.13	0	0.19	0.28	0.36	0.89
Share Within 20%	0.50	0.15	0.06	0.40	0.50	0.58	1
Panel C: Demographics							
Male	0.38	0.49	0	0	0	1	1
Single Person Household	0.12	0.33	0	0	0	0	1
Has Children	0.42	0.49	0	0	0	1	1
Full-Time Employment	0.61	0.49	0	0	1	1	1
Higher Education	0.71	0.45	0	0	1	1	1
Rural	0.30	0.46	0	0	0	1	1
Household Income > 950' NOK	0.54	0.50	0	0	1	1	1
Age: Below 40 Years	0.23	0.42	0	0	0	0	1
Age: 40-59 Years	0.47	0.50	0	0	0	1	1
Age: Above 60 Years	0.30	0.46	0	0	0	1	1
Panel D: Shopping Behavior and Attitudes							
# Frequented Grocery Chains	2.98	1.39	1	2	3	4	9
Mainly uses discount stores	0.86	0.35	0	1	1	1	1
Average Weekly Store Visits	2.90	1.80	0.09	1.69	2.58	3.65	16.53
Main Grocery Shopper	0.74	0.44	0	0	1	1	1
Average Weekly Variety	27.98	13.23	1.39	18.80	26.15	35.09	102.72
Three Customer Membership Programs	0.48	0.50	0	0	0	1	1
Price Consciousness	3.81	0.75	1	3.36	3.91	4.38	5
Price Sensitivity	4.12	0.79	1	3.74	4.26	4.74	5
Price Learning Effort	3.52	0.89	1	2.95	3.62	4.22	5

Notes: Summary statistics for the 2,028 respondents in the sample. Discount Utilization is the percent the price paid falls below a product's reference shelf price, averaged over the consumer's chain-by-product cells with equal (Simple) or spending (Sales-Weighted) weights; the reference is quarterly (baseline), recomputed over the indicated window in the window rows, and Timing-Only uses the posted shelf price instead of the price paid. Share of purchases on sale rows give the share of purchases at least the stated percent below the product's 90-day rolling-median shelf price; Basket Coefficient of Variation is the consumer's mean shelf-price CV across purchased products. Price Knowledge (PK) and Price Bias (PD) are the negative mean absolute and the mean signed percentage deviation of recalled from true shelf prices; Share Over-, Under-estimated, Correct, and Within are the shares of recalls too high, too low, exact, or within the stated percent. Panel C rows are demographic indicators. In Panel D, Average Weekly Store Visits and Variety are mean weekly trips and distinct products bought, Three Customer Membership Programs flags at least three memberships, and Price Consciousness, Sensitivity, and Learning Effort are 1–5 survey indices. SD, P25, P75: standard deviation and 25th/75th percentiles.

Table A.2: Summary statistics: individual × chain × product-level variables

	Mean	SD	Min	P25	Median	P75	Max
Panel A: Price Knowledge							
Price Knowledge (PK)	−0.27	0.26	−1.99	−0.38	−0.20	−0.08	0
Price Bias (PD)	0.02	0.37	−0.90	−0.22	−0.03	0.17	1.99
Share Over-estimated	0.46	0.50	0	0	0	1	1
Share Under-estimated	0.53	0.50	0	0	1	1	1
Share Correct	0.01	0.11	0	0	0	0	1
Share Within 1%	0.06	0.24	0	0	0	0	1
Share Within 5%	0.17	0.37	0	0	0	0	1
Share Within 10%	0.29	0.45	0	0	0	1	1
Share Within 20%	0.50	0.50	0	0	1	1	1
Panel B: Demographics							
Male	0.38	0.49	0	0	0	1	1
Single Person Household	0.12	0.33	0	0	0	0	1
Has Children	0.42	0.49	0	0	0	1	1
Full-Time Employment	0.61	0.49	0	0	1	1	1
Higher Education	0.71	0.45	0	0	1	1	1
Rural	0.30	0.46	0	0	0	1	1
Household Income > 950’ NOK	0.54	0.50	0	0	1	1	1
Age: Below 40 Years	0.23	0.42	0	0	0	0	1
Age: 40-59 Years	0.47	0.50	0	0	0	1	1
Age: Above 60 Years	0.30	0.46	0	0	0	1	1
Panel C: Shopping Behavior and Attitudes							
# Frequented Grocery Chains	2.98	1.39	1	2	3	4	9
Mainly uses discount stores	0.86	0.35	0	1	1	1	1
Main Grocery Shopper	0.74	0.44	0	0	1	1	1
Three Customer Membership Programs	0.48	0.50	0	0	0	1	1
Price Consciousness	3.81	0.75	1	3.36	3.91	4.38	5
Price Sensitivity	4.12	0.79	1	3.74	4.26	4.74	5
Price Learning Effort	3.52	0.89	1	2.95	3.62	4.22	5
Panel D: Chain, Product, and Purchase Characteristics							
Supermarket Chain	0.50	0.50	0	0	0.50	1	1
Private Label	0.08	0.28	0	0	0	0	1
Average Weekly Chain Visits	1.45	1.56	0	0.23	0.99	2.21	15.56
Average Weekly Chain Variety	13.99	14.37	0	1.42	9.67	23.01	89.13
Category Size	19.34	19.11	2	7	12.92	25	128
Price Coefficient of Variation	0.11	0.06	0.01	0.06	0.11	0.15	0.33
Sales	92.33	363.26	0	0	0	31.56	18492.96
Has Purchased Product at Chain	0.31	0.46	0	0	0	1	1
Months Since Last Purchase	4.68	4.80	0	0.73	2.73	7.50	17.53

Notes: Summary statistics at the individual × chain × product (recall) level for the 72,392 price recalls. Price Knowledge (PK) and Price Bias (PD) are the negative absolute and the signed percentage deviation of the recalled from the true shelf price; Share Over-, Under-estimated, Correct, and Within flag a recall too high, too low, exact, or within the stated percent. Panel B rows are demographic indicators. In Panel C, Three Customer Membership Programs flags at least three memberships and Price Consciousness, Sensitivity, and Learning Effort are 1–5 survey indices. In Panel D, Supermarket Chain and Private Label flag the full-service chain and a private-label product; Average Weekly Chain Visits and Variety are mean weekly trips to and distinct products bought at the chain; Category Size is the count of products in the category, Price Coefficient of Variation the product’s daily shelf-price SD over its mean, Sales the consumer’s total spending on the product at the chain, Has Purchased Product at Chain a purchase indicator, and Months Since Last Purchase the time since the last purchase. SD, P25, P75: standard deviation and 25th/75th percentiles.

lie within 5% of the true price. This is a slightly poorer performance than our sample, with an average PK of 27% and a share of recalled prices within 5% of the true price of 17%. The magnitudes are fairly similar, but of course, the results are sensitive to the products asked about (see Figure A.1a), the population surveyed, and the setting in which the respondents were surveyed (see Evanschitzky et al., 2004 for a discussion).

Equation 2.3 measures price knowledge without distinguishing between over- and underestimation. To measure the direction of the respondents' price bias, we can calculate:

$$\mathrm{PD}_{ijc} = \frac{P^{\mathrm{Recall}}_{ijc} - P^{\mathrm{Shelf}}_{jcmt}}{P^{\mathrm{Shelf}}_{jcmt}}$$

The respondent's overall price bias is then given by $\mathrm{PD}_i = \frac{1}{K_i}\sum_{k=1}^{K_i}\mathrm{PD}_{ijc}$. Finally, averaging indicators of the sign of PD_{ijc} lets us calculate the respondent's share of over- and underestimations.

