Do Pharmacists Buy Bayer?
Sophisticated Shoppers and the Brand Premium

Bart J. Bronnenberg,* *CentER and Tilburg*

Jean-Pierre Dubé, *Chicago Booth and NBER*

Matthew Gentzkow, *Chicago Booth and NBER*

Jesse M. Shapiro, *Chicago Booth and NBER*

June 2013

Abstract

We estimate the effect of information on consumers' willingness to pay for branded goods in physically homogeneous consumer packaged goods categories. In a case study of headache remedies, we find that college education, working in a healthcare occupation, and other proxies for product knowledge predict more purchases of private labels relative to brands. Pharmacists devote almost 90 percent of headache remedy purchases to private labels, against 71 percent for the average consumer. The effect of knowledge is similar across a broad set of health products, and in a set of relatively homogeneous food products, but smaller for food and drink products overall. We conclude that a significant share of the willingness to pay for brands in these categories would disappear in a world where consumers were fully informed.

keywords: branding, private label, painkillers

JEL codes: D12, D83, L66

---

*We are grateful to Kevin Murphy for inspiring this project. J-F Houde and Allan Collard-Wexler provided especially valuable comments. We are also grateful for the comments from seminar participants at the 2011 workshop on the economics of advertising in Moscow, the 2012 Consumer Insights Conference at Yale, the 2013 ZICE conference at the University of Zurich, Cheung Kong GSB, Indiana University, New York University, Santa Clara University, the University of Colorado at Boulder, and the University of Illinois at Urbana-Champaign. We are grateful to Jeff Eastman, Grace Hyatt and Todd Kaiser at Nielsen, for their assistance with the collection of the PanelViews Survey and to Art Middlebrooks, Svetlozar Nestorov, and the James M. Kilts Center for Marketing at the University of Chicago Booth School of Business for their help with the Homescan and RMS data. We thank the Neubauer Family Foundation, the Initiative on Global Markets at the University of Chicago Booth School of Business, the Marketing Science Institute, and the Netherlands Foundation for Scientific Research (NWO, Vici grant) for financial support. Results in this paper are calculated based on data from The Nielsen Company (US), LLC.

E-mail addresses: bart.bronnenberg@tilburguniversity.edu jdube@chicagobooth.edu, matthew.gentzkow@chicagobooth.edu, jesse.shapiro@chicagobooth.edu.

1 Introduction

A 100-tablet package of 325mg Bayer Aspirin costs $6.29 at cvs.com. A 100-tablet package of 325mg CVS private label aspirin costs $1.99 (CVS 2013). The two goods share the same dosage, directions, and active ingredient: acetyl salicylic acid. Aspirin has been sold in the United States for more than 100 years, CVS directs consumers to compare Bayer to the CVS alternative, and CVS is the largest pharmacy chain in the country, with presumably little incentive to sell a faulty product. Yet the prevailing prices are evidence that some consumers are willing to pay a three-fold premium to buy the branded product. Indeed, in data we introduce below, 24 percent of aspirin sales by volume (and 59 percent by expenditure) are to branded products. Research shows that the prices of automobiles (Sullivan 1998), index funds (Hortaçsu and Syverson 2004), and online books (Smith and Brynjolfsson 2001) all embed substantial brand premia even within groups of physically homogeneous products.

Why are consumers willing to pay more to buy brands? One possibility is that they are simply making a mistake. Economists have long hypothesized that the demand for brands may reflect misinformation induced by advertising.[1] Recent theoretical work models competition and regulation in markets in which sophisticated firms compete for the business of uninformed or manipulable consumers (Gabaix and Laibson 2006; Piccione and Spiegler 2012; Ellison and Wolitzky 2012). The ability of firms to mislead consumers is a key motivation for federal regulation of advertising (Federal Trade Commission 1999). A second possibility is that brands actually produce more utility. Becker and Murphy (1993) develop a model in which advertising is a complement to consumption. Kamenica et al. (2013) present evidence for such complementarity in the case of drugs, suggesting that advertising may enhance placebo-like effects and so make branded drugs more clinically effective. Even seemingly homogeneous products may differ in subtle ways that affect objective quality.[2]

In this paper, we ask how much of the brand premium for drug-store and supermarket products results from lack of knowledge or sophistication. We match individual purchase data from the 2004-2010 Nielsen

[1] Braithwaite (1928) writes that advertisements "exaggerate the uses and merits" of brands, citing aspirin and soap flakes as examples. Henry Simons (1948) advocates government regulation of advertising to help mitigate "the uninformed consumer's rational disposition to 'play safe' by buying recognized, national brands" (p. 247). Scherer (1970) discusses premium prices for branded drugs and bleach, and writes that "it is hard to avoid concluding that if the housewife-consumer were informed about the merits of alternative products by some medium more objective than advertising and other image-enhancing devices, her readiness to pay price premiums as large as those observed here would be attenuated" (pp. 329-332).

[2] Branded and generic drugs may differ in size and shape, as well as in non-active ingredients such as those in their coatings. In one instance, the FDA determined that a generic antidepressant performed less well than its branded counterpart, likely due to differences in their "extended release" coatings (Thomas 2012). A widely publicized 2006 recall of store-brand acetaminophen resulted from the discovery that some pills could contain metal fragments (Associated Press 2006); such risks could conceivably be lower for brands. Hortaçsu and Syverson (2004) conclude that purchases of high-cost "brand name" index funds partly reflect willingness to pay for non-financial objective attributes such as tax exposure and the number of other funds in the same family.

Homescan panel to a new survey containing proxies for consumer information, and to separate data on store-level quantities and prices. We estimate the effect of our knowledge measures on the propensity to choose private labels over brands, and study the choices of experts such as pharmacists and physicians as an approximation to behavior under perfect information. We then use these estimates to ask how aggregate outcomes would differ in a world in which all consumers were perfectly informed.

The economic value at stake is large. We estimate that consumers spend \$166 billion in product categories where a comparable private label exists to branded goods. In these categories, households would reduce expenditure by \$32 billion, or 19 percent of total expenditure, by switching to buying all private labels at current prices. More broadly, knowing the source of the brand premium is an important input to evaluating the welfare effects of the roughly \$140 billion spent each year on advertising (Kantar Media 2013).

Our main identification challenge is separating the effect of consumer knowledge from correlated differences in other drivers of choice such as preferences. We limit the scope for preference differences by focusing on choices between branded and private label products that are identical on all physical attributes measured by Nielsen. We include detailed controls for income and other demographics, and compare occupations such as physicians and lawyers with similar socio-economic status but different levels of product-specific expertise. We show that high- and low-knowledge consumers look similar in their preferences for measured product attributes, making it more plausible that they are similar in their preferences for any unmeasured attributes. To address differences in private label availability, we compare consumers shopping at the same chain in the same market in the same week. To address confounds related directly to consumers' workplaces (e.g., pharmacists receive free samples or discounts that affect their purchasing behavior), we look at expert consumers who are not currently employed. Though we cannot rule out all possible confounds, the pattern of evidence suggests our estimates mainly capture the causal effect of information.

We begin our analysis with a detailed case study of headache remedies. As indirect measures of knowledge, we use information on occupation, along with educational attainment and college major. As a direct measure, we use a survey question answered by a subset of our panelists that asked them to name the active ingredient in various branded headache remedies. The average respondent in this subset can correctly identify the active ingredient 59 percent of the time. For the college-educated, the fraction rises to 61 percent. For those whose major was science or health, the fraction is 73 percent. For registered nurses, it is 85 percent, for pharmacists it is 89 percent, and for physicians it is 90 percent. Occupational specialty is important enough to outweigh large differences in general education. For example, registered nurses are far better informed about headache remedies than lawyers, despite having completed less schooling and earning less in the labor market on average.

The propensity to choose private label over branded headache remedies closely tracks knowledge. The average household devotes 71 percent of headache remedy purchases to private labels. Controlling for household income, other demographics, and market-chain fixed effects, a shopper who identifies all active ingredients is 20 percentage points more likely to purchase a private label than a shopper who identifies none. With the same controls, having a college-educated primary shopper predicts an increase of 4 percentage points, having a primary shopper with a healthcare occupation other than physician or pharmacist predicts an increase of 9 percentage points, and having a primary shopper who is a physician or pharmacist predicts an increase of 18 percentage points. Primary shoppers with health or science majors buy more private labels than those with other college degrees, and the effects of healthcare occupations are sizable among consumers not currently employed.

We predict that in a world where all households had a primary shopper as informed as a pharmacist or physician, the market share of branded headache remedies would fall by 55 percent. Expenditure on branded products would fall by 50 percent, and total headache remedy expenditure would fall by 15 percent. If we could isolate purchases for which both private label and branded alternatives are available, the effects would likely be even bigger. We conclude that a substantial majority of the brand premium in these categories is due to misinformation.

We next expand the analysis to a larger set of 34 health-related product categories, where we find effects that are qualitatively similar. The effects of knowledge of headache remedy active ingredients, college education and working in health care occupations are positive for the overwhelming majority of categories. Making all consumers as informed as a pharmacist would reduce the market share of brands in these categories by 21 percent, expenditure on brands by 24 percent, and total expenditure by 5 percent.

In the following section of the paper, we turn from health products to food and drink products. We present a case study of pantry staples such as salt, sugar, and baking powder, using working as a chef or food preparer as a proxy for knowledge, again finding consistent positive effects. Chefs devote more than 80 percent of their purchases to private labels in these categories, as compared to 61 percent for the average consumer. Controlling for other drivers of choice, chefs are 13 percentage points more likely to purchase private labels than the average consumer, and food preparers who are not chefs are 3 percentage points more likely to buy private labels. Making all consumers as informed as a chef would reduce the market share of brands and expenditure on brands in these categories by 24 percent. Because private labels are only 8 percent cheaper than brands in these categories on average, the effect on total expenditure is a reduction of only 1 percent.

We then present results for 243 other food and drink categories. Here, neither college education nor working as a chef or other food preparer has consistently positive effects on private label purchases. The coefficient on being a chef, for example, is positive for 147 categories and negative for 96, with 25 of the positive coefficients and 16 of the negative coefficients significantly different from zero. Making all consumers as informed as a chef would reduce expenditure on brands and brand market share by a significant but economically small 2 percent.

Taken together, the results suggest that in categories such as headache remedies and pantry staples which we singled out for having particularly homogeneous products, the share of brand premia due to misinformation is large and the real utility differences are small or negligible. In other categories such as carbonated soft drinks, yogurt, and many other food and drink products, the real utility differences may be bigger, and the role of misinformation consequently less. In aggregate, we estimate that change in expenditure in a world where all consumers are well-informed would be \$410 million in headache remedies, \$340 million in other healthcare categories, \$20 million in pantry staples, and \$340 million in other food and drink categories, for a total of \$1.1 billion. Thus, the total effect of misinformation is significant, but far smaller than the \$32 billion cost we would infer if all private labels produced equivalent utility to brands.

In the final section of the paper, we ask to what extent the knowledge effects we measure are domain-specific. We show that neither knowledge of headache remedy active ingredients nor working in a healthcare occupation predict private label purchases in pantry staple categories. Similarly, working in a food preparer occupation other than chef does not predict private label headache remedy purchases. We do find that chefs buy more private label headache remedies, possibly suggesting that some of their knowledge is transferable across domains.

The primary contribution of this study is to use novel data and methods to estimate the importance of information in consumer choice.[3] We add to existing survey and experimental evidence[4] by exploiting the large variation in consumer information induced by occupational expertise.[5] Of particular relevance

[3] A sizable existing literature examines the demographic and attitudinal correlates of purchasing private label consumer packaged goods (e.g., Berges et al. 2009; Dick et al. 1995; Richardson et al. 1996; Burton et al. 1998; Sethuraman and Cole 1999; Kumar and Steenkamp 2007; Steenkamp et al. 2010) and generic prescription drugs (e.g. Shrank et al. 2009). A literature on blind taste tests finds that consumers cannot distinguish among brands (Husband and Godfrey 1934; Allison and Uhl 1964) or between branded and generic goods (Pronko and Bowles 1949), though there are exceptions (Mason and Batch 2009). Wills and Mueller (1989) and Caves and Greene (1996) use aggregate data to estimate the role of advertising and quality in brand premia. Sethuraman and Cole (1999) analyze the drivers of willingness to pay for brands using hypothetical choices reported on a survey.

[4] Existing evidence indicates that perceptions of similarity between branded and private label painkillers are correlated with stated purchase intentions (Cox et al. 1983; Sullivan et al. 1994). Cox et al. (1983) find that informing consumers of active ingredient similarity does not have a discernible effect on purchase selections.

