

# UC Berkeley

## CUDARE Working Papers

### Title

Generic Aversion and Observational Learning in the Over-the-Counter Drug Market

### Permalink

<https://escholarship.org/uc/item/7ks7s9jf>

### Authors

Carrera, Mariana  
Villas-Boas, Sofia B

### Publication Date

2020-12-23

# Generic Aversion and Observational Learning in the Over-the-Counter Drug Market<sup>1</sup>

Mariana Carrera<sup>2</sup>

Montana State University

Sofia Villas-Boas<sup>3</sup>

University of California- Berkeley

Dec. 16, 2020

## ABSTRACT

Through a labeling intervention at a national retailer, we test three hypotheses for consumer aversion to generic over the counter (OTC) drugs: lack of information on the comparability of generic and brand drugs, inattention to their price differences, and uncertainty about generic quality that can be reduced with information on peer purchase rates. With a difference-in-differences strategy, we find that posted information on the purchases of other customers increases generic purchase shares significantly, while the other treatments do not. Consumers without prior generic purchases appear particularly responsive to this information. These findings have policy implications for promoting evidence-based, cost-effective choices.

---

<sup>1</sup>We thank Gordon Dahl, Stefano DellaVigna, Ulrike Malmendier, Gautam Rao, Mark Votruba, and seminar participants at U.C. Berkeley, Case Western, Kent State, Vanderbilt, and Duke University for helpful comments. We thank a national retailer, whom we cannot name, for sharing data and allowing us to conduct a field experiment. We thank the Giannini Foundation for support.

<sup>2</sup>Department of Agricultural Economics and Economics (DAEE), Montana State University | P.O. Box 172920 | Bozeman, MT 59717-2920. [mariana.carrera@montana.edu](mailto:mariana.carrera@montana.edu)

<sup>3</sup>Department of Agricultural and Resource Economics, University of California, Berkeley, Berkeley, CA 94720-3310. [sberto@berkeley.edu](mailto:sberto@berkeley.edu)

## **1. Introduction**

Consumers' choices are influenced by several non-standard factors including the salience of prices and other product attributes (Chetty, Looney, and Kroft, 2009; Bordalo, Gennaioli and Schleifer, 2013), the difficulty of comparing attributes across alternatives (Allcott, 2013; Hossain and Morgan, 2006), and potentially biased beliefs (Bollinger, Leslie and Sorensen, 2010). Beliefs about efficacy and risks are acutely important when consumers choose experience goods or credence goods, which encompass most health care treatments. Mistaken beliefs can lead to treatment overuse (e.g., antibiotics) or underutilization (e.g., chronic preventative drugs