The price recalls overestimate the correct prices by 2% on average. Consistent with the findings of Aalto-Setälä and Raijas (2003), who measured price knowledge among Finnish grocery shoppers, the bias of the respondents as a group is low, despite substantial errors on any given recall. However, this breaks down somewhat at the chain and product level (see Figure A.1b). Although respondents only slightly underestimate the prices of the discount chain on average ($PD = -1.5\%$), they overestimate the prices of the supermarket chain by 5.2%. This likely reflects the considerable efforts that discount chains in Norway, who serve over 60% of the market, put into building and maintaining a favorable price image. Supermarkets, on the other hand, focus more on service, assortment size, and quality. Furthermore, price bias varies substantially across products. For example, the respondents are far off target, on average, for strawberry jam (60%), crispbread (49%), potato chips (33%), and frozen cod (−40%), whereas the bias is quite low for eggs (−2%), smoothies (−2%), and 10 other products with less than 10% of price bias.

The share of overestimated recalls is 46%, and the share of underestimated recalls is 53%, despite the average price bias being positive. This stems from the overestimated price recalls being larger in magnitude, which is also reflected in the median price bias (−3%) being lower than the average (2%). The shares of over- and underestimation differ substantially from the findings of Loy et al. (2020), who find that 56% of price recalls overestimate the true price, and only 36% underestimate true prices. However, they find that 8% of recalls are exact, which we suspect can be explained by differences in the number of decimal places accounted for. We do not know if some of the 8% would be counted as over- or underestimations in our study. However, if we allow for slight errors within 1%, we get a success rate of 6%, which is much closer to the results of Loy et al. (2020).

The explanatory variables that we use in the main regression analyses are also included in Tables A.1 and A.2. To simplify the analysis, we mostly use binary indicators derived from the detailed questions included in the survey. Both tables show that the sample is quite representative in terms of demographics, although compared to the general population, more respondents have higher education (71%) and are older than 40 years old (77%), and there are more female than male respondents (62%). From the averages of price consciousness (3.81), sensitivity (4.12), and learning effort (3.52), we understand that most respondents agree or strongly agree on statements related to price sensitivity, whereas respondents are more mixed on activities that generate price knowledge (see Figure A.2).

The variables that are unique to Table A.1 are those related to discount utilization and shopping basket price variation, which are calculated from the full purchasing histories of the respondents, not just the products included in the survey. We see that most consumers have positive discount utilization, meaning that they achieve lower prices than they would by shopping at random. Furthermore, simple discount utilization is higher across the quartiles than sales-weighted discount utilization, which suggests that the products that consumers spend the most on are not the products they buy the most at discounted prices, despite being incentivized to look for sales on products that they can make the most savings on. One explanation is that the products people spend the most on have fewer and smaller discounts for consumers to act on. This is supported by the lower average price coefficient of variation for the respondents' sales-weighted shopping baskets, at 0.1, compared to the simple average across products, at 0.11, indicating that the products consumers spend the most on vary slightly less in price over the sample period.

Table A.2 includes variables that are unique to the recall-level analysis. The share of price recalls for the supermarket and the discount chain is even, and 8% of recalls were on private label products (2 out of the 24 included products). The average weekly number of visits per chain is 1.45 and the average weekly number of unique products purchased at a given chain is 14. The average recalled price is related to a product that has 19 close substitutes, defined by the category it belongs to, and that has an average price coefficient of variation, now measured for a specific product $\times$ chain combination included in the survey, of 0.11, which happens to be identical to the simple average across all products that the respondents purchase. This indicates that the list of products that we include in the survey is representative of all products that consumers buy in terms of price variation.

Out of all the recalled prices that are included in the analysis, 31% were given by respondents who at some point purchased the product at the relevant chain during the sample period. Among those, the average time since purchase is 4.7 months whereas the median is 2.7 months. This might suggest that most respondents buy the same products more frequently, but that a few respondents go much longer between purchases or are even one-time buyers, pulling up the average. The average spending on the product being recalled during the sample period of 18 months is 92 NOK, but this includes the 69%

of recalls for products on which there is zero spending, which is reflected by the median spending of 0 NOK. The total spending distribution among those that bought the product is also skewed, with the average and 3rd quartile spending both being 300 NOK, the 95th percentile being 1200 NOK, and the observed maximum being 18500 NOK.

Table A.3: Factor loadings and index weights for price consciousness, price sensitivity, and price learning effort

Likert Scale Question	**Price Consciousness**		**Price Sensitivity**		**Price Learning Effort**	
	Loading	**Weight**	**Loading**	**Weight**	**Loading**	**Weight**
1. Price is crucial for my choice of products.	0.73	15.4%	0.883	39.7%	–	–
2. Price is crucial for my choice of store.	0.719	15.2%	0.771	34.7%	–	–
3. I buy on Sales and promotions as often as I can.	0.74	15.7%	0.571	25.7%	–	–
4. I usually compare prices between the grocery chains.	0.81	17.1%	–	–	0.86	32.9%
5. I usually check the grocery chains flyers for prices.	0.717	15.2%	–	–	0.762	29.1%
6. I usually check the results of VGs grocery price comparison tests.	0.387	8.2%	–	–	0.411	15.7%
7. I usually check the shelf price of the product before I buy it.	0.624	13.2%	–	–	0.584	22.3%
Sum	4.727	100%	2.225	100%	2.617	100%

Notes: This table shows estimated factor loadings for each Likert scale item using diagonally weighted least squares, as implemented by the *lavaan* package in *R*. The weights put on each likert scale question in the indices for price consciousness, sensitivity, and learning effort are calculated by dividing the factor loadings by the sum of the factor loadings.