[5] We are not aware of other research on the brand preferences of health-care professionals. An existing literature examines the health behaviors of doctors (Glanz et al. 1982), including their propensities to use certain categories of medications like sleeping pills (Domenighetti et al. 1991). Most studies of the relationship between occupation and private label purchases code occupation at a high level of aggregation (white collar, blue collar, etc.) without reference to specific expertise (see Szymanski and Busch 1987

is concurrent work by Carrera and Villas-Boas (2013), who use a field experiment to assess the impact of informative product labels on the propensity to purchase private label headache remedies. Our approach of comparing the choices of experts to those made by average consumers is close in spirit to Levitt and Syverson (2008), who look at real estate agents selling their own homes, and to Johnson and Rehavi (2013), who look at the frequency with which physicians give birth by caesarean section.

Methodologically, we build on the large literature using discrete-choice demand models to predict brand choice and evaluate the sources of household purchase heterogeneity (Guadagni and Little 1983, Rossi et al. 1996). Our counterfactuals build especially on recent work that uses an equilibrium framework to evaluate the size and determinants of brand premia (Goldfarb et al. 2009). Our methods also relate to earlier work that uses the behavior of informed agents to predict the underlying preferences of uninformed agents, as in Bartels' (1996) study of voting.

Our results on health-related products also speak to issues in health policy. Purchases of branded prescription drugs in categories where generic alternatives are available is a significant component of health costs (Haas et al. 2005). A range of policies including mandatory substitution (NIHCM 2002) and financial incentives for physicians (Endsley et al. 2006) and patients (Huskamp et al. 2003) have been used in an effort to increase the generic share. Although we focus on over-the-counter products, understanding the sources of willingness to pay for private label drugs is relevant to the design of such policies and to evaluating the welfare costs of current spending on brands.

The remainder of the paper is organized as follows. Section 2 describes our data. Section 3 lays out our empirical strategy. Section 4 presents our results for health products and section 5 presents our results for food categories. Section 6 presents evidence on domain specificity. Section 7 concludes.

2 Data

2.1 The Nielsen Homescan Panel

The backbone of our data is the Nielsen Homescan Panel, which we obtained through a partnership between the Nielsen Company and the James M. Kilts Center for Marketing at the University of Chicago Booth School of Business.[6] The data include purchases made on more than 66 million shopping trips by 112,921 households from 2004 to 2010. Panelist households are given optical scanners and are asked to scan the barcodes of all consumer packaged goods they purchase, regardless of outlet or store format. The data

for a review). An exception is Darden and Howell (1987), who study the effect of retail work experience on elements of "shopping orientation," such as attitudes toward store clerks.

[6]Information on access to the data is available at http://research.chicagobooth.edu/nielsen/.

include purchases not only from supermarkets, but also from convenience stores, mass merchandisers and club stores, drug stores, and so on. See Einav et al. (2010) for a recent validation study of the Homescan panel.

For each panelist household, we have a set of demographics including the education of the household head, a categorical measure of household income, number of adults, race, age, household composition, and home ownership. We also know the Scantrak market in which each household resides. A Scantrak market is a geographic area defined by Nielsen and is either a metropolitan area (e.g., Chicago), a combination of nearby cities (e.g., Hartford - New Haven), or a part of a state (e.g., West-Texas). We will refer to Scantrak markets as markets for the remainder of the paper.

For each purchase, we know the date, the universal product classification (UPC) code, the transaction price, an identifier for the store chain in which the purchase was made, and the size of the UPC in so-called "equivalent units" specific to product categories, e.g., pill counts for headache remedies or ounces for salt.

2.2 PanelViews Surveys

We conducted two surveys of Homescan Panelists using Nielsen's PanelViews survey, a monthly survey administered to the Nielsen Homescan panel. The first survey was sent electronically to 75,221 households in October of 2008 with the request that each adult in the household complete the survey separately. In total, 80,077 individuals in 48,501 households responded to the survey for a household response rate of 64.5 percent. The second survey was sent electronically to 90,393 households in November 2011 with the request that each adult in the household complete the survey separately. In total, 80,205 individuals in 56,258 households responded to the survey for a household response rate of 62.2 percent.

Both surveys asked for the respondent's current or most recent occupation, classified according to the 2002 Bureau of Labor Statistics (BLS) codes (BLS 2002).[7] We match these to data on the median earnings of full-time full-year workers in each occupation in 1999 from the US Census (2000). We group occupations into categories (health care, food preparer) using a combination of BLS-provided hierarchies and subjective judgment. The online appendix lists the occupations in these groupings.

The first survey included a set of additional questions relating to household migration patterns. These questions were used in the analysis of Bronnenberg et al. (2012). We ignore them in the present analysis.

The second survey, designed for this study, included a series of questions about households' knowledge and attitudes toward various products. In particular, for each of five national brands of headache remedy

[7] In the small number of cases where an individual provided conflicting responses to the occupation question across the two surveys we use the value from the second survey.

(Advil, Aleve, Bayer, Excedrin, Tylenol), we asked each respondent who indicated familiarity with a brand to identify its active ingredient from a list of six possible choices, or state that they "don't know."[8] The correct active ingredients are ibuprofen (Advil), naproxen (Aleve), aspirin (Bayer), aspirin-acetaminophen-caffeine (Excedrin), and acetaminophen (Tylenol). For each respondent we calculate the number of correct responses, treating "don't know" as incorrect. We also asked respondents whether they agreed or disagreed with a series of statements, including "Store brand products for headache remedy / pain relievers are just as safe as the brand name products," with responses on a 1 (agree) to 7 (disagree) scale. For each respondent, we code a dummy variable equal to one if they chose 1 (the strongest possible agreement) and zero otherwise.

The second survey also asked respondents about their college major using codes from the National Center for Education Statistics (U.S. Department of Education 2012). We define two groups of majors for analysis: health majors, which include all majors with the word "health" in their description, and non-health science majors, which include all majors in the physical and biological sciences.

Both surveys asked respondents to indicate whether they are their household's "primary shopper" and whether they are the "head of the household." For each household we identify a single shopper whose characteristics we use in the analysis, following the criteria used in Bronnenberg et al. (2012). We start with all individuals within a household who respond to the survey. We then apply the following criteria in order, stopping at the point when only a single individual is left: (i) keep only primary shopper(s) if at least one exists; (ii) keep only household head(s) if at least one exists; (iii) keep only the female household head if both a female and a male head exist; (iv) keep the oldest individual; (v) drop responses that appear to be duplicate responses by the same individual; (vi) select one respondent randomly.

Throughout the paper, we restrict attention to households that answered the occupation question in one or both of our PanelViews surveys.

2.3 Product Classification

Nielsen provides a set of attribute variables for each UPC code purchased by a Homescan panelist. Some of these, such as size, are available for all categories. Others are category-specific. For example the data include a variable that encodes the active ingredient for each headache remedy in the data. We harmonize the codes for essentially identical descriptors (e.g., "ACET" and "ACETAMINOPHEN" both become "ACETAMINOPHEN").

We use these descriptors to aggregate UPCs into *products*. A product is a group of UPCs that are identical on all non-size attributes provided by Nielsen. For instance, in the case of headache remedies, a

[8] In each case, the six choices were the five correct active ingredients plus the analgesic hydrocodone.

product is a combination of an active ingredient (e.g., aspirin, naproxen), form (e.g., tablet, gelcap), formula (e.g., regular strength, extra strength), and brand (e.g., Bayer, Aleve, private label).

To compare private label and branded goods we aggregate products into *comparable product groups*, which are sets of products that are identical on all product attributes except for brand and size.

To perform our analysis we consider comparable product groups in which we observe at least 500 purchases with at least some purchases going to both branded and private label products.[9] We eliminate categories in which the available attribute descriptors do not provide sufficient information to identify comparable products.[10] This leaves us with 395 comparable product groups.

For our case study of headache remedies we consider the subset of these comparable product groups classified by Nielsen as adult daytime headache remedies.

For our case study of pantry staples we consider the subset of these comparable product groups classified by Nielsen as table salt, sugar, or baking soda.

In our analysis, we restrict attention to transactions such that at least one comparable branded purchase and at least one comparable private label purchase are observed in the Homescan data in the same retail chain and quarter as the given transaction. We use this restriction to limit the likelihood that a branded product is purchased because no private label alternative is available (or vice versa).

2.4 Retail Scanner Data

To estimate prices and aggregate expenditure, we use 2008 store-level scanner data from the Nielsen Retail Measurement Services (RMS) files, which, like the Homescan data, we obtained through a partnership between Nielsen and Chicago Booth's Kilts Center for marketing. These data contain store-level revenue and volume by UPC and week for approximately 35,000 stores in over 100 retail chains. We use our product classification to aggregate UPCs into products.

For each comparable product group, we compute average price per equivalent unit for brands and private labels respectively as the ratio of total expenditure to total equivalent units across all grocery, drug, and mass merchandise stores across all weeks in 2008. We also estimate total US expenditure on brands and generics respectively by multiplying the number of equivalent units purchased in the Homescan data by (i) the ratio of total equivalent units for the comparable group in RMS and Homescan, (ii) the average price per equivalent unit, (iii) the ratio of 2008 US food, drug, and mass merchandise sales to total 2008 expenditure measured

[9] We further eliminate comparable product groups in which fewer than 50 retail chains ever sell a private label according to the retail scanner data we discuss in section 2.4 below.

[10] These are: deli products, fresh produce, nutritional supplements, miscellaneous vitamins, and anti-sleep products.

in RMS.[11]

The sum of estimated total US expenditure across the comparable product groups in our sample is \$166 billion. If all observed equivalent units were purchased at the average price per equivalent unit of generics, this sum would fall by \$32 billion or 19 percent.

3 Empirical Strategy

Our goal is to estimate the effect of information on the propensity to buy private label products. To establish the direction of this effect, we look at the relationship between purchases and a range of information proxies, including knowledge of active ingredients, completed schooling, college major, and occupation. To pin down the overall share of brand premia due to limited information, we focus on the choices of occupational experts such as pharmacists, physicians, and chefs, whose behavior we take as an approximation to choices under full information.

The main specification we estimate is a linear probability model:

$$Pr(y_{ij} = 1) = \alpha_j + S_i\beta_j + X_i\gamma_j, \tag{1}$$

where an observation is a purchase by consumer i in comparable product group j, y_{ij} is an indicator equal to one if the purchase was a private label and zero if it was a brand, S_i is a vector of information proxies of interest, X_i is a vector of controls, and β_j and γ_j are vectors of parameters. In our preferred specification, X_i includes detailed demographic and income controls as well as chain-by-market fixed effects. We weight observations by volume purchased and cluster standard errors by household. Since we treat S_i as proxies, the parameters of interest β_j include both the causal effect of the components S_i per se (e.g., knowing that Tylenol's active ingredient is acetaminophen directly affects purchases) and the effect of information that is correlated with S_i (e.g., consumers who know Tylenol's active ingredient also tend to be well informed about other characteristics of headache remedies). The key assumption necessary to interpret the estimated β_j as causal is therefore that the information proxies in S_i are orthogonal to *non-informational* drivers of choice conditional on X_i.

Including chain-by-market fixed effects addresses possible correlation between S_i and the shopping environment. Our results are robust to the possibility that sophisticated shoppers systematically choose chains

[11] The Annual Retail Trade Survey of the United States Census Bureau reports 2008 annual sales in grocery stores, pharmacies and drug stores, and warehouse clubs and superstores of \$511 billion, \$211 billion, and \$352 billion, respectively, totalling \$1,075 billion (U.S. Census, 2013).

that differ in the quality or availability of private labels, or in their prices, promotions, and so forth. We show in the appendix that our results survive even richer controls for the timing and location of purchases.

A different concern is that S_i may be correlated with preferences. More sophisticated shoppers may have different tastes for product attributes rather than different beliefs or information. We limit the scope for this kind of confound by studying the choice of private label vs. branded products among groups of products whose external attributes are identical. Variation in preferences over these measured attributes thus cannot explain variation in the propensity to purchase private label products within our comparable product groups. We address some of the remaining confounds by the controls in Z_i and by comparing occupational groups such as physicians and lawyers with similar socio-economic status but different expertise. We also show that high- and low-knowledge consumers look similar in their preferences for measured product attributes.

Though we argue that this evidence taken together strongly points toward an effect of information, we cannot rule out all possible confounds. What remains is the possibility that brands and private labels within our comparable product groups do differ in objective quality, and that the willingness to pay for quality is correlated with expertise. We note that under the plausible assumption that any such correlation is positive (physicians have if anything a greater taste for high-quality medicine, and chefs have if anything a greater taste for high-quality food), our estimates will be a lower bound on the true effect.