Comparison of Mean Price Error and Bias Between Products

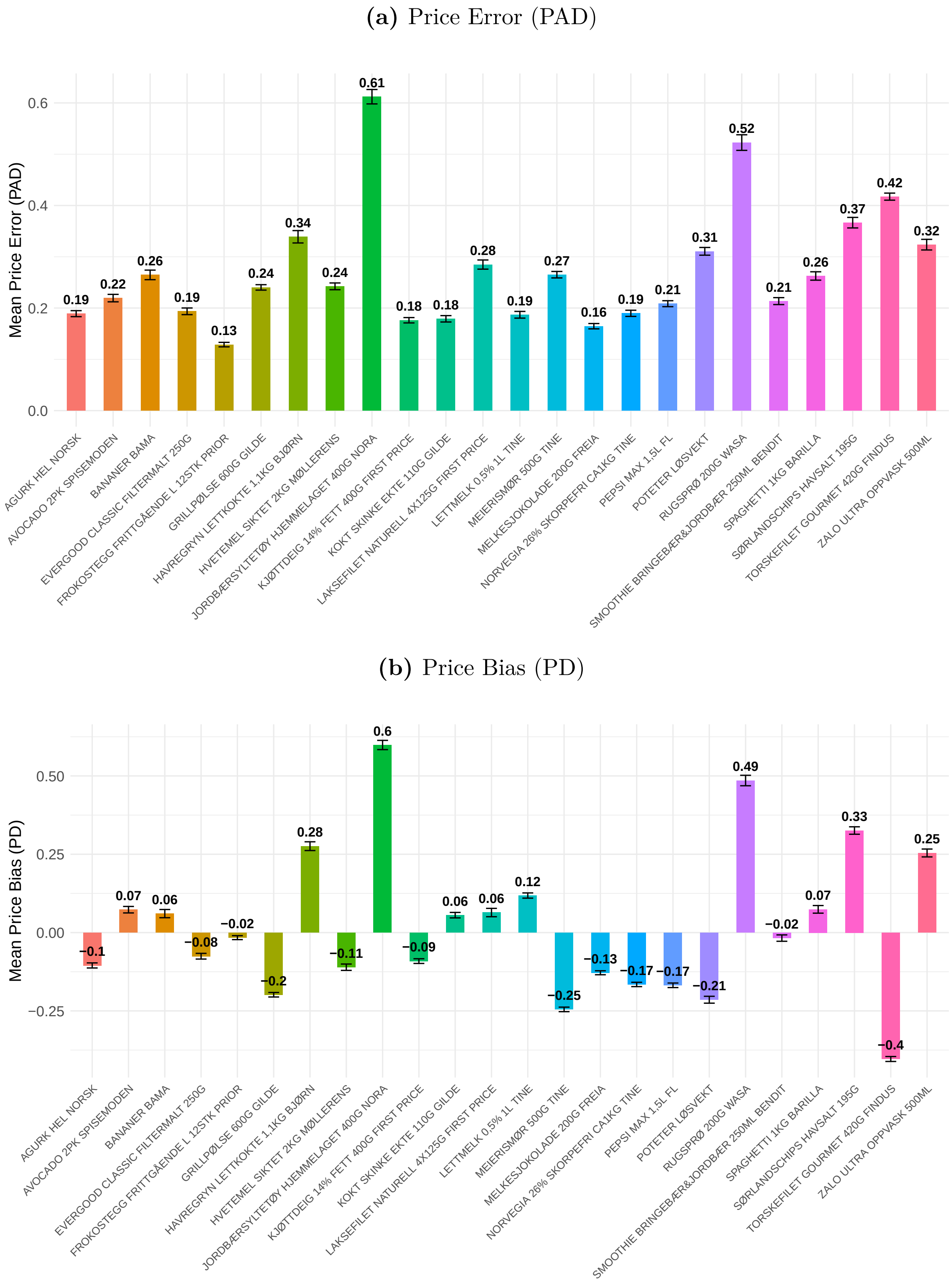


Figure A.1: Price error is measured as the percentage absolute deviation between the respondent's recalled price and the true shelf price. Price bias is measured as the percentage deviation between the respondent's recalled price and the true shelf price. Confidence intervals at 95% are calculated using the Student's t-distribution.

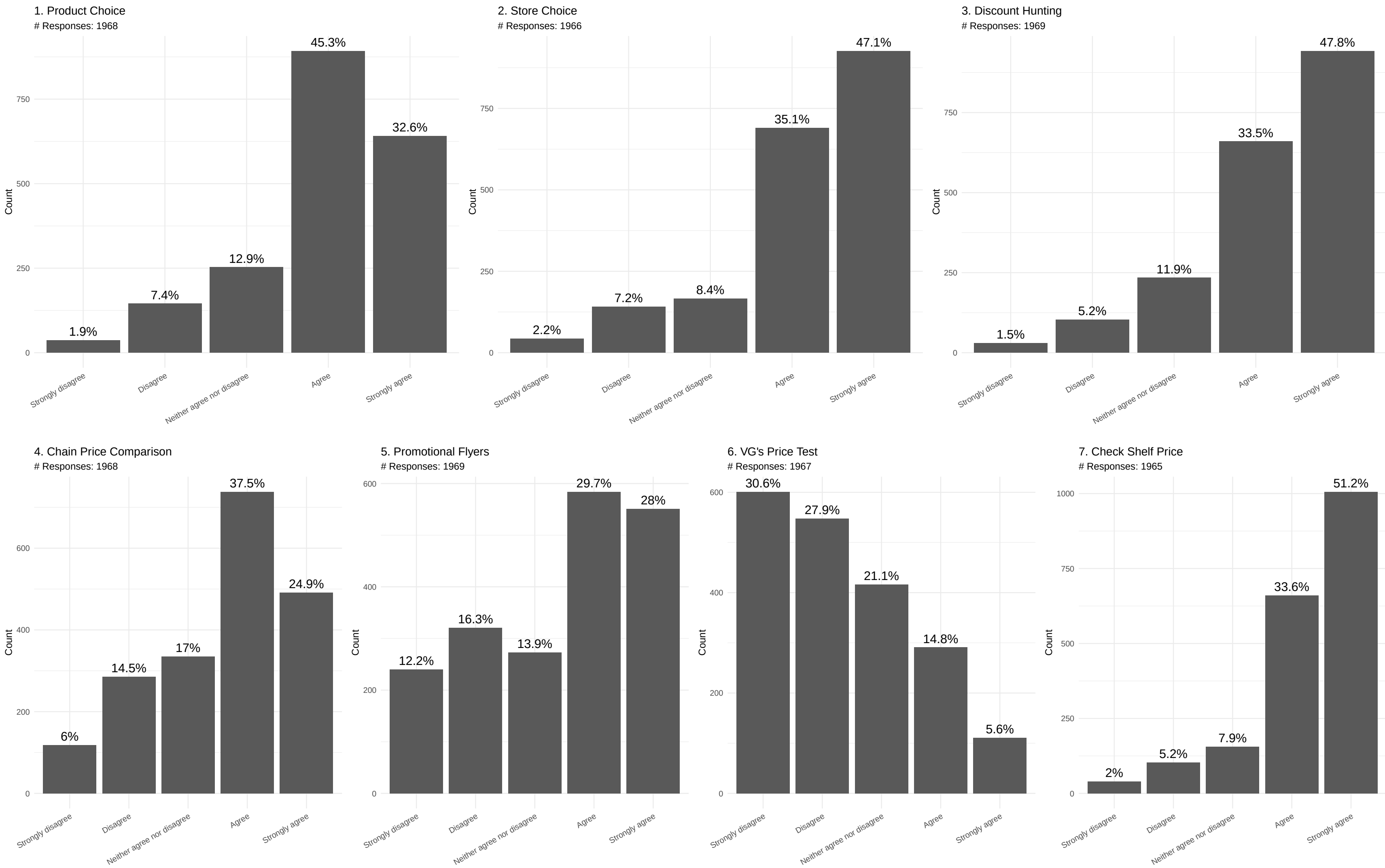


Figure A.2: This figure contains bar charts displaying the distribution of answers for each Likert-scale question. The number of answers is given by the height of the bars, and the percentage share is indicated above the bars.

A.3 Frequency of Purchases and Sales

Where and what consumers buy shapes how often they can purchase on sale, so we briefly summarize the composition of purchases and its relation to sale frequency. Table A.4 shows that respondents concentrate their shopping at the discount chain, which accounts for 83.5% of transactions, 82.5% of units, and 81.6% of spending. The supermarket chain's share is somewhat larger in spending (18.4%) than in transactions (16.5%), reflecting its larger and more expensive baskets. This concentration matters for sale exposure, because the two chains discount very differently: products at the supermarket chain are on temporary sale on about 19% of days, against only about 11% at the discount chain, even though discount-chain sales last longer (Table 2.1).

Table A.4: Purchases by chain: transactions, units, and gross sales

	Transactions		Units		Gross sales (NOK)	
	N	Share	N	Share	NOK	Share
Discount chain	464,950	83.5%	543,689	82.5%	20,708,253	81.6%
Supermarket chain	92,194	16.5%	115,254	17.5%	4,657,808	18.4%
Total	557,144	100.0%	658,943	100.0%	25,366,061	100.0%

Notes: Distribution of tracked respondent purchases across the discount chain and the supermarket chain, before each respondent's survey date. Transactions counts purchase records, Units sums quantity purchased, and Gross sales sums spending. Share columns are the chain's percent of the column total.

Table A.5 describes how often the products respondents buy are on sale, and whether that varies with how much they buy them. Across the 1035 product-by-chain series in the purchase data, the average share of days on sale is 15.8% (Panel A), essentially the same as for the full assortment in Table2.1: viewed product by product, the goods consumers buy are discounted about as often as the grocery products in general. Purchases are, however, highly concentrated—the median series accounts for 49 transactions against a mean of 537—and Panel B shows that the most heavily purchased series are on sale *less* often than others, with a rank correlation between purchases and the share of days on sale of -0.25 across chains (negative within each chain as well). The items that make up the bulk of consumers' baskets are therefore discounted somewhat less frequently than the assortment average, so the sale availability consumers face in practice is somewhat less than 15.8%. That respondents nonetheless buy on sale on about 16% of occasions (Table3.4), despite concentrating their baskets on the discount chain and on less-discounted items, indicates that they time purchases toward sales, consistent with the positive discount utilization documented in Section 3.

Table A.5: Purchase and on-sale frequency for purchased products

Panel A: Distribution across purchased product-by-chain series				
	Mean	Median	Min	Max
Purchases per series	537	49	1	39,692
Units per series	634	55	0	41,546
Share of days on sale	0.158	0.138	0.000	0.514
Panel B: Correlation (purchases × share of days on sale)				
	Pearson	Spearman	# Series	
Both chains	-0.129	-0.253	1,035	
Discount chain	-0.082	-0.112	396	
Supermarket chain	-0.099	-0.152	639	

Notes: Descriptives across the 1035 product-by-chain series purchased by the respondents in our sample. Purchases per series is the number of tracked transactions and Units per series sums quantity. Share of days on sale uses the same definition and unit as the price-dynamics summary table (Table 2.1), the per-series share of days at least 3% below the 90-day rolling-median reference shelf price; Table 2.1 covers all products with shelf-price data, whereas this table restricts to those the sampled respondents purchased. Panel B correlates purchases with the share of days on sale across series: Spearman is the rank correlation, Pearson the linear correlation; the chain rows restrict to the discount chain and the supermarket chain.

B Robustness Checks and Extensions

B.1 Discount Utilization for Perishable vs. Non-Perishable Products

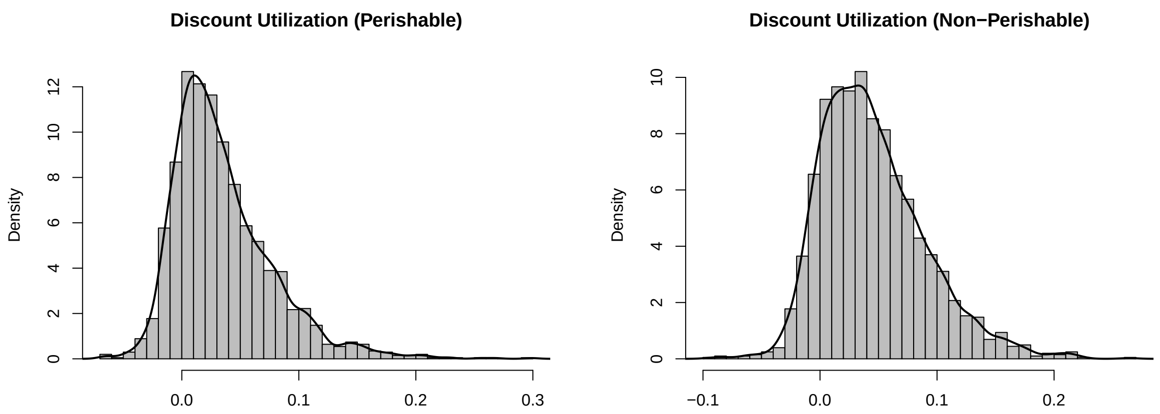


Figure B.1: Histograms and density estimates of discount utilization (DU) for perishable and non-perishable products. DU is computed as a simple average across the perishable and non-perishable products the consumer purchased, respectively. Histogram bin sizes follow Scott's rule and density bandwidths the Sheather–Jones plug-in estimator.

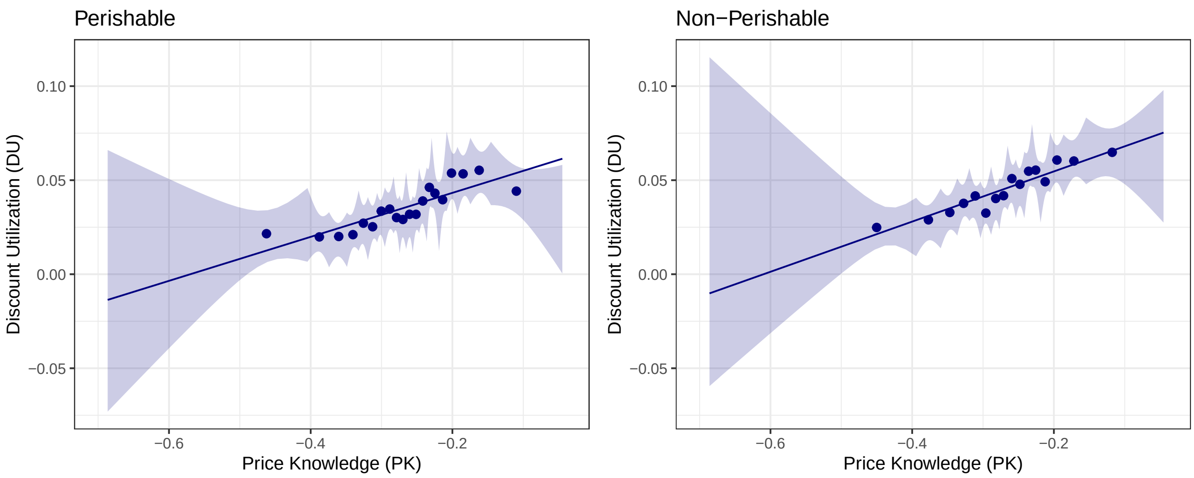


Figure B.2: Estimate of the expected value of discount utilization (DU) conditional on price knowledge (PK), estimated separately for perishable and non-perishable products, from binned scatterplots with a linear fit. PK is the simple average of recall accuracy across the 24 survey products; DU is computed as a simple average across the perishable and non-perishable products the consumer purchased, respectively. The number of bins is chosen to minimize the integrated mean squared error, and 95% confidence bands are calculated using piecewise polynomials.

Table B.1: Discount utilization regressions: perishable vs. non-perishable products

	DU Perishable (1)	DU Non-Perishable (2)	DU Perishable (3)	DU Non-Perishable (4)
Price Knowledge	0.117*** (0.012)	0.133*** (0.013)	0.075*** (0.011)	0.101*** (0.012)
Male			0.000 (0.002)	0.002 (0.002)
Single Person Household			0.002 (0.003)	0.006 (0.003)
Has Children			0.007*** (0.002)	0.003 (0.002)
Full-Time Employment			-0.004 (0.002)	-0.003 (0.002)
Higher Education			0.002 (0.002)	0.002 (0.002)
Rural			-0.004* (0.002)	-0.004* (0.002)
Household Income > 950' NOK			-0.002 (0.002)	0.001 (0.002)
Age: 40-59 Years			0.008*** (0.002)	0.006* (0.002)
Age: Above 60 Years			0.009** (0.003)	0.008** (0.003)
Price Sensitivity			0.003 (0.001)	0.003 (0.001)
Price Learning Effort			0.007*** (0.001)	0.008*** (0.001)
# Frequented Grocery Chains			0.000 (0.001)	0.000 (0.001)
Mainly Uses Discount Stores			-0.004 (0.003)	-0.001 (0.003)
Average Weekly Store Visits			0.004*** (0.001)	0.001 (0.001)
Main Grocery Shopper			0.004 (0.002)	0.000 (0.002)
Average Weekly Variety (per 10 products)			-0.003*** (0.001)	-0.003*** (0.001)
Three Customer Program Memberships			0.006*** (0.002)	0.004* (0.002)
Basket Coefficient of Variation			1.122*** (0.087)	1.327*** (0.098)
Constant	0.067*** (0.003)	0.081*** (0.004)	-0.116*** (0.012)	-0.116*** (0.013)
Observations	2,028	2,028	1,928	1,928
Dependent variable mean	0.04	0.05	0.04	0.05
R^2	0.05	0.06	0.26	0.26

Notes: This table reports results from estimation of separate discount utilization regressions for perishable and non-perishable products. Discount utilization is calculated as the percentage difference between the quarterly average of daily modal (shelf) prices and the average price paid on the purchase day, averaged across purchase days and then across product × chain combinations (simple weights). Price knowledge is calculated as the negative average percentage absolute deviation between the respondent's recalled prices and true shelf prices. The average weekly variety is calculated as the average number of unique products purchased per week over the sample period. Price sensitivity and learning effort are calculated as the factor-weighted averages of Likert-scale survey items 1–3 and 4–7, respectively. Basket CV is calculated as the coefficient of variation of shelf prices over the sample period, averaged across all products in the consumer's purchase history, using simple weights. Columns 1–2 include the simple regression on price knowledge for perishable and non-perishable products, and Columns 3–4 adds controls for demographics and shopping habits. The standard errors shown in parentheses are robust to heteroskedasticity (HC1). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

B.2 Discount Utilization for Non-Survey Products

To address a potential concern with reverse causality, we re-estimate the relationship between price knowledge and discount utilization, but we calculate the dependent variable using only the products in the respondents' purchase history that were not included in the survey. Thus, we separate the measurement of price knowledge (based on the 18 surveyed items) from the behavior it predicts (discount utilization across the remaining ~700 items in the purchase history).