4 Results: Health Products

4.1 Headache Remedies

We begin our analysis with a case study of adult daytime headache remedies. The six relevant comparable groups in our sample span four active ingredients, each associated with a familiar brand: aspirin (Bayer), acetaminophen (Tylenol), ibuprofen (Advil), and naproxen (Aleve). For acetaminophen and ibuprofen, there are separate groups for tablets and gelcaps, while for aspirin and naproxen we consider only tablets. Both acetaminophen groups are "extra strength." All other groups are "regular strength."

The first rows of table 1 show summary statistics for these comparable groups. We estimate total annual expenditure on headache remedies to be \$2.67 billion. Private label purchases account for 71 percent of quantity and 49 percent of dollars.

On average, the per-pill price of a private label is 41 percent of the price of an equivalent brand. For aspirin, a mature product that has been off patent since 1917, private labels are 78 percent cheaper than Bayer. These price differences are not due to differences in where these products are sold or to volume discounts: among market-store-weeks in which we observe at least one branded and one private label purchase for the same active ingredient and package size, the per-pill price for private labels is 33 percent of the price of an equivalent brand. The median gap is 37 percent, and the branded good is cheaper in only 5 percent of cases.

Private label alternatives for branded headache remedies are widely available. Using our store-level data, we estimate that 83 percent of the time that a branded headache remedy is purchased, a private label with the same active ingredient and form and at least as many pills is available at a lower price. In our PanelViews survey data, only 3.6 percent of households report that no private label alternative was available at their last purchase.

In figure 1 we look at the relationship between knowledge of active ingredients and our indirect knowledge proxies — completed schooling, occupation, and college major. The relationships are strongly positive in each case. Panel A shows that shoppers with a college education correctly identify the active ingredient in 61 percent of cases, as against 52 percent for those with a high school degree or less. Panel B shows that pharmacists correctly identify the active ingredient in 89 percent of cases, nurses in 85 percent, doctors in 90 percent. Panel C shows that shoppers whose college major is health or science related are more informed than other shoppers. In the online appendix, we confirm these relationships in a regression framework, showing that they remain strong even after controlling for income and a rich set of other household demographics.

Having validated our proxies, we turn next to our main question of interest: the impact of knowledge on the share of purchases that go to private labels. Figure 2 shows that greater knowledge of active ingredients predicts. Those who can name no active ingredients buy less than 60 percent private labels. Those who can name all five active ingredients buy more than 80 percent private labels.

Figure 3 shows the relationship between private label share and completed schooling. With no controls, we see that those with education beyond high school buy more private labels than those with a high school degree or less, but that there is no clear difference between those with some college, a college degree, or more than a college degree. The main confound here is income, which is strongly negatively correlated with private label purchases (see appendix figure 1). Once we add income controls, we see there is a monotonic positive relationship between completed schooling and private label share.

Figure 4 shows the relationship between private label share and occupation. Here we see the strong negative relationship between private label share and income — in this case represented as median income by occupation. Households whose primary shopper is a healthcare professional buy far more private labels than others of similar income. Pharmacists, physicians, and nurses buy more private labels than lawyers, who have high levels of schooling but different occupational expertise.

Pharmacists, who stand out in the survey data in figure 1 as among the most informed about active ingredients, also stand out for having the largest private label share among large healthcare occupations. Only 12 percent of volume bought by pharmacists are branded headache remedies, an amount small enough to be explained by the occasional stock outs of private labels, and the fact that some purchases are made by the non-pharmacist member of a pharmacist's household.

Table 2 presents the relationship between private label share and knowledge of active ingredients in a regression framework. The table presents estimates of equation 1, where the information variables of interest S_i are a dummy for college education and the share of active ingredients known. Column (1) includes in X_i controls for demographic characteristics, product comparable group fixed effects, and market fixed effects. Column (2) adds flexible income controls, and column (3) replaces market fixed effects with market-chain fixed effects. The effect of college education increases when income controls are added and is equal to 3 percentage points in the preferred specification of column (3). Moving the share of active ingredients known from zero to one increases increases the private label share by 20 percentage points, an effect which is consistent across specifications.

The final column of table 2 adds to S_i an additional knowledge proxy: a dummy for whether consumers report believing that private labels are "just as safe" as brands. This is in some sense less clean as a measure of knowledge than naming active ingredients, since it is possible to argue that the correct answer is ambiguous. Believing private labels are just as safe as brands has an additional effect of 22 percentage points over and above the effect of active ingredient knowledge. The effect of having this belief *and* being able to name all active ingredients correctly is 37 percentage points.

Table 3 presents regression evidence on the effect of occupation. The model and controls in the first three columns are the same as in table 2, with the information variables of interest S_i now being a college education dummy, a dummy for pharmacist or physician, and a dummy for healthcare occupations other than pharmacist or physician. The estimated occupation effects remain stable as we add controls. In the preferred specification of column (3) we find that the effect of being a pharmacist or physician is 18 percentage points, and the effect of other healthcare occupations is 9 percentage points.

Column (4) of table 3 presents evidence on the role of college major. We restrict the sample to respondents who completed college and who reported their college major in our survey. We find that health majors buy an insignificant 0.2 percentage points more private labels and non-health science majors buy a highly significant 7 percentage points more private labels. Column (5) of table 3 presents occupation results for the subsample of respondents who are not currently employed for pay. (Recall that our occupation variables are defined based on the most recent employment spell.) The coefficient on other healthcare occupations is unchanged in this sample and remains highly significant. The pharmacist or physician coefficient is imprecisely estimated because we have few unemployed pharmacists or physicians. Taken together, columns (4) and (5) suggest our results are unlikely to be driven by factors specific to current employment in a healthcare profession, such as the availability of employee discounts or free samples.

Table 4 presents evidence on the extent to which our direct and indirect knowledge measures capture the same underlying variation. Column (1) repeats the preferred specification of table 3 column (3), this time restricting to respondents who participated in the wave of our survey where we asked direct knowledge questions. Column (2) restricts the sample the shoppers who named all active ingredients correctly. Column (3) adds the additional restriction that the respondent believes private labels are "just as safe" as brands. Restricting attention to well informed consumers reduces the estimated effect of education and occupation substantially, while only slightly reducing precision. In the final column, the occupation coefficients are reduced by more than 70 percent and are statistically indistinguishable from zero. These findings suggest that all of our measures capture variation along a common dimension, which we interpret as knowledge. They suggest that consumers who answer all our direct knowledge questions correctly are approximately as well informed as the average pharmacist or physician. And they suggest that our occupation results cannot be driven by correlation with unobserved preferences unless these preferences are similarly correlated with knowledge.

As further support for our identifying assumptions, appendix figure 2 shows that healthcare professionals and other consumers look similar in their choices of observed product attributes such as active ingredient and physical form.

To summarize the implications of these results, table 5 shows how overall purchases would change in a world where all consumers were as informed as a pharmacist or physician. To perform this counterfactual, we assume that the probability that consumers who are not pharmacists or physicians choose a private label on a given purchase occasion increases by the estimated effect from our preferred specification of table 3 column (3).[12] We find that both the market share of brands and expenditure on brands would fall by half, and that total expenditure on headache remedies would fall by approximately $400 million or 15 percent.

4.2 Other Health Products

We next turn to analyzing the other healthcare-related comparable product groups in our data. We restrict attention to the 34 comparable groups for which we observe at least $5,000$ purchases by households with non-missing values of our demographic controls. These include other medications such as cold remedies,

[12] For comparability with later counterfactuals, we do not use the coefficient from table 3 column (3) directly, but run separate regressions for each headache remedy comparable product group and then average the coefficients weighting by purchase volume. The individual coefficients by comparable group are shown in figure 6 below.

first aid products such as bandages, and miscellaneous products such as vitamins and contact lens solution.

Collectively these categories account for \$6.52 billion of expenditure per year. Private label purchases account for 56 percent of volume. Private label prices are half of branded prices on average.

For each comparable product group, we run a separate regression to estimate the effect of knowledge of headache remedy active ingredients (using the specification in column (3) of table 2) and of occupation (using the specification in column (3) of table 3).

Figure 5 presents coefficients and 95% confidence intervals for the estimated effect of knowledge of headache remedy active ingredients. We show results for the 6 headache remedy comparable groups as well as the 34 "other health" comparable groups. Although knowledge of these active ingredients is obviously most relevant to headache remedy purchases, we expect it to also be a good proxy for more general knowledge relevant to the other health categories. The knowledge effect is positive in 33 cases and negative in 7. Twenty-one of the positive coefficients and none of the negative coefficients are significantly different from zero. We note the coefficients tend to be larger and more significant for medications and vitamins, and relatively smaller for first aid and eye care products, perhaps suggesting larger real quality differentials among the latter.[13] Four of the five largest effects are for headache remedies.

Figures 6 and 7 present analogous coefficients for the effect of being a pharmacist or physician and the effect of other healthcare professions respectively. We see broadly similar patterns to the knowledge results, with somewhat less precision. The pharmacist and physician effect is positive in 31 cases and negative in 9. Thirteen of the positive coefficients and two of the negative coefficients are significant.[14] The other healthcare professional effect is positive in 36 cases and negative in 4, with 17 of the positive coefficients and none of the negative coefficients significantly different from zero. The pattern of larger coefficients for medications relative to first aid and eye care products remains.

We report coefficients on college education in both the knowledge and occupation specifications in the online appendix. These coefficients, too, are overwhelmingly positive.

Figure 8 plots the distribution of t-statistics on knowledge of active ingredients for health categories and non-health categories respectively. While the distribution for health categories is shifted far to the right relative to what we would expect under the null hypothesis of no effect, the distribution for non-health categories is much closer to the null hypothesis. Figure 9 shows a similar, though less dramatic,

[13]Contact lens solutions are the only health care product we have identified where some medical professionals recommend patients buy brands due to quality concerns with private labels. See Secor (2002).

[14]The comparable groups with significant negative coefficients are "laxative salts" and "alkalizing effervescents." We are not aware of any evidence that private label quality in these categories is unusually low. Given that these categories do not show negative effects of either active ingredient knowledge or other healthcare occupations, we suspect the negative pharmacist / physician effects may be attributable to sampling variation.

pattern for the coefficients on being a pharmacist or physician. Together, these results suggest that the health-related knowledge captured by our active ingredient measure may be domain specific, rather than representing general human capital that applies equally to all product types. We present more evidence on this in section 6. It may also reflect larger true quality differences between brands and private labels in non-health categories, a point we return to below.

To summarize the implications of these results, table 6 shows how purchases in our 34 non-headache-remedy health categories would change in a world where all consumers were as informed as a pharmacist or physician, following the same approach as in table 5. We find that expenditure on brands would fall by 24 percent and that the market share of brands would fall by 21 percent. Total expenditure across these categories would fall by approximately $340 million or 5 percent.

5 Results: Food and Drink Products

5.1 Pantry Staples

We now turn to the analysis of food purchases. Here our proxies for knowledge are indicators for whether the primary shopper is a chef or other food preparer. Unlike the health care case, we do not have reliable direct measures of food-related knowledge that correlate strongly with the occupational proxies.[15] We begin with a case study of pantry staples: salt, sugar, and baking soda. We choose these products because they are uniform in chemical composition and purpose, and thus analogous to headache remedies in being relatively homogeneous.

The bottom section of table 1 includes summary statistics for the six comparable groups we classify as pantry staples: baking soda; regular iodized and plain salt (sold in boxes); and regular granulated, light brown, and powdered sugar (sold in bags). Collectively, these categories account for $1.81 billion of expenditure. Private label purchases account for 60 percent of volume and 58 percent of expenditure. On average, the price-per-equivalent-volume for private labels is 8 percent less than for brands. The private-label-to-brand price ratios range from a low of 0.75 for plain salt to 0.92 for granulated sugar, the category which dominates pantry staple expenditure.

Figure 10 shows shows the relationship between private label share and occupation. As with headache remedies, we see a strong negative relationship between private label share and median occupational income.

[15] Our 2011 PanelViews survey asked respondents to identify the most common additive to table salt (correct answer: iodine), the scientific name for baking soda (correct answer: sodium bicarbonate), and the most common ingredient of granulated sugar (correct answer: sucrose). The share of these questions answered correctly is positively correlated with working as a chef or food preparer and with purchasing private label pantry staples, but neither relationship is consistently statistically significant. Results for these knowledge measures are presented in the online appendix.

Households whose primary shopper is a food preparer or manager buy more private labels than others of similar occupational income. Chefs — the occupational group we would have expected *ex ante* to be most informed about the quality of food products — buy more than 80 percent private labels in these categories, more than any of the other occupation categories in the figure.