The results are presented in Table B.2. We find that the association between price knowledge and discount utilization remains robust and highly statistically significant when excluding the survey products. In fact, the estimated coefficient on price knowledge is slightly larger in this specification than in the baseline results.

Table B.2: Discount utilization regressions: non-survey products

	Discount Utilization		
	(1)	(2)	(3)
Price Knowledge	0.137^{***} (0.013)	0.133^{***} (0.012)	0.097^{***} (0.011)
Male		0.002 (0.002)	0.003 (0.002)
Single Person Household		0.004 (0.003)	0.004 (0.003)
Has Children		0.005^{*} (0.002)	0.006^{**} (0.002)
Full-Time Employment		-0.005^{*} (0.002)	-0.003 (0.002)
Higher Education		0.001 (0.002)	0.003 (0.002)
Rural		-0.007^{***} (0.002)	-0.005^{**} (0.002)
Household Income > 950' NOK		-0.002 (0.002)	-0.001 (0.002)
Age: 40-59 Years		0.007^{**} (0.002)	0.007^{***} (0.002)
Age: Above 60 Years		0.006^{*} (0.003)	0.007^{*} (0.003)
Price Sensitivity		0.008^{***} (0.001)	0.004^{**} (0.001)
Price Learning Effort			0.008^{***} (0.001)
# Frequented Grocery Chains			0.000 (0.001)
Mainly uses discount stores			-0.006 (0.003)
Average Weekly Store Visits			0.003^{***} (0.001)
Main Grocery Shopper			0.001 (0.002)
Average Weekly Variety (per 10 products)			-0.004^{***} (0.001)
Three Customer Program Memberships			0.006^{***} (0.002)
Basket Coefficient of Variation			1.360^{***} (0.087)
Constant	0.080^{***} (0.004)	0.042^{***} (0.007)	-0.129^{***} (0.012)
Observations	2,028	1,935	1,928
Dependent variable mean	0.04	0.04	0.04
R^2	0.07	0.11	0.33

Notes: This table reports results from estimation of Equation 2.1. The dependent variable, individual-level discount utilization, is the percentage difference between the quarterly average of daily modal (shelf) prices and the average price paid on the purchase day, averaged across purchase days and then across non-survey product × chain combinations (simple weights). Price knowledge is calculated as the negative average percentage absolute deviation between the respondent's recalled prices and true shelf prices. The average weekly variety is calculated as the average number of unique products purchased per week over the sample period. Price sensitivity and learning effort are calculated as the factor-weighted averages of Likert-scale survey items 1–3 and 4–7, respectively. Basket CV is calculated as the coefficient of variation of shelf prices over the sample period, averaged across all products in the consumer's purchase history, using simple weights. Column 1 reports the simple regression on price knowledge (PK), Column 2 adds controls for demographics and price sensitivity, and Column 3 adds controls for shopping habits and price learning effort. The standard errors shown in parentheses are robust to heteroskedasticity (HC1). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

B.3 Decomposition of Discount Utilization and Alternative Reference-Price Windows

Table B.3: Discount utilization: decomposition and reference-price window

	Decomposition of baseline DU			Alternative reference-price windows				
	Quarterly (1)	Quarterly (Timing Only) (2)	Daily (3)	1 Week (4)	2 Weeks (5)	Monthly (6)	6 Months (7)	Full Period (8)
Price Knowledge (PK)	0.125***	0.109***	0.020***	0.027***	0.053***	0.090***	0.136***	0.145***
	(0.011)	(0.009)	(0.004)	(0.005)	(0.006)	(0.009)	(0.012)	(0.013)
Constant	0.074***	0.066***	0.008***	0.014***	0.030***	0.051***	0.080***	0.083***
	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.004)
Observations	2,028	2,028	2,028	2,027	2,027	2,028	2,027	2,028
Dependent variable mean	0.04	0.04	0.00	0.01	0.02	0.03	0.04	0.04
R^2	0.07	0.08	0.01	0.02	0.04	0.06	0.08	0.08

Notes: Robustness analysis for the main discount-utilization result, in two column panels. The decomposition panel breaks the paper's baseline DU into two channels. Column 1 is baseline DU — the purchase price relative to the quarterly-average shelf (regular) price. Column 2 (timing only) re-values each purchase at the posted shelf price instead of the paid price, capturing the saving from buying when the shelf price is temporarily low (sale timing); column 3 (daily) measures the purchase price against the same day's shelf price, capturing coupons and personal discounts net of timing. Baseline DU is approximately the sum of columns 2 and 3. The window panel re-averages the reference (regular) price over alternative calendar windows, from one week to the full 18-month panel. Heteroskedasticity-robust (HC1) standard errors in parentheses. $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$.

B.4 Discount Utilization and Price Knowledge Quintiles

To assess the functional form of the relationship between price knowledge and discount utilization, and to evaluate the empirical validity of the binary distinction between "informed" and "uninformed" consumers often found in search-theoretic models (e.g., Varian (1980)), we re-estimate the model using quintile indicators of price knowledge. In this specification, the bottom quintile (lowest price knowledge) serves as the reference category. The results are presented in Table B.4, and the corresponding quintile intervals for price knowledge are reported in Table B.5.

Table B.4: Discount utilization regressions: price knowledge quintiles

	Discount Utilization		
	(1)	(2)	(3)
Price Knowledge Q2 (20–40%)	0.010*** (0.002)	0.010*** (0.002)	0.010*** (0.002)
Price Knowledge Q3 (40–60%)	0.012*** (0.002)	0.011*** (0.002)	0.010*** (0.002)
Price Knowledge Q4 (60–80%)	0.023*** (0.003)	0.023*** (0.003)	0.018*** (0.002)
Price Knowledge Q5 (80–100%)	0.032*** (0.003)	0.030*** (0.003)	0.022*** (0.002)
Male		0.000 (0.002)	0.001 (0.002)
Single Person Household		0.004 (0.003)	0.004 (0.003)
Has Children		0.004* (0.002)	0.005** (0.002)
Full-Time Employment		-0.005* (0.002)	-0.003 (0.002)
Higher Education		0.001 (0.002)	0.002 (0.002)
Rural		-0.006*** (0.002)	-0.005** (0.002)
Household Income > 950' NOK		-0.003 (0.002)	-0.001 (0.002)
Age: 40-59 Years		0.007*** (0.002)	0.007*** (0.002)
Age: Above 60 Years		0.007** (0.003)	0.008*** (0.002)
Price Sensitivity		0.007*** (0.001)	0.003** (0.001)
Price Learning Effort			0.007*** (0.001)
# Frequented Grocery Chains			0.000 (0.001)
Mainly uses discount stores			-0.003 (0.002)
Average Weekly Store Visits			0.003*** (0.001)
Main Grocery Shopper			0.002 (0.002)
Average Weekly Variety (per 10 products)			-0.003*** (0.001)
Three Customer Program Memberships			0.005*** (0.002)
Basket Coefficient of Variation			1.205*** (0.074)
Constant	0.025*** (0.002)	-0.006 (0.005)	-0.149*** (0.010)
Observations	2,028	1,935	1,928
Dependent variable mean	0.04	0.04	0.04
R^2	0.09	0.13	0.36

Notes: This table reports results from estimation of discount utilization regressions using indicators for the quintiles of price knowledge. The dependent variable, individual-level discount utilization, is the percentage difference between the quarterly average of daily modal (shelf) prices and the average price paid on the purchase day, averaged across purchase days and then across product × chain combinations (simple weights). Price knowledge is calculated as the negative average percentage absolute deviation between the respondent's recalled prices and true shelf prices. The average weekly variety is calculated as the average number of unique products purchased per week over the sample period. Price sensitivity and learning effort are calculated as the factor-weighted averages of Likert-scale survey items 1–3 and 4–7, respectively. Basket CV is calculated as the coefficient of variation of shelf prices over the sample period, averaged across all products in the consumer's purchase history, using simple weights. Column 1 reports the simple regression on price knowledge, Column 2 adds controls for demographics and price sensitivity, and Column 3 adds controls for shopping habits and price learning effort. The standard errors shown in parentheses are robust to heteroskedasticity (HC1). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table B.5: Quintiles for Price Knowledge

Quintile	PK interval
Q1 (0–20%)	$[-0.70, -0.33]$
Q2 (20–40%)	$(-0.33, -0.28]$
Q3 (40–60%)	$(-0.28, -0.25]$
Q4 (60–80%)	$(-0.25, -0.21]$
Q5 (80–100%)	$(-0.21, -0.04]$

The results confirm a monotonic positive relationship between price knowledge and discount utilization. Relative to the reference category, discount utilization increases as we move to higher quintiles. Entering the second quintile is associated with a 1.0 percentage point increase in discount utilization. While the increase from the second to the third quintile is small (0.2 percentage points), the subsequent increases to the fourth and fifth quintiles are substantial (1.1 and 0.9 percentage points, respectively).

The quintiles around the center of mass of the distribution (PK $= -0.27$) are quite narrow, particularly the third quintile (Q3), which spans only 3 percentage points (from -0.28 to -0.25). The flattening of the effect in this region may therefore reflect that the estimated Q3 coefficient is sensitive to local idiosyncrasies rather than indicating a genuine nonlinearity in the underlying relationship between price knowledge and discount utilization. This interpretation is supported by the non-parametric evidence in the binned scatterplot (Figure 3.2), which shows a fairly linear trend through the bulk of the data despite some local variability. Consequently, we conclude that the linear assumption employed in our main specification remains a reasonable and robust approximation.

We further note that the coefficients on the demographic and behavioral control variables remain stable in magnitude and statistical significance relative to the baseline linear specification (Table 3.1).

Finally, we find no evidence of distinct thresholds of price knowledge associated with sharp changes in discount utilization. Instead, discount utilization increases gradually across the distribution. This pattern suggests that the binary classification used in theoretical models is a simplification of real-world behavior. Price knowledge is better viewed as a continuous trait, where incremental improvements lead to incremental reductions in prices paid.

B.5 Sales-Weighted Discount Utilization Regression

Table B.6: Discount utilization regressions: sales weights

	Discount Utilization		
	(1)	(2)	(3)
Price Knowledge	0.070*** (0.010)	0.073*** (0.007)	0.054*** (0.007)
Male		0.000 (0.001)	0.000 (0.001)
Single Person Household		0.000 (0.002)	0.001 (0.002)
Has Children		-0.001 (0.001)	0.002 (0.001)
Full-Time Employment		-0.004* (0.002)	-0.002 (0.001)
Higher Education		-0.001 (0.001)	0.000 (0.001)
Rural		-0.006*** (0.001)	-0.003* (0.001)
Household Income > 950' NOK		0.000 (0.001)	0.001 (0.001)
Age: 40-59 Years		0.003* (0.001)	0.004** (0.001)
Age: Above 60 Years		0.006** (0.002)	0.005** (0.002)
Price Sensitivity		0.003** (0.001)	0.001 (0.001)
Price Learning Effort			0.005*** (0.001)
# Frequented Grocery Chains			0.000 (0.000)
Mainly uses discount stores			-0.008*** (0.002)
Average Weekly Store Visits			0.003*** (0.001)
Main Grocery Shopper			0.002 (0.001)
Average Weekly Variety (per 10 products)			-0.003*** (0.001)
Three Customer Program Memberships			0.003** (0.001)
Basket Coefficient of Variation			0.566*** (0.051)
Constant	0.040*** (0.003)	0.032*** (0.005)	-0.042*** (0.007)
Observations	2,028	1,935	1,928
Dependent variable mean	0.02	0.02	0.02
R^2	0.04	0.08	0.31

Notes: This table reports results from estimation of Equation 2.1. The dependent variable, individual-level discount utilization, is the percentage difference between the quarterly average of daily modal (shelf) prices and the average price paid on the purchase day, averaged across purchase days and then across product × chain combinations, with the latter using sales weights. Price knowledge is calculated as the negative average percentage absolute deviation between the respondent's recalled prices and true shelf prices. The average weekly variety is calculated as the average number of unique products purchased per week over the sample period. Price sensitivity and learning effort are calculated as the factor-weighted averages of Likert-scale survey items 1–3 and 4–7, respectively. Basket CV is calculated as the coefficient of variation of shelf prices over the sample period, averaged across all products in the consumer's purchase history, using sales weights. Column 1 reports the simple regression on price knowledge, Column 2 adds controls for demographics and price sensitivity, and Column 3 adds controls for shopping habits and price learning effort. The standard errors shown in parentheses are robust to heteroskedasticity (HC1). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table B.7: Discount utilization regressions: sales-weighted basket CV interactions

	Discount Utilization	
	(1)	(2)
Price Knowledge	0.070***	0.046***
	(0.010)	(0.005)
High Basket CV		0.029***
		(0.005)
Price Knowledge × High Basket CV		0.039*
		(0.018)
Constant	0.040***	0.024***
	(0.003)	(0.002)
Observations	2,028	2,028
Dependent variable mean	0.02	0.02
R^2	0.04	0.16

Notes: This table reports results from estimation of discount utilization regressions using an interaction between price knowledge and an indicator for having above-median basket price variation. Discount utilization is calculated as the percentage difference between the quarterly average of daily modal (shelf) prices and the average price paid on the purchase day, averaged across purchase days and then across product × chain combinations (sales weights). Price knowledge (PK) is calculated as the negative average percentage absolute deviation between the respondent's recalled prices and true shelf prices. Basket CV is calculated as the coefficient of variation of shelf prices over the sample period, averaged across all products in the consumer's purchase history, using sales weights. Column 1 includes the simple regression on price knowledge and Column 2 adds an interaction with Basket CV. The standard errors shown in parentheses are robust to heteroskedasticity (HC1). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

B.6 Discount Utilization and the Share of Correct Price Recalls

Table B.8 shows the OLS results of a simple regression of discount utilization on the share of recalled prices within some error bound, using the sample including outliers in PK, for comparison with the results in Table 3.1:

$$\mathrm{DU}(Q)_i^{\mathrm{Simple}} = \alpha + \delta \cdot W_i^e + \epsilon_i \tag{B.1}$$

where W_i^e is the share of correctly recalled prices within some error bound e. Like a high PK, a high share of correctly recalled prices is predictive of more discount utilization. This is reassuring, since the share of correct recalls is robust to extreme price recalls, for example, from erroneous entry or inattention, that could significantly affect the PK.

The explained variance is larger for the share of roughly correct recalls, i.e., within 20%. This suggests that consumers do not need precise knowledge of prices in order to act on sales—a rough idea of price levels suffices. Still, the PK performs even better with an R^2 of 7% compared to the R^2 of 5% obtained by using the share of roughly correct recalls.[35] Finally, note that the larger coefficients on precise relative to rough recalls do not signify that precise recalls are more important, since a given increase in the share of precise recalls is much harder to obtain than an equal increase in the share of roughly correct recalls.

[35]This also holds when comparing like-for-like on the sample without outliers.