Table 7 shows the relationship with occupation in a regression framework. The specifications in the five columns are the same as in table 3, with the information proxies of interest S_i now consisting of a dummy for college education, a dummy for being a chef, and a dummy for being a food preparer but not a chef. In our preferred specification of column (3), we estimate that being a chef increases the probability of buying private labels by 13 percentage points, and working in a non-chef food preparation occupation increases this probability by 3 percentage points. The magnitude of these effects are somewhat smaller than in the specifications without controls. In contrast to headache remedies, we do not find any clear effect of college education. Column (4) shows that non-health science majors buy significantly more private label pantry staples, while health majors look similar to other consumers. Column (5) shows that the coefficients on being a chef and other food preparer increase substantially when we focus on shoppers who are not currently employed, suggesting our effects are not driven by mechanical effects of employment in the food industry.

We summarize the implications of these results in table 8, which shows how pantry staple purchases would change in a world where all consumers were as informed as a chef, following the same approach as in table 5. We find that the market share of brands and expenditure on brands would both fall by a quarter in these categories. The decline in total expenditure is relatively small — \$20 million or 1 percent — because private labels in these categories are only 8 percent cheaper than brands on average.

5.2 Other Food and Drink Products

Finally, we consider the remaining comparable groups for food and drink products in our data. We restrict attention to the 243 comparable groups for which we observe at least $5,000$ purchases by households with non-missing values of our demographic controls. They comprise a broad cross-section of supermarket products, from milk and eggs, to carbonated beverages, to ready-to-eat cereal.

Collectively, these categories account for \$122 billion of expenditure. Private label purchases account for 42 percent of volume. On average, the price-per-equivalent-volume for private labels is 29 percent less than for brands.

For each comparable product group, we run a separate regression to estimate the effect of working as a chef or other food preparer on private label purchases (using the specification in column (3) of table 7).

Figure 11 summarizes the estimated coefficients and 95% confidence intervals. Rather than try to present all 243 coefficients in a single figure, we aggregate all comparable groups other than pantry staples up to a higher level of aggregation that Nielsen calls "product groups," weighting the individual comparables by precision and computing the aggregate standard error assuming that the individual coefficients are statistically independent. Thus, for example, the comparables for cola, diet cola, lemon-lime soda, and so forth are combined into the Nielsen product group "carbonated beverages." For pantry staples, we present individual comparables separately.

The estimated effects of knowledge on private label purchases in these categories are less overwhelmingly positive than what we saw for health products and pantry staples. The coefficients on working as a chef are positive for 147 comparables and negative for 96, with 25 of the positive coefficients and 16 of the negative coefficients significantly different from zero. Figure 11 shows that those effects which are significant are generally small in magnitude. The pantry staples categories stand out as having among the most positive and significant coefficients: plain salt has the largest coefficient in the figure, and four of the top eight coefficients are pantry staples. In the online appendix, we present the distribution of coefficients for working in other food preparation occupations and for college education.

Table 9 summarizes the implications of these results, showing how other food product purchases would change in a world where all consumers were as informed as a chef. We again follow the same approach as in table 5. We find that the market share of brands and expenditure on brands would fall statistically significant but economically small 2 percent. The decline in total expenditure is estimated to be an imprecisely estimated $340 million per year.

6 Evidence on Domain Specificity

The results above suggest that health experts purchase more private label health products and food experts purchase more private label food products. A natural follow-up question is to what extent experts' knowledge is transferable outside of their domain of expertise. Perhaps pharmacists' understanding of the equivalence of branded and private label headache remedies leads them to also recognize the likely equivalence of branded and private label baking soda. Or perhaps their understanding does not translate beyond the categories they are directly familiar with.

Table 10 presents evidence on domain specificity. The first two columns look at the effect of healthcare expertise on pantry staple purchases. Column (1) shows that the share of headache remedy active ingredients known has no significant effect on the probability of purchasing private label pantry staples, with a confidence interval that rules out coefficients greater than 1.5 percentage points. Column (2) shows that pharmacists, physicians, and other health professionals are also not significantly more likely to buy private label pantry staples. The confidence intervals on the pharmacist-physician and other health occupation coefficients rule out effects greater than 7 percentage points and 2.3 percentage points respectively. We can strongly reject the hypothesis that these effects are as large as the effects we estimate for headache remedy purchases. The evidence thus suggests that healthcare expertise does not translate to behavior outside the health domain, consistent with past evidence on the domain specificity of expertise (Levitt et al. 2010).

The final column of table 10 looks at the effect of food preparation expertise on headache remedy purchases. Here, we do see some evidence of transferability: chefs are a statistically significant 10 percentage points more likely to buy private label headache remedies than other consumers. There is no significant effect for food preparers other than chefs.

7 Conclusions

We use scanner data coupled with a novel survey to estimate the effect of consumer sophistication on the propensity to purchase private label products in physically homogeneous product categories. In a detailed case study of headache remedies we find that college education, working in a healthcare occupation, and other proxies for product knowledge predict private label purchases. Pharmacists devote almost 90 percent of category purchases to private labels, against 71 percent overall. A second case study of pantry staples (salt, sugar and flour) shows effects of knowledge that are qualitatively similar but smaller in magnitude. We extend the analysis to a broad set of categories to compute the overall change in grocery purchases and spending from a shift to more informed consumers.

References

Allison, Ralph I. and Kenneth P. Uhl. 1964. Influence of beer brand identification on taste perception. *Journal of Marketing Research* 1(3): 36-39.

Associated Press. 2006. Contaminated with metal, a painkiller is recalled. Accessed at http://www.nytimes.com/2006/11/10/health/10recall.html on March 29, 2013.

Bartels, Larry M. 1996. Uninformed votes: Information effects in presidential elections. *American Journal of Political Science* 40(1): 194-230.

Becker, Gary S. and Kevin M. Murphy. 1993. A simple theory of advertising as a good or bad. *The Quarterly Journal of Economics* 108(4): 941-964.

Berges, Fabian, Daniel Hassen, Sylvette Monier-Dilhan, and Hélène Raynal. 2009. Consumers' decision between private labels and national brands in a retailer's chain: A mixed multinomial logit application. *Gestion 2000* 3: 41-57.

BLS. 2002. *Census Occupation Classification*. Accesed at http://www.bls.gov/cps/cenocc.pdf in 2010.

Braithwaite, Dorothea. 1928. The economic effects of advertisement. *The Economic Journal* 38(149): 16-37.

Bronnenberg, Bart J., Jean-Pierre Dubé, and Matthew Gentzkow. 2012. The evolution of brand preferences: Evidence from consumer packaged goods. *American Economic Review* 102(6): 2472-2508.

Burton, Scot, Donald R. Lichtenstein, Richard G. Netemeyer and Judith A. Garretson. 1998. A scale for measuring attitude toward private label products and an examination of its psychological and behavioral correlates. *Journal of the Academy of Marketing Science* 26(4): 293-306.

Carrera, Mariana and Sofia Villas-Boas. 2013. Generic aversion and observational learning in the over-the-counter drug market. Case Western mimeo.

Caves, Richard E. and David P. Greene. 1996. Brands' quality levels, prices, and advertising outlays: Empirical evidence on signals and information costs. *International Journal of Industrial Organization* 14(1): 29-52.

Cox, Steven R., Kenneth A. Coney, and Peter F. Ruppe. 1983. The impact of comparative product ingredient information. *Journal of Public Policy & Marketing* 2, 57-69.

CVS. 2013. *CVS Shop*. Accessed at http://www.cvs.com/shop/product-detail/Bayer-Aspirin-Tablets-Easy-Open-Cap?skuId=100073 on March 29, 2013.

Darden, William R. and Roy D. Howell. 1987. Socialization effects of retail work experience on shopping orientations. *Journal of the Academy of Marketing Science* 15(3): 52-63.

Dick, Alan, Arun Jain, and Paul Richardson. 1995. Correlates of store brand proneness: Some empirical observations. *Journal of Product & Brand Management* 4(4): 15-22.

Domenighetti, Gianfranco, Michele Tomamichel, Felix Gutzwiller, Silvio Berthoud, and Antoine Casabianca. 1991. Psychoactive drug use among medical doctors is higher than in the general population 33(3): 269-274.

Einav, Liran, Ephraim Leibtag, and Aviv Nevo. 2010. Recording discrepancies in Nielsen Homescan data: Are they present and do they matter? *Quantitative Marketing and Economics* 8(2): 207-239.

Ellison, Glenn and Alexander Wolitzky. 2012. A search cost model of obfuscation. *RAND Journal of Economics* 43(3): 417-441.

Endsley, Scott, Geof Baker, Bernard A. Kershner, and Kathleen Curtin. 2006. What family physicians need to know about pay for performance. *Family Practice Management* 13(7): 69-74.

Federal Trade Commission. 1999. *Final Order In the Matter of Novartis Corporation, et al., Docket 9279*. Accessed at http://www.ftc.gov/os/decisions/docs/Volume127.pdf on January 18, 2013.

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

Glanz, Karen, Stanley B. Fiel, Lawrence R. Walker, and Marvin R. Levy. 1982. Preventive health behavior of physicians. *Journal of Medical Education* 57(8): 637-639.

Goldfarb, Avi, Quiang Lu, and Sridhar Moorthy. 2009. Measuring brand value in an equilibrium framework. *Marketing Science* 28(1): 69-86.

Guadagni, Peter M., and John D.C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Science* 2(3): 203-238.

Haas, Jennifer S., Kathryn A. Phillips, Eric P. Gerstenberger, and Andrew C. Seger. 2005. Potential savings from substituting generic drugs for brand-name drugs: Medicare expenditure panel survey, 1997-2000. *Annals of Internal Medicine.* 142: 891-897.

Hortaçsu, Ali and Chad Syverson. 2004. Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds. *The Quarterly Journal of Economics* 119(2): 403-456.

Husband, R.W. and J. Godfrey. 1934. An experimental study of cigarette identification. *Journal of Applied Psychology* 18(2): 220-223.

Huskamp, Haiden A., Patricia A. Deverka, Arnold M. Epstein, Robert S. Epstein, Kimberly A. McGuigan, and Richard G. Frank. 2003. The effect of incentive-based formularies on prescription-drug utilization and spending. *New England Journal of Medicine* 349(23): 2224-2232.

Johnson, Erin M. and M. Marit Rehavi. 2013. Physicians treating physicians: Information and incentives in childbirth. MIT mimeo.

Kamenica, Emir, Robert Naclerio, and Anup Malani. 2013. Advertisements impact the physiological efficacy of a branded drug. Working Paper.

Kantar Media. 2013. Kantar Media reports U.S. advertising expenditures increased 3 percent in 2012. March 11. Accessed at http://kantarmediana.com/intelligence/press/us-advertising-expenditures-increased-3-percent-2012 on April 23, 2013.

Kumar, Nirmalya and Jan-Benedict E.M. Steenkamp. 2007. *Private Label Strategy*. Cambridge, MA: Harvard Business School Press.

Levitt, Steven D., John A. List, and David H. Reiley. 2010. What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments. *Econometrica* 78(4): 1413-1434.

Levitt, Steven D. and Chad Syverson. 2008. Market distortions when agents are better informed: The value of information in real estate transactions. *Review of Economics Studies* 90(4): 599-611.

Mason, Kevin and Alice Batch. 2009. What's in a (brand) name? *Proceedings of ASBBS* 16(1).

National Institute for Health Care Management (NIHCM). 2002. *A Primer: Generic Drugs, Patents, and the Pharmaceutical Marketplace*. Washington DC: National Institute for Health Care Management Foundation.

Piccione, Michele and Ran Spiegler. 2012. Price competition under limited comparability. *The Quarterly Journal of Economics* 127(1): 97-135.

Pronko, N.H. and J.W. Bowles Jr. 1949. Identification of cola beverages. III. A final study. *Journal of Applied Psychology* 33(6): 605-608.

Richardson, Paul S., Arun K. Jain, and Alan Dick. 1996. Household store brand proneness: A framework. *Journal of Retailing* 72(2): 159-185.

Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Science* 15(4): 321-340.

Scherer, Frederic M. 1970. *Industrial Market Structure and Economic Performance*. Chicago, IL: Rand McNally.

Secor, Glenda B. 2002. Educate patients about inconsistency of generic contact lens solutions. *Primary Care Optometry News*. May. Accessed at http://www.healio.com/optometry/contact-lenses-eyewear/news/print/primary-care-optometry-news/%7Ba6af06d9-ffe3-4a21-88a2-10f963172c5a%7D/educate-patients-about-inconsistency-of-generic-contact-lens-solutions on May 24, 2013.

Sethuraman, Raj and Catherine Cole. 1999. Factors influencing the price premiums that consumers pay for national brands over store brands. *Journal of Product & Brand Management* 8(4): 340-351.

Shrank, William H., Emily R. Cox, Michael A. Fischer, Jyostna Mehta, and Niteesh K. Choudhry. 2009. Patients' perceptions of generic medications. *Health Affairs* 28(2): 546-556.

Simons, Henry. 1948. *Economic Policy for a Free Society*. Chicago, IL: University of Chicago Press.

Smith, Michael D. and Erik Brynjolfsson. 2001. Consumer decision-making at an Internet shopbot: Brand still matters. *Journal of Industrial Economics* 49(4): 541-558.

Steenkamp, Jan-Benedict E.M., Harold J. van Heerde, and Inge Geyskens. 2010. What makes consumers willing to pay a premium for national brands over private labels. *Journal of Marketing Research* 47(6): 1011-1024.

Sullivan, Donald L., Stephen W. Birdwell, and Suzan N. Kucukarslan. 1994. An assessment of consumer purchasing behavior for private-label vs. brand-name over-the-counter products in chain pharmacies. *Journal of Pharmaceutical Marketing & Management* 8(1): 85-108.

Sullivan, Mary W. 1998. How brand names affect the demand for twin automobiles. *Journal of Marketing Research* 35(2): 154-165.

Szymanski, David M. and Paul S. Busch. 1987. Identifying the generics-prone consumer: A meta-analysis. *Journal of Marketing Research* 24(4): 425-431.

Thomas, Katie. 2012. An increase in scrutiny for generics. *New York Times*. October 31. Accessed at http://www.nytimes.com/2012/11/01/business/fda-increases-scrutiny-of-some-generic-drugs.html on March 29, 2013.

U.S. Census. 2000. Earnings by occupation and education, 1999. Accessed at http://www.census.gov/hhes/www/income/earnings/earnings.html on December 17, 2009.

U.S. Census. 2013. Monthly retail trade report: Estimates of monthly retail and food services sales by kind of business: 2008. Accessed at http://www.census.gov/retail/ on March 31, 2013.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics (NCES). 2012. CIP user site. Accessed at http://nces.ed.gov/ipeds/cipcode/Default.aspx?y=55 on Decemeber 28, 2012.

Wills, Robert L. and Willard F. Mueller. 1989. Brand pricing and advertising. *Southern Economic Journal* 56(2): 383-395.

Table 1: Summary statistics

	Total expenditure ($bn / year)	Private label share (volume)	Private label share ($)	Price ratio (private label / brand)
Headache remedies				
Acetaminophen gelcaps	$0.15	0.43	0.35	0.72
Ibuprofen gelcaps	$0.47	0.25	0.18	0.67
Acetaminophen tablets	$0.46	0.76	0.53	0.37
Aspirin tablets	$0.24	0.76	0.41	0.22
Ibuprofen tablets	$0.97	0.78	0.56	0.36
Naproxen sodium tablets	$0.37	0.56	0.43	0.61
Total (6)	$2.67	0.71	0.49	0.41
Other healthcare products (71)	$8.82	0.58	0.48	0.54
Pantry staples				
Baking soda	$0.13	0.33	0.29	0.80
Salt (iodized)	$0.07	0.53	0.46	0.76
Salt (plain)	$0.04	0.46	0.39	0.75
Sugar (brown)	$0.17	0.70	0.65	0.82
Sugar (granulated)	$1.27	0.62	0.60	0.92
Sugar (powdered)	$0.13	0.72	0.69	0.89
Total (6)	$1.81	0.60	0.58	0.92
Other food & drink products (260)	$126.81	0.41	0.35	0.71
Remaining products (52)	$25.65	0.36	0.30	0.74

Notes: Total expenditure is 2008 expenditure for each category in all grocery, drug, and mass merchandise stores in the US, estimated from Nielsen Homescan and RMS data as described in section 2.4. Private label share (volume) is the share of equivalent quantity units (pills, ounces, gallons, etc.) in each category devoted to private labels in our Nielsen Homescan sample. Private label share ($) is the share of expenditure devoted to private labels in our Nielsen Homescan sample. Price ratio is the average price per equivalent quantity unit observed in the Nielsen RMS data for private labels divided by the analogous average price for brands. Rows for "headache remedies" and "pantry staples" each correspond to a single comparable product group. Rows for "other healthcare products," "other food products," and "non-food products" aggregate over multiple products groups, with the number of such groups shown in parentheses. In columns two through four, these aggregates average over comparable product groups weighting by expenditure.

Table 2: Knowledge and headache remedy purchases

Dependent variable: Purchase is a private label

Primary shopper characteristics	(1)	(2)	(3)	(4)
College education	0.0074	0.0254	0.0289	0.0230
	(0.0085)	(0.0087)	(0.0081)	(0.0076)
Share of active ingredients known	0.1929	0.1946	0.1974	0.1501
	(0.0129)	(0.0128)	(0.0118)	(0.0116)
Believe private labels are "just as safe"				0.2169
				(0.0078)
Demographic controls?	X	X	X	X
Product group & market fixed effects?	X	X	X	X
Income controls?		X	X	X
Market-chain fixed effects?			X	X
Sample	Second survey wave	Second survey wave	Second survey wave	Second survey wave
Mean of dependent variable	0.7020	0.7020	0.7020	0.7020
R^2	0.1335	0.1385	0.2036	0.2510
Number of households	20696	20696	20696	20696
Number of purchase occasions	145233	145233	145233	145233

Notes: Unit of observation is a purchase of a headache remedy by a household. Observations are weighted by volume (number of pills). Standard errors in parentheses are clustered by household. Independent variables apply to the "primary shopper" in each household, defined in section 2.2. Product type fixed effects are dummies for comparable product groups. Income controls are dummies for 16 household income categories. Demographic controls are dummies for categories of race, age, household composition, housing ownership, and geographic region.

Table 3: Occupation and headache remedy purchases

Dependent variable: Purchase is a private label

Primary shopper characteristics:	(1)	(2)	(3)	(4)	(5)
College education	0.0160	0.0320	0.0362		0.0847
	(0.0070)	(0.0072)	(0.0067)		(0.0140)
Pharmacist-physician	0.1566	0.1788	0.1753	0.1857	0.0746
	(0.0327)	(0.0323)	(0.0328)	(0.0393)	(0.1266)
Other health occupation	0.0912	0.0975	0.0929	0.0801	0.0796
	(0.0116)	(0.0115)	(0.0111)	(0.0189)	(0.0216)
Health major				0.0022	
				(0.0185)	
Non-health science major				0.0665	
				(0.0237)	
Demographic controls?	X	X	X	X	X
Product group & market fixed effects?	X	X	X	X	X
Income controls?		X	X	X	X
Market-chain fixed effects?			X	X	X
Sample	All	All	All	College major reported	Not currently employed
Mean of dependent variable	0.7095	0.7095	0.7095	0.7183	0.7058
R^2	0.1142	0.1183	0.1736	0.2283	0.2921
Number of households	32912	32912	32912	11136	5534
Number of purchase occasions	217429	217429	217429	68670	44809

Notes: Unit of observation is a purchase of a headache remedy by a household. Observations are weighted by volume (number of pills). Standard errors in parentheses are clustered by household. Independent variables apply to the "primary shopper" in each household, defined in section 2.2. Occupation is defined as of the primary shopper's most recent employment spell. "Pharmacist-physician" means occupation is either "pharmacist" or "physician or surgeon." "Other health occupation" means a health occupation other than "pharmacist" or "physician." "Health major" and "non-health science major" refer to primary shopper's reported college major. Product type fixed effects are dummies for comparable product groups. Income controls are dummies for 16 household income categories. Demographic controls are dummies for categories of race, age, household composition, housing ownership, and geographic region.

Table 4: Occupation and headache remedy purchases by well-informed consumers

Dependent variable: Purchase is a private label

Primary shopper characteristics:	(1)	(2)	(3)
College education	0.0347	0.0109	0.0086
	(0.0083)	(0.0133)	(0.0122)
Pharmacist-physician	0.1730	0.1392	0.0586
	(0.0382)	(0.0444)	(0.0394)
Other health occupation	0.0859	0.0569	0.0243
	(0.0140)	(0.0168)	(0.0161)
Sample	Second survey wave	Second survey wave	Second survey wave
Primary shopper survey response:			
Know all active ingredients		X	X
Believe private labels are "just as safe"			X
Mean of dependent variable	0.7020	0.7720	0.8492
R^2	0.1891	0.2838	0.3074
Number of households	20696	5250	3220
Number of purchase occasions	145233	38522	24189

Notes: Unit of observation is a purchase of a headache remedy by a household. Observations are weighted by volume (number of pills). Standard errors in parentheses are clustered by household. Independent variables apply to the "primary shopper" in each household, defined in section 2.2. Occupation variables are as in table 3. All specifications include demographic controls, income controls, product type dummies, and market-chain fixed effects as in column (3) of table 3. "Know all active ingredients" means the primary shopper correctly identified the active ingredient in all five headache remedies. "Believe private labels are 'just as safe"' means the primary shopper chose "agree" (1) on a 1-7 agree/disagree scale in response to the statement "Store brand products for headache remedies/pain relievers are just as safe as the brand name products."

Table 5: Headache remedy purchases under full information

	Total expenditure ($bn / year)	Expenditure on brands ($bn / year)	Brand quantity share
Baseline	$2.67	$1.48	0.291
All consumers as informed as a pharmacist / physician	$2.26 (0.07)	$0.74 (0.12)	0.131 (0.028)
Percent change	-0.154 (0.028)	-0.499 (0.084)	-0.548 (0.097)
Number of comparables: 6			
Price ratio (private label / brand): 0.41			

Notes: "Baseline" repeats summary information from table 1 for our six headache remedy comparable groups. Total expenditure and expenditure on brands are estimated 2008 totals for all grocery, drug, and mass merchandise stores in the US. The "all consumers as informed as a pharmacist / physician" counterfactual is based on separate regressions by comparable group using the specification of table 3 column (3). "Brand quantity share" in the counterfactual is computed by subtracting the quantity-weighted average physician-pharmacist coefficient from the observed brand quantity share for all non-pharmacist / physician shoppers in the sample. Counterfactual brand and total expenditure are computed from the counterfactual brand quantity share assuming brand and generic prices remain unchanged.

Table 6: Other health category purchases under full information

	Total expenditure ($bn / year)	Expenditure on brands ($bn / year)	Brand quantity share
Baseline	$6.52	$3.52	0.435
All consumers as informed as a pharmacist / physician	$6.18 (0.06)	$2.67 (0.15)	0.345 (0.018)
Percent change	-0.052 (0.010)	-0.240 (0.043)	-0.206 (0.042)
Number of comparables: 34			
Price ratio (private label / brand): 0.51			

Notes: Restricted to "other health" comparable groups with observed purchases by at least 5,000 households. "Baseline" presents summary information analogous to table 1. Total expenditure and expenditure on brands are estimated 2008 totals for all grocery, drug, and mass merchandise stores in the US. The "all consumers as informed as a pharmacist / physician" counterfactual is based on separate regressions by comparable group using the specification of table 3 column (3). "Brand quantity share" in the counterfactual is computed by subtracting the expenditure-weighted average physician-pharmacist coefficient from the observed brand quantity share for all non-pharmacist / physician shoppers in the sample. Counterfactual brand and total expenditure are computed from the counterfactual brand quantity share assuming brand and generic prices remain unchanged.

Table 7: Occupation and pantry staple purchases

Dependent variable: Purchase is a private label

Primary shopper characteristics:	(1)	(2)	(3)	(4)	(5)
College education	-0.0184	-0.0002	-0.0051		0.0178
	(0.0054)	(0.0055)	(0.0044)		(0.0093)
Chef	0.1697	0.1563	0.1296	0.2484	0.1827
	(0.0210)	(0.0216)	(0.0260)	(0.0727)	(0.0454)
Other food preparer	0.0481	0.0385	0.0305	0.0388	0.0430
	(0.0147)	(0.0139)	(0.0117)	(0.0204)	(0.0225)
Health major				0.0000	
				(0.0108)	
Non-health science major				0.0312	
				(0.0180)	
Demographic controls?	X	X	X	X	X
Product group & market fixed effects?	X	X	X	X	X
Income controls?		X	X	X	X
Market-chain fixed effects?			X	X	X
Sample	All	All	All	College major reported	Not currently employed
Mean of dependent variable	0.6052	0.6052	0.6052	0.5895	0.5945
R^2	0.0770	0.0815	0.2523	0.2627	0.2946
Number of households	37211	37211	37211	12567	6032
Number of purchase occasions	464369	464369	464369	144262	96757

Notes: Unit of observation is a purchase of a pantry staple by a household. Observations are weighted by volume (pounds). Standard errors in parentheses are clustered by household. Independent variables apply to the "primary shopper" in each household, defined in section 2.2. Occupation is defined as of the primary shopper's most recent employment spell. "Chef" means occupation is "chef or head cook." "Other food preparer" means a food preparation occupation other than "chef or head cook." "Health major" and "non-health science major" refer to primary shopper's reported college major. Product type fixed effects are dummies for comparable product groups. Income controls are dummies for 16 household income categories. Demographic controls are dummies for categories of race, age, household composition, housing ownership, and geographic region.