Table B.8: Discount utilization regressions: within percentage error bounds

	Discount Utilization			
	(1)	(2)	(3)	(4)
Share Within 1%	0.064***			
	(0.012)			
Share Within 5%		0.060***		
		(0.007)		
Share Within 10%			0.051***	
			(0.006)	
Share Within 20%				0.051***
				(0.005)
Constant	0.035***	0.030***	0.025***	0.014***
	(0.001)	(0.001)	(0.002)	(0.002)
Observations	2,285	2,285	2,285	2,285
Dependent variable mean	0.04	0.04	0.04	0.04
R^2	0.02	0.03	0.04	0.05

Notes: This table reports results from estimation of discount utilization regressions using the share of correct price recalls within different error bounds and the sample including outliers in price knowledge. Discount utilization is calculated as the percentage difference between the quarterly average of daily modal (shelf) prices and the average price paid on the purchase day, averaged across purchase days and then across product × chain combinations using simple weights. Columns (1-4) uses error bounds of 1, 5, 10, and 20 percent as explanatory variables, respectively. The standard errors shown in parentheses are robust to heteroskedasticity (HC1). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

B.7 Explaining the Likelihood of Correct Price Recalls

Table B.9 reports average marginal effects from logistic regressions, estimated by maximum likelihood, in which the dependent variable is an indicator for whether a price recall falls within varying error bounds. The explanatory variables are the same as those in Column 2 of Table 4.1. Estimates are based on the full sample of price recalls, including outliers, using the following specification:

$$Pr(W^e_{ijc} = 1|\{\theta, \mathbf{X}_{ijc}\}) = \Lambda(\theta + \gamma^{'}\mathbf{X}_{ijc}) = \frac{1}{1 + e^{-(\theta+\gamma^{'}\mathbf{X}_{ijc})}} \tag{B.2}$$

The results are qualitatively similar to those reported in our main specification (Table 4.1) in terms of both sign and statistical significance, particularly when using the more lenient error bounds. This consistency is reassuring, as the binary indicators for correct recall are robust to outliers, confirming that our primary results are not driven by our outlier exclusion criteria.

The results in Table B.9 also illustrate how the determinants of price knowledge differ between precise and approximate recall. In most cases, the average marginal effects on the probability of a correct recall share the same sign across specifications; however, the magnitudes are naturally larger for the more lenient error bounds, reflecting the higher baseline probability of achieving a rough recall.[36]

Notably, we observe a divergence regarding the time elapsed since the last purchase. While the probability of a precise recall (within 1%) diminishes significantly as time passes, we find no statistically significant decay for roughly correct recalls (within 20%). Among consumers who have purchased a product, the probability of a precise recall drops by 2.6 percentage points (from 9.1% to 6.5%) after 12 months. In contrast, the probability of a roughly correct recall decreases by only 1 percentage point (from 59% to 58%) over the same period. This suggests that while exact price information held in short-term memory fades, a general sense of the price level persists in long-term memory. This finding aligns with Vanhuele and Drèze (2002), who show that consumers can recognize deal prices even when unable to recall exact prices.

[36] This is consistent with the sigmoid shape of the logistic function, given that precise recalls are significantly rarer than rough recalls.

Table B.9: Logistic price knowledge regressions: correct price recalls

	Pr(Within 1%) (1)	Pr(Within 5%) (2)	Pr(Within 10%) (3)	Pr(Within 20%) (4)
Male	-0.118* (0.055) [-0.7%]	-0.134*** (0.036) [-1.8%]	-0.131*** (0.030) [-2.6%]	-0.133*** (0.028) [-3.2%]
Single Person Household	0.094 (0.104) [0.5%]	-0.002 (0.064) [0.0%]	-0.032 (0.053) [-0.6%]	-0.037 (0.047) [-0.9%]
Has Children	0.000 (0.062) [0.0%]	0.033 (0.042) [0.4%]	0.017 (0.035) [0.3%]	0.005 (0.033) [0.1%]
Full-Time Employment	-0.040 (0.072) [-0.2%]	-0.007 (0.045) [-0.1%]	-0.039 (0.038) [-0.8%]	-0.028 (0.034) [-0.7%]
Higher Education	-0.183** (0.061) [-1.0%]	-0.172*** (0.039) [-2.4%]	-0.135*** (0.033) [-2.7%]	-0.092** (0.031) [-2.2%]
Rural	0.003 (0.059) [0.0%]	-0.025 (0.038) [-0.3%]	-0.013 (0.032) [-0.3%]	-0.010 (0.030) [-0.2%]
Household Income > 950' NOK	0.060 (0.064) [0.3%]	0.083* (0.041) [1.1%]	0.068* (0.034) [1.4%]	0.076* (0.031) [1.8%]
Age: 40-59 Years	0.061 (0.063) [0.3%]	0.030 (0.040) [0.4%]	-0.002 (0.034) [0.0%]	-0.023 (0.032) [-0.6%]
Age: Above 60 Years	-0.083 (0.100) [-0.5%]	-0.039 (0.062) [-0.5%]	-0.104* (0.052) [-2.0%]	-0.108* (0.047) [-2.6%]
# Frequented Grocery Chains	0.015 (0.019) [0.1%]	0.014 (0.012) [0.2%]	0.014 (0.010) [0.3%]	0.009 (0.010) [0.2%]
Mainly Uses Discount Stores	-0.102 (0.085) [-0.6%]	-0.005 (0.054) [-0.1%]	0.009 (0.046) [0.2%]	0.014 (0.042) [0.3%]
Main Grocery Shopper	0.067 (0.053) [0.4%]	0.069 (0.037) [0.9%]	0.074* (0.032) [1.5%]	0.098** (0.030) [2.4%]
Supermarket Chain	-0.132* (0.059) [-0.7%]	0.035 (0.028) [0.5%]	0.118*** (0.019) [2.3%]	0.200*** (0.016) [4.8%]
Private Label	-1.182*** (0.107) [-4.3%]	0.200*** (0.039) [2.8%]	0.128*** (0.033) [2.6%]	0.118*** (0.034) [2.8%]
Three Customer Program Memberships	0.110* (0.052) [0.6%]	0.055 (0.034) [0.7%]	0.084** (0.029) [1.7%]	0.084** (0.027) [2.0%]
Price Sensitivity	-0.061 (0.045) [-0.3%]	-0.024 (0.029) [-0.3%]	-0.019 (0.023) [-0.4%]	-0.013 (0.020) [-0.3%]
Price Learning Effort	0.126*** (0.037) [0.7%]	0.122*** (0.023) [1.6%]	0.130*** (0.020) [2.6%]	0.135*** (0.018) [3.2%]
Average Weekly Chain Visits	0.075* (0.031) [0.4%]	0.055** (0.017) [0.7%]	0.051*** (0.013) [1.0%]	0.054*** (0.012) [1.3%]
Average Weekly Chain Variety (per 10 products)	-0.075* (0.031) [-0.4%]	-0.069*** (0.018) [-0.9%]	-0.077*** (0.014) [-1.5%]	-0.091*** (0.012) [-2.2%]
Category Size	-0.062*** (0.011) [-0.3%]	-0.064*** (0.007) [-0.9%]	-0.049*** (0.005) [-1.0%]	-0.093*** (0.005) [-2.2%]
Price Coefficient of Variation	-1.381*** (0.305) [-7.7%]	-1.245*** (0.197) [-16.6%]	-0.766*** (0.172) [-15.3%]	-1.251*** (0.163) [-30.0%]
Purchased	0.417*** (0.049) [2.4%]	0.465*** (0.034) [6.6%]	0.418*** (0.030) [9.2%]	0.455*** (0.030) [12.2%]
Purchased × Sales in 1000 NOK	0.303*** (0.038) [2.2%]	0.278*** (0.034) [4.5%]	0.449*** (0.037) [10.0%]	0.588*** (0.049) [13.8%]
Purchased × Months Since Last Purchase	-0.030*** (0.006) [-0.2%]	-0.020*** (0.004) [-0.3%]	-0.009** (0.003) [-0.2%]	-0.004 (0.003) [-0.1%]
Constant	-2.630*** (0.198)	-1.888*** (0.130)	-1.304*** (0.112)	-0.377*** (0.100)
Observations	73,548	73,548	73,548	73,548
Baseline Probability	6%	16%	29%	49%
Pseudo R^2	0.03	0.02	0.02	0.03

Notes: This table reports results from estimation of Equation B.2 using maximum likelihood. Price knowledge (within) is an indicator for a recalled price being within a percentage error bound of the true shelf price. Price sensitivity and learning effort are calculated as the factor-weighted average of Likert items 1–3 and 4–7, respectively. The average weekly chain variety is calculated as the average number of unique products purchased at each chain per week over the sample period. Category size is the chain-specific number of close substitutes for the product as defined by the grocery chain. The price coefficient of variation is calculated for each product × chain combination over the sample period. Purchased is an indicator of whether the respondent bought the product at a given chain during the sample period, and sales is the consumer's total product–chain spending. Columns 1–4 parallel Column 3 in Table 4.1 using error bounds of 1, 5, 10, and 20 percent, respectively. The standard errors shown in parentheses are clustered by individual, and the average marginal effects are shown in square brackets. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

C Full Consumer Survey

The survey used in this article was originally presented to respondents in Norwegian. Below, we present an English version of the complete survey, translated using Google Translate. We have not fine-tuned the translated version to match natural English usage; consequently, some formulations may appear awkward or unpolished.