Table 8: Pantry staple purchases under full information

	Total expenditure ($bn / year)	Expenditure on brands ($bn / year)	Brand quantity share
Baseline	$1.81	$0.77	0.398
All consumers as informed as a chef	$1.79 (0.00)	$0.58 (0.04)	0.301 (0.020)
Percent change	-0.011 (0.002)	-0.243 (0.050)	-0.244 (0.050)
Number of comparables: 6			
Price ratio (private label / brand): 0.92			

Notes: "Baseline" repeats summary information from table 1 for our six pantry staple comparable groups. The "all consumers as informed as a chef" counterfactual is based on separate regressions by comparable group using the specification of table 7 column (3). Total expenditure and expenditure on brands are estimated 2008 totals for all grocery, drug, and mass merchandise stores in the US. "Brand quantity share" in the counterfactual is computed by subtracting the expenditure-weighted average chef coefficient from the observed brand quantity share for all non-chef shoppers in the sample. Counterfactual brand and total expenditure are computed from the counterfactual brand quantity share assuming brand and generic prices remain unchanged.

Table 9: Other food and drink purchases under full information

	Total expenditure ($bn / year)	Expenditure on brands ($bn / year)	Brand quantity share
Baseline	$122.30	$77.72	0.576
All consumers as informed as a chef	$121.96 (0.30)	$75.89 (0.88)	0.563 (0.006)
Percent change	-0.003 (0.002)	-0.023 (0.011)	-0.022 (0.011)
Number of comparables: 243			
Price ratio (private label / brand): 0.71			

Notes: Restricted to "other food and drink" comparable groups with observed purchases by at least 5,000 households. "Baseline" presents summary information analogous to table 1. Total expenditure and expenditure on brands are estimated 2008 totals for all grocery, drug, and mass merchandise stores in the US. The "all consumers as informed as a chef" counterfactual is based on separate regressions by comparable group using the specification of table 7 column (3). "Brand quantity share" in the counterfactual is computed by subtracting the expenditure-weighted average chef coefficient from the observed brand quantity share for all non-chef shoppers in the sample. Counterfactual brand and total expenditure are computed from the counterfactual brand quantity share assuming brand and generic prices remain unchanged.

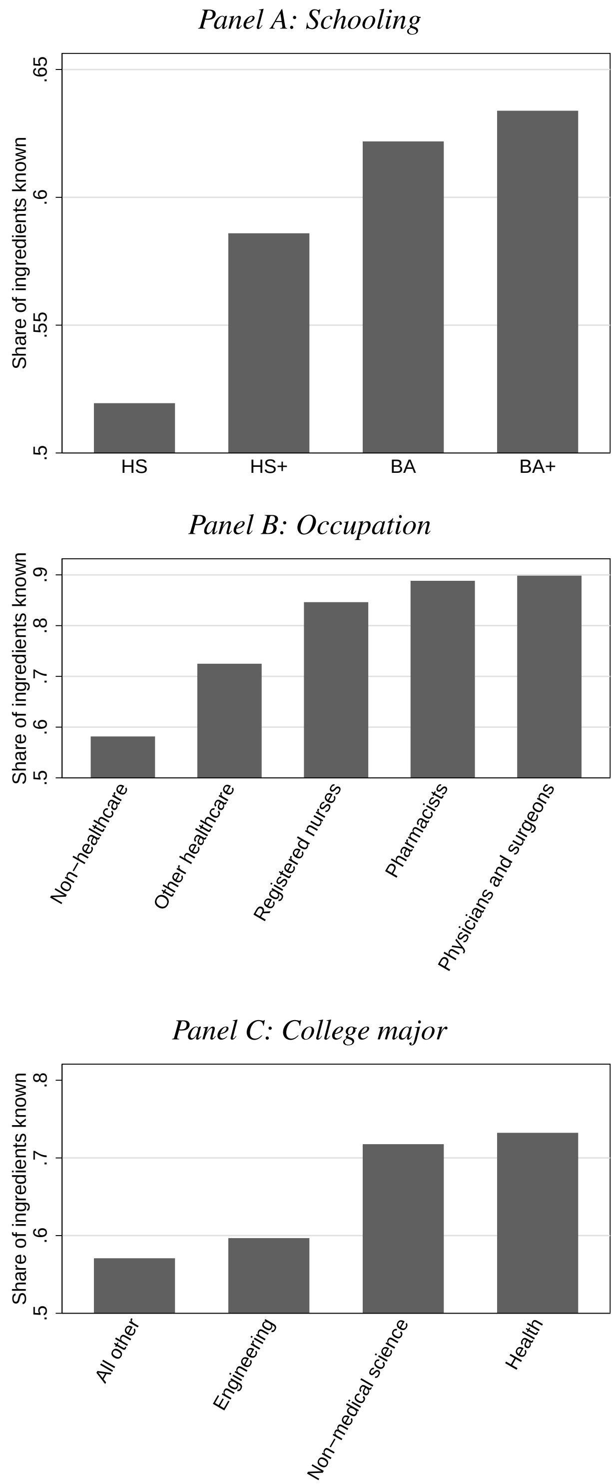
Table 10: Evidence on domain specificity

Dependent variable: Purchase is a private label

Primary shopper characteristics:	(1)	(2)	(3)
College education	-0.0045	-0.0063	0.0452
	(0.0054)	(0.0044)	(0.0066)
Share of active ingredients known	-0.0023		
	(0.0075)		
Pharmacist-physician		0.0170	
		(0.0275)	
Other health occupation		0.0043	
		(0.0094)	
Chef			0.0994
			(0.0429)
Other food preparer			-0.0027
			(0.0182)
Products	Pantry Staples	Pantry Staples	Headache Remedies
Mean of dependent variable	0.6044	0.6052	0.7095
R^2	0.2590	0.2520	0.1704
Number of households	23226	37211	32912
Number of purchase occasions	306090	464369	217429

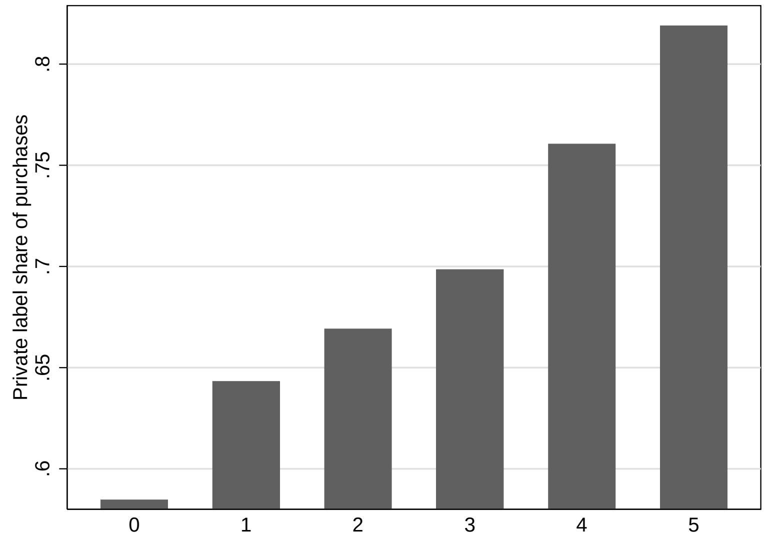
Notes: Unit of observation is a purchase of a pantry staple (first two columns) or headache remedy (third column) by a household. Observations are weighted by volume (pounds or number of pills). Standard errors in parentheses are clustered by household. Independent variables apply to the "primary shopper" in each household, defined as described in section 2.2. Occupation variables are as in tables 3 and 7. All specifications include demographic controls, income controls, product type dummies, and market-chain fixed effects as in column (3) of tables 3 and 7.

Figure 1: Product knowledge, headache remedies



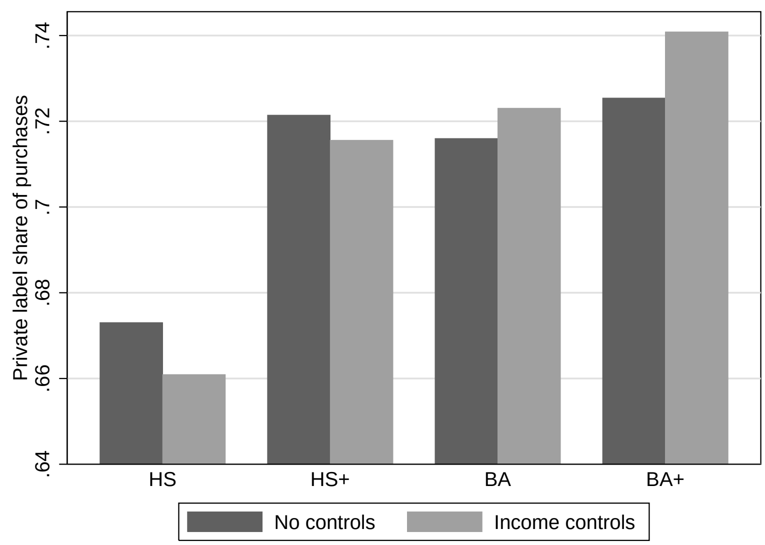
Note: Figure shows the mean share of six headache remedy active ingredients correctly identified by each group of respondents in 2011 PanelViews survey, among those who answered all six questions.

Figure 2: Private label purchases and knowledge, headache remedies



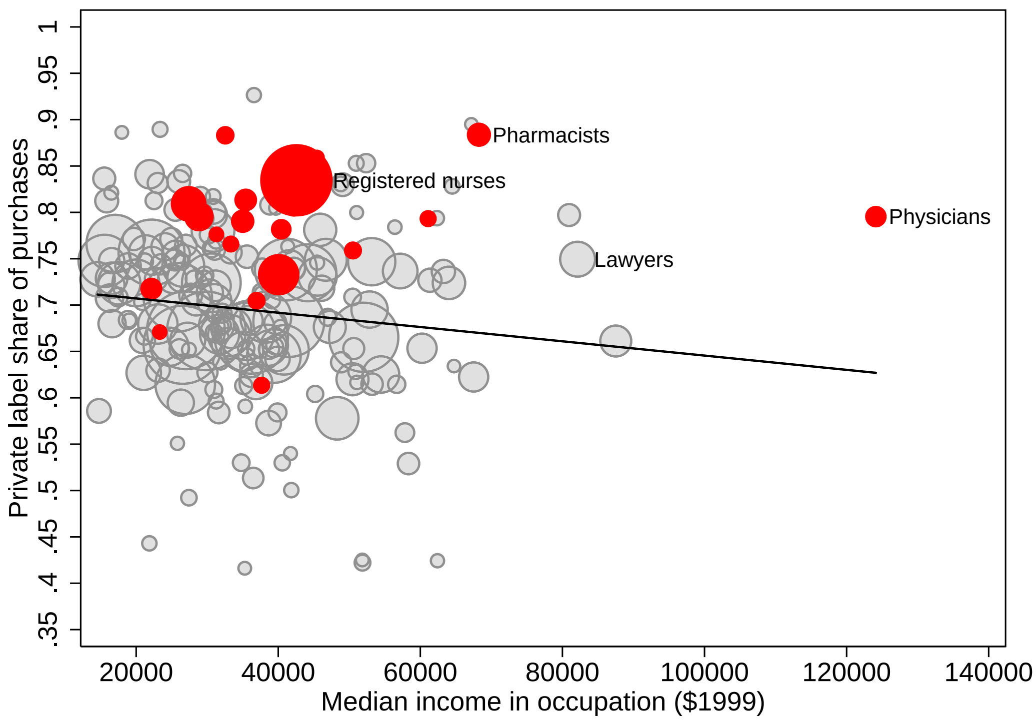
Notes: Horizontal axis gives the number of headache remedy active ingredients correctly identified in 2011 PanelViews survey. The bars show the volume-weighted private label share of headache remedies for households in each category.

Figure 3: Private label purchases and education, headache remedies



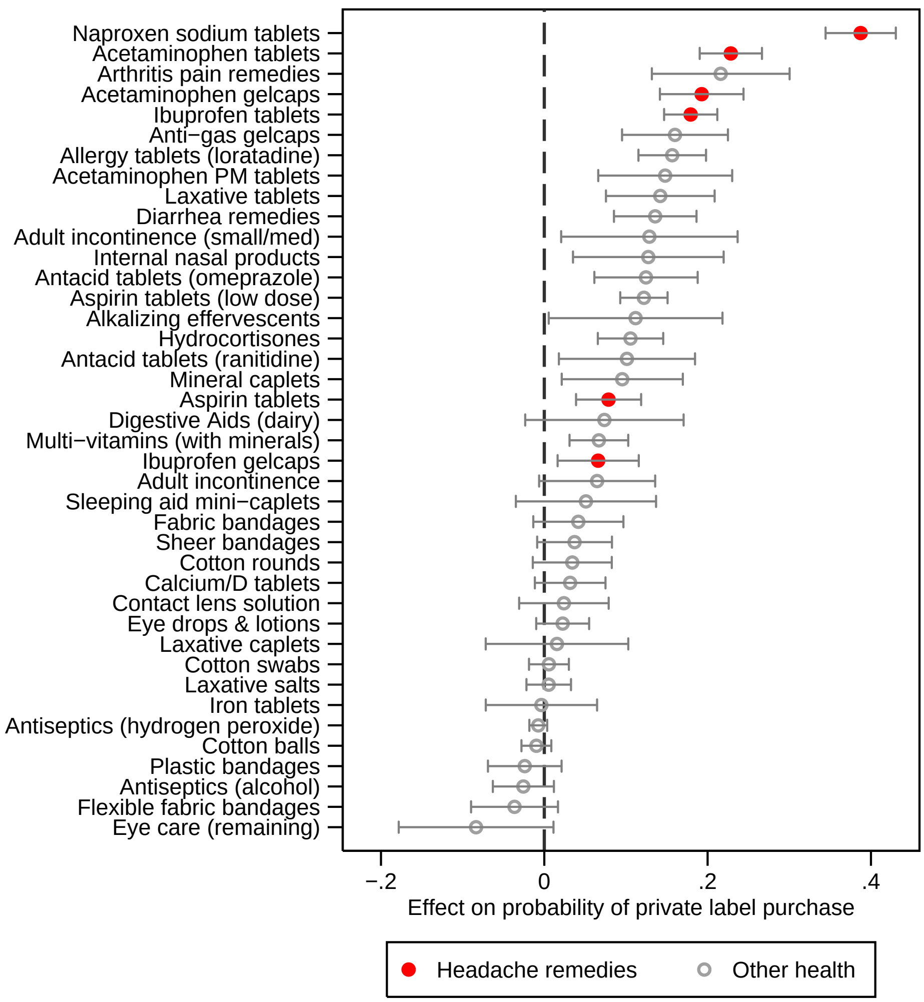
Notes: Dark bars shows the volume-weighted private label share of headache remedies for households in each education category. Light bars show the predicted volume-weighted private label share in each education category from a regression on dummies for 16 household income categories, with the predicted values computed at the means of the covariates.

Figure 4: Private label purchases and occupation, headache remedies



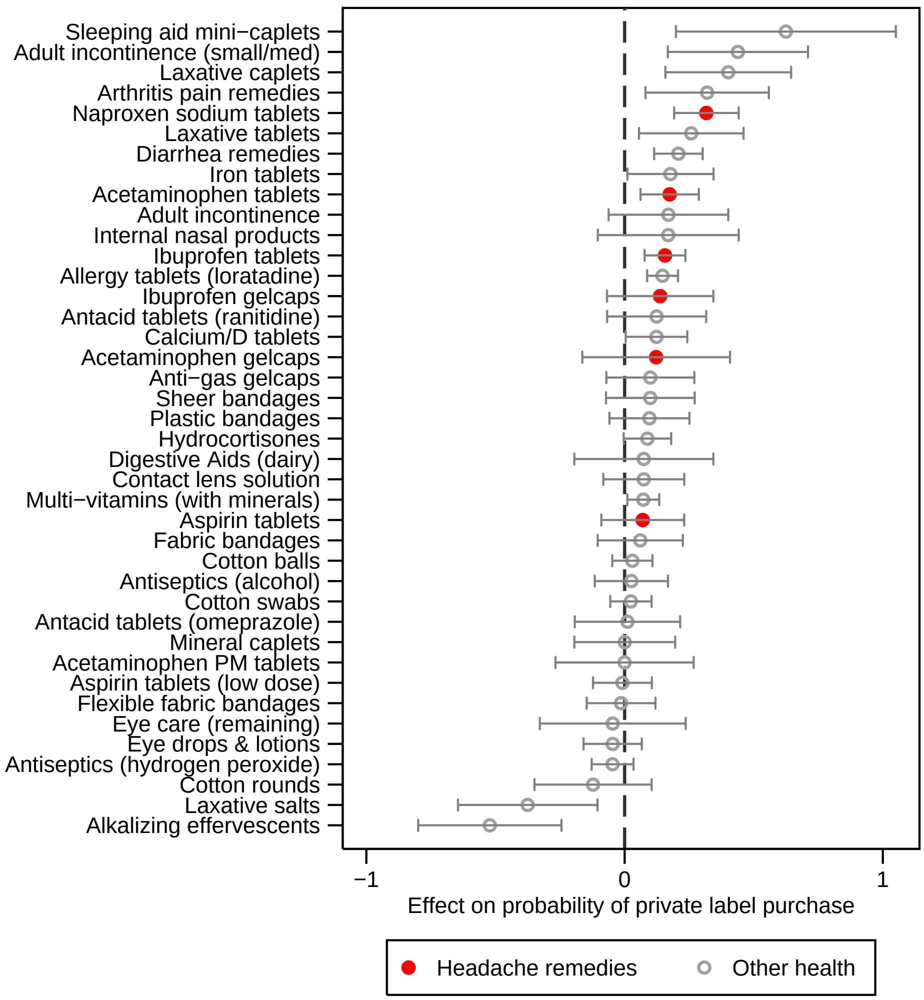
Notes: Figure shows volume-weighted private label share of headache remedy purchases by occupation (y-axis) and median earnings for full-time full-year workers in 1999 by occupation (x-axis). Filled (colored) circles represent healthcare occupations. The area of each circle is proportional to the number of households whose primary shopper has the given occupation in our sample, with different scale for healthcare and non-healthcare occupations. Occupations with fewer than 25 such households are excluded from the figure.

Figure 5: Active ingredient knowledge coefficients



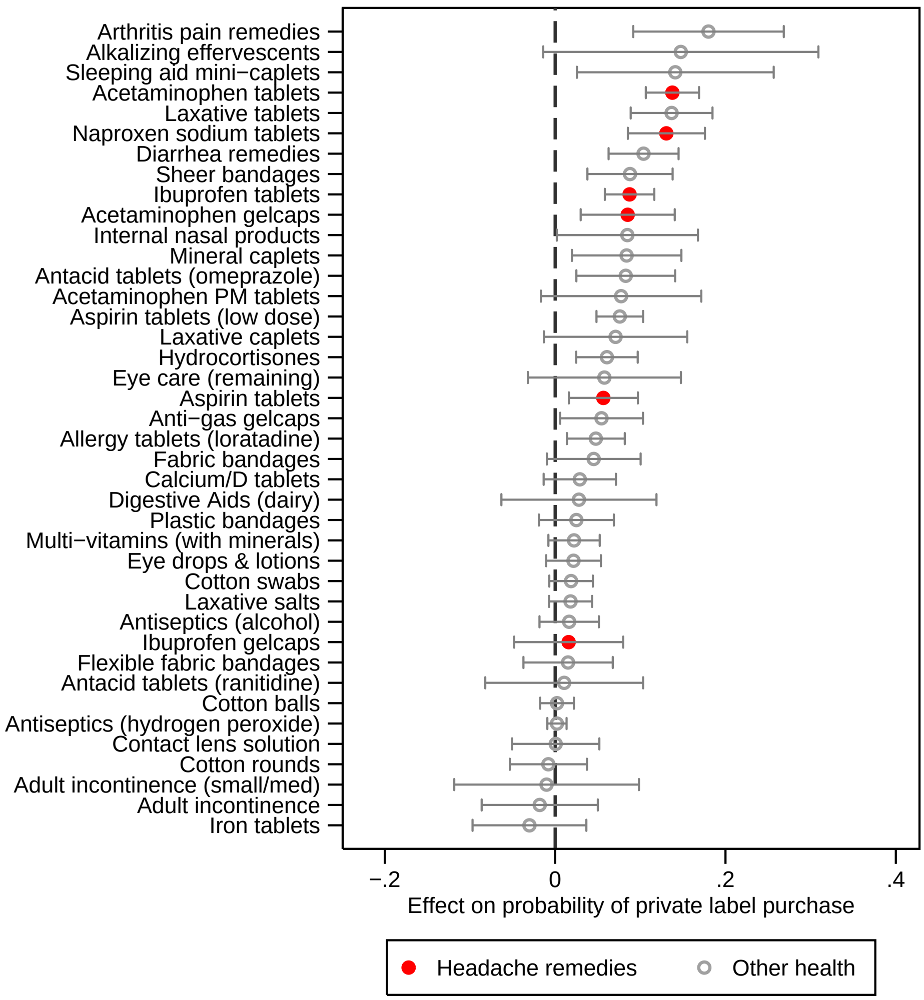
Notes: Figure plots coefficients and standard errors on "share of active ingredients known" for each health-related comparable product group in our sample from a regression following the specification of table 2 column (3).

Figure 6: Pharmacist / physician occupation coefficients



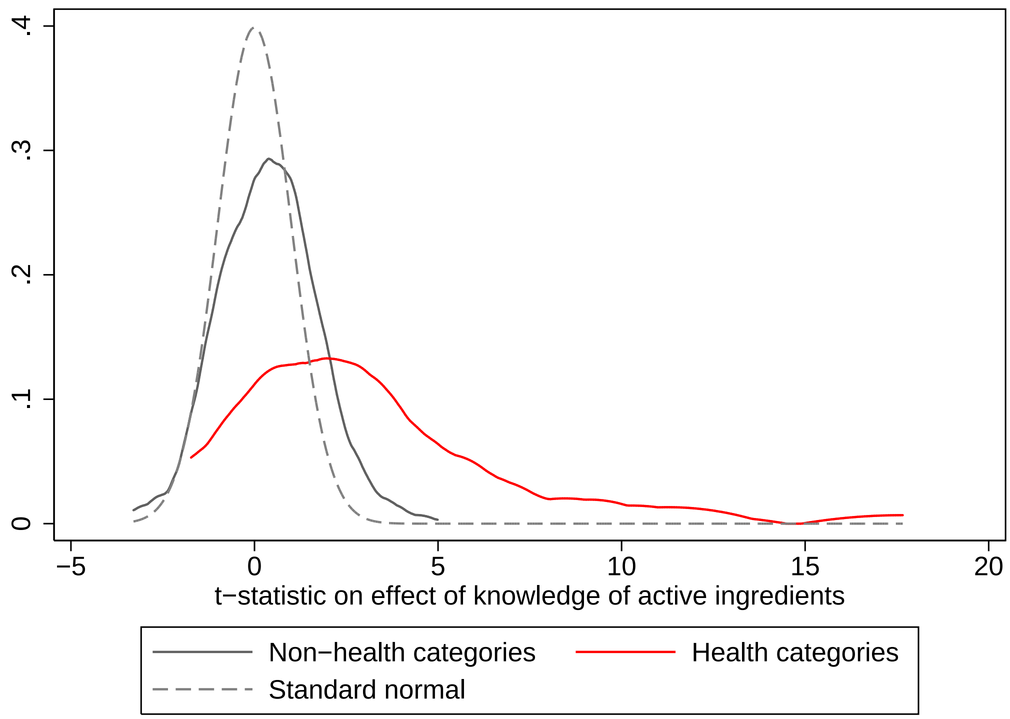
Notes: Figure plots coefficients and standard errors on “pharmacist-physician” for each health-related comparable product group in our sample from a regression following the specification of table 3 column (3).