The survey was distributed to Trumf members through a newsletter over the period 4–12 June 2024. The median respondent spent 8.7 minutes completing the questionnaire. Out of the 2505 respondents, 418 did not complete the survey in its entirety, yielding an attrition rate of 16.7%. Most respondents who dropped out did so before the price recall section.

There were a few inconsistencies in the Likert-scale response options. Two different Norwegian synonyms for "Strongly" ("Svært" and "Veldig") were used for the extreme response categories, and the adverb "Slightly" appeared only in the "Agree" option. It is difficult to determine whether this influenced respondents' choices, but we suspect it may have slightly shifted some answers between the two outer categories on either side.

English

Introduksjon

Price knowledge and purchasing behavior

You are hereby invited to participate in a research project led by FOOD, a collaboration between the Norwegian School of Economics (NHH) and NorgesGruppen. The study aims to explore the relationship between grocery customers' price knowledge and their purchasing behavior.

Project overview

Participation involves answering questions about demographics, shopping habits, and price awareness. No directly identifiable information about you will be shared with the FOOD project.

A selection of your purchase history from Trumf, which will be linked to the answers you provide in the survey, will be shared with the FOOD project. This includes information about selected products from product groups in the survey, sales amount, quantity/weight, chain and date of purchase, and number of purchases and average shopping cart size per week from January 1, 2023 to June 12, 2024.

Trumf will process the sales data and link it to your survey responses, but they will not have access to your responses.

Participation details

Participation in the survey is voluntary. The information is used only in the FOOD project and is treated confidentially in accordance with data protection regulations. The project will end in 2024, after which only anonymized data will be stored on a secure server at NHH, accessible to FOOD researchers.

Once you have completed the survey, you will be entered into the draw to win:

- Main prize: Trumf bonus of NOK 5,000
- Partial prizes for 10 pieces: Trumf bonus of 500 kr

The survey takes approximately **5-10 minutes** to complete.

By answering the survey, you agree to participate in the research project. If you do not wish to participate, please close the tab.

We greatly appreciate your participation!

Demografi

What is your age?

Under 30 years old

30-39 years old

40-49 years

50-59 years

60-69 years

70 years or over

What gender are you?

Woman

Man

Other:

How many members are there in your household?

1

2

3

4

5

6 or more

How many children under the age of 18 are there in your household?

0

1

2

3

4 or more

Choose the option that *best* describes your work situation:

Full-time employee

Part-time employee

Self-employed

Not at work

Student

Retired

Other:

What is your highest completed education?

Primary school

High school

Vocational school

University/college

Other:

What is your household's total annual income before taxes?

Total income includes all wage income, business income, capital income, pensions, and other taxable and tax-exempt transfers.

Under 300,000 NOK

300,000 - 600,000 kr

600,000 - 950,000 kr

950,000 - 1,350,000 kr

1,350,000 - 1,800,000 kr

Over 1,800,000 kr

How would you describe the area you live in?

City

Suburb

Rural

Handlevaner

Which of these grocery chains do you usually shop at during a typical month?

Check multiple boxes if applicable.

Kiwi

Rema 1000

Coop Extra

Menu

Spar/Eurospar

Coop Note

Coop Mega

Coop Price

Joker

Bottom price

Coop Market

The convenience store

The food corner

Oda

You did not check these chains. Which of them would it have been practical for you to shop at in your everyday life?

Check multiple boxes if applicable.

» Kiwi

» Rema 1000

» Coop Extra

» Menu

» Spar/Eurospar

» Coop Note

» Coop Mega

- Coop Price
- Joker
- Bottom price
- Coop Market
- The convenience store
- The food corner
- Oda

Which of these chains do you shop at the most?

- Kiwi
- Rema 1000
- Coop Extra
- Menu
- Spar/Eurospar
- Coop Note
- Coop Mega
- Coop Price
- Joker
- Bottom price
- Coop Market
- The convenience store
- The food corner
- Oda

How many times do you shop for groceries (on average) per week?

0-1

2-3

4-5

6 or more

How many loyalty programs are you a member of (Æ, Trumf, Coop)?

1

2

3 or more

Are you primarily responsible for purchasing groceries for the household?

Yes

No

Shared responsibility

Priskunnskap - Intro

You will now be asked what you think the current shelf price is for 18 different grocery products.

Any offers or promotions should be taken into account, but you can disregard personal offers and coupons.

We ask that you answer as best you can based on your own memory.

Frokostprodukter

What is the current shelf price of **200g Wasa Rye Crackers** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **400g Nora homemade strawberry jam** in the following stores?

Number of kroner:

Kiwi

Menu

What is the current shelf price of **110g sliced cooked ham from Gilde** in the following stores?

Number of kroner:

Kiwi

Menu

What is the current shelf price of **1.1kg Bjørn easy-cooked oatmeal from AXA** in the following stores?

Number of kroner:

Kiwi

Menu

Kjøtt og fisk

What is the current shelf price of **400g of minced meat from First Price** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **600g Gilde grill sausages** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **4x125g salmon fillets from First Price** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **420g cod fillet from Findus** in the following stores?

	Number of kroner:
Kiwi	
Menu	

Meieriprodukter

What is the current shelf price of **1l Tine low-fat milk 0.5% fat** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **12 breakfast eggs from Prior** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **500g Tine dairy butter** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **1kg of Norvegia yellow cheese** in the following stores?

	Number of kroner:
Kiwi	
Menu	

Husholdningsprodukter og tørrvarer

What is the current shelf price of **2kg of Møllerens sifted wheat flour** in the following stores?

	Number of kroner:
Kiwi	
Menu	

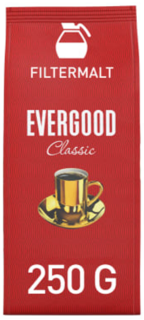
What is the current shelf price of **500ml of Zalo Ultra dishwashing detergent** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **1kg of spaghetti from Barilla** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **250g Evergood Classic filter ground coffee** in the following stores?



	Number of kroner:
Kiwi	
Menu	

Brus, snacks og søtsaker

What is the current shelf price of **200g Freia milk chocolate** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **1.5l Pepsi Max** in the following stores (without deposit)?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **195g Sørlandschips sea salt** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **250ml Bendit smoothie with raspberries and strawberries** in the following stores?

	Number of kroner:
Kiwi	
Menu	

Frukt og grønt

What is the current shelf price of **a Norwegian cucumber from Gartner** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price for **2 packs of ripe avocados from Bama** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **1kg of bananas from Bama** in the following stores?

	Number of kroner:
Kiwi	
Menu	

What is the current shelf price of **1kg of potatoes in bulk** in the following stores?

	Number of kroner:
Kiwi	
Menu	

Prisbevissthet - 1-by-1

Finally, we ask you to indicate the extent to which you agree with 10 different statements.

Price is crucial for my choice of *products* .

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

Price is crucial for my choice of *store* .

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

I buy on offers and promotions as often as I can.

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

I usually compare prices between chains.

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

I usually check the grocery chains' customer flyers for prices.

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

I usually check the results in VG's price comparison test.

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

I usually check the shelf price of the product before I buy it.

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

I have good knowledge of grocery prices.

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

I have better knowledge about the prices of groceries I buy frequently.

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree

I have better knowledge about the prices of groceries that cost a lot.

Strongly disagree	Disagree	Neither agree nor disagree	Slightly agree	Strongly agree