Figure 7: Other healthcare occupation coefficients



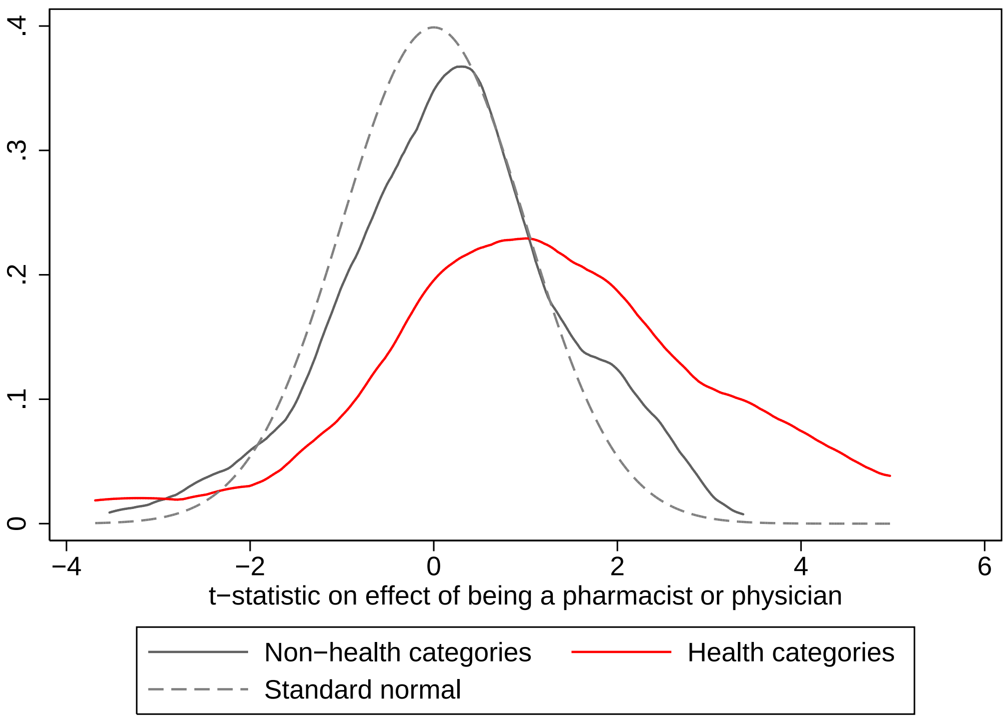
Notes: Figure plots coefficients and standard errors on "other health occupation" for each health-related comparable product group in our sample from a regression following the specification of table 3 column (3).

Figure 8: Active ingredient knowledge coefficients, health vs. non-health products



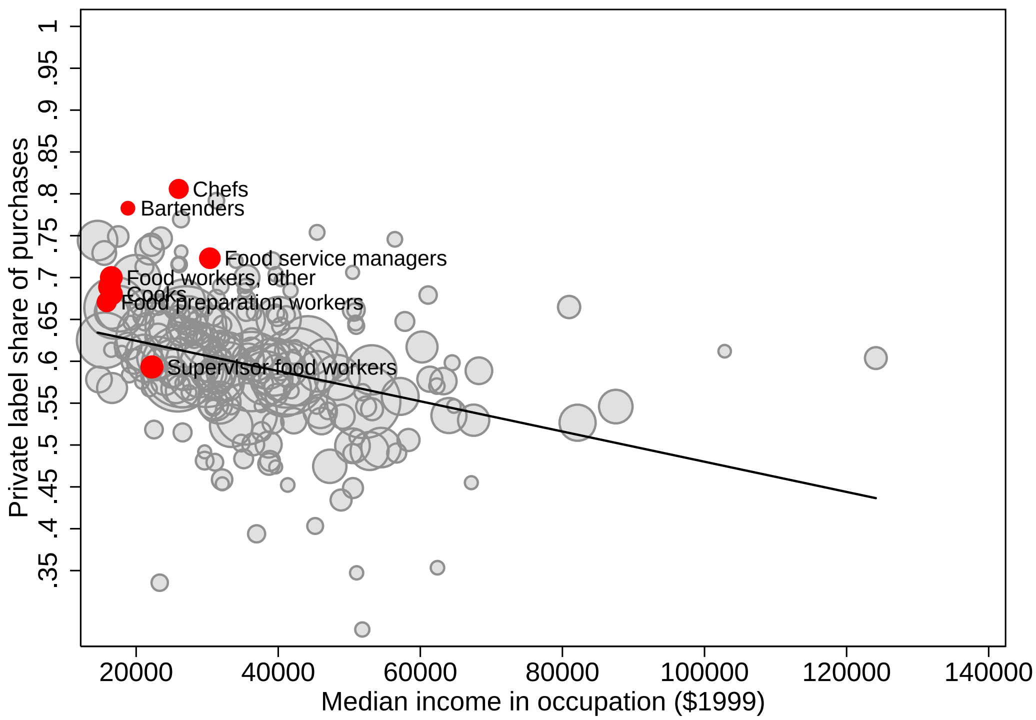
Notes: Figure plots the distribution of t-statistics on "share of active ingredients known" for all health-related and non-health-related comparable products groups in our sample from a regression following the specification of table 2 column (3). Distribution is estimated using an Epanechnikov kernel with optimal bandwidth. The standard normal density ("null hypothesis") is plotted with dashed lines.

Figure 9: Pharmacist / physician coefficients, health vs. non-health products



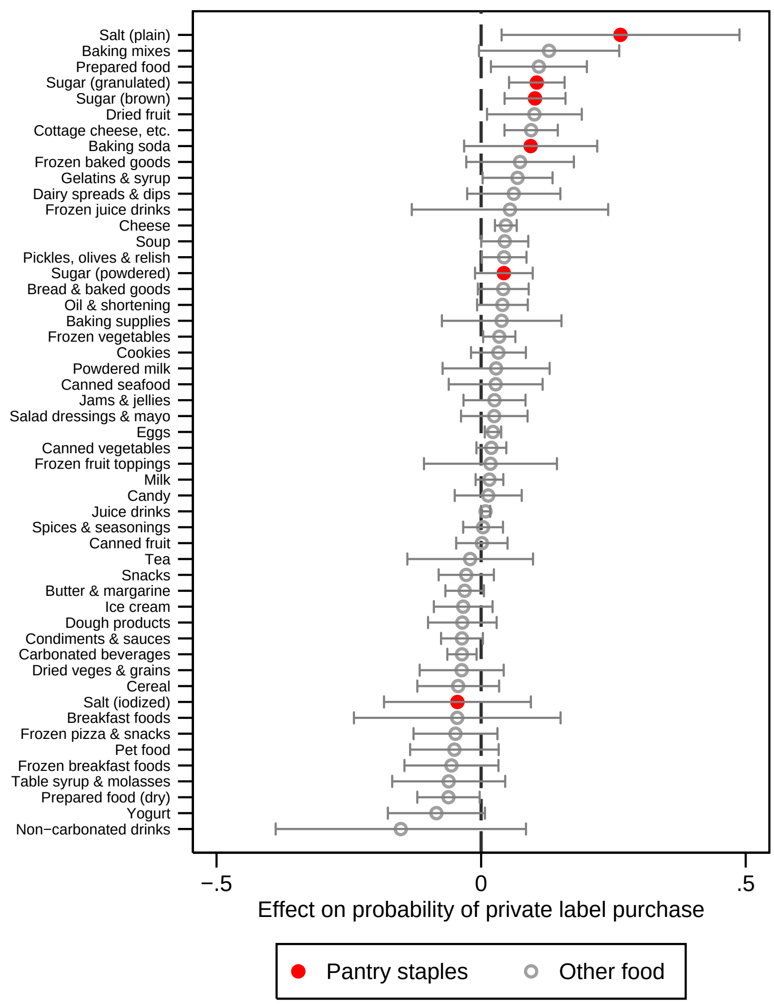
Notes: Figure plots the distribution of t-statistics on "pharmacist-physician" for all health-related and non-health-related comparable products groups in our sample from a regression following the specification of table 3 column (3). Distribution is estimated using an Epanechnikov kernel with optimal bandwidth. The standard normal density ("null hypothesis") is plotted with dashed lines.

Figure 10: Private label purchases and occupation, pantry staples



Notes: Figure shows volume-weighted private label share of pantry staple purchases by occupation (y-axis) and median earnings for full-time full-year workers in 1999 by occupation (x-axis). Filled (colored) circles represent food preparer occupations. The area of each circle is proportional to the number of households whose primary shopper has the given occupation in our sample, with different scale for food preparer and non-food-preparer occupations. Occupations with fewer than 25 such households are excluded from the figure.

Figure 11: Chef coefficients



Notes: Figure plots coefficients and standard errors on "chef" for each food and drink category in our sample from a regression following the specification of table 7 column (3). Coefficients for pantry staples are plotted individually by comparable product groups. Coefficients for other categories are aggregated to the level of Nielsen "product groups," which may include multiple comparable product groups, weighting coefficients by precision and averaging standard errors assuming independence across comparable product groups.

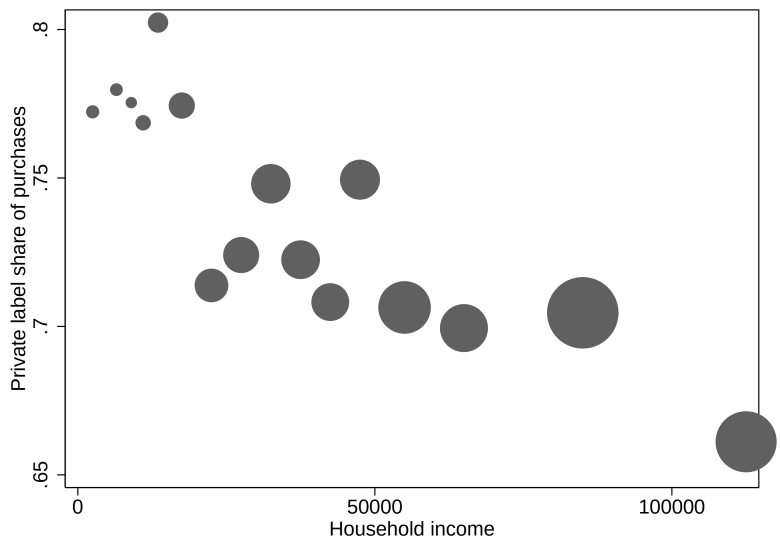
Appendix Table 1: Knowledge and headache remedy purchases, robustness

Dependent variable: Purchase is a private label

	Headache remedies		Pantry staples
	Share of active ingredients	Pharmacist-Physician	Chef
Specification	coefficient	coefficient	coefficient
Baseline	0.1974	0.1753	0.1296
	(0.0118)	(0.0328)	(0.0260)
Control for market-chain-week	0.2162	0.2080	0.1158
	(0.0176)	(0.0493)	(0.0240)
Control for total volume purchased	0.1947	0.1690	0.1341
	(0.0117)	(0.0318)	(0.0257)

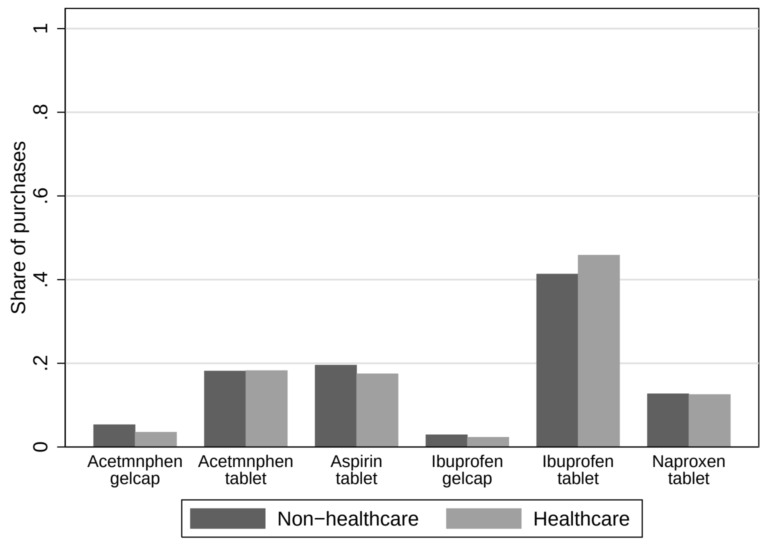
Note: Each row gives (i) the coefficient on "share of active ingredients known" from a specification analogous to table 2 column (3); (ii) the coefficient on "pharmacist-physician" from a specification analogous to table 3 column (3); and (iii) the coefficient on "chef" from a specification analogous to table 7 column (3). The first row repeats the results from our main specifications. The second row is the same as the baseline but replaces market-chain fixed effects with market-chain-week fixed effects. The third row is the same as the baseline but adds a control for the total volume of headache remedies purchased by the household.

Appendix Figure 1: Private label purchases and household income, headache remedies



Note: Figure shows the volume-weighted private label share of headache remedies for households in each income category. Household income is imputed at the midpoint of the range for each category, with the top category imputed at 120,000. The area of each circle is proportional to the number of households in the income category in our sample.

Appendix Figure 2: Physical attribute choice and occupation, headache remedies



Notes: Probability of purchase is computed from a set of linear probability models of the likelihood of purchasing the given product. Bars labeled “healthcare” show the predicted probability from the given model for purchases made by households whose primary shopper is in a healthcare occupation. Bars labeled “not healthcare” show the predicted probability for the same purchases under the counterfactual in which the household’s primary shopper is not in a healthcare occupation. Each linear probability model’s unit of observation is the purchase occasion. Observations are weighted by volume (in number of pills). All specifications include a dummy for college completion, income controls, demographic controls, and market-chain fixed effects. Income controls are dummies for 16 household income categories. Demographic controls are dummies for categories of race, age, household composition, housing ownership, and geographic region. Predicted probabilities set the market-chain fixed effect so that the mean predicted probability is equal to the empirical share. See the online appendix for a supporting table with additional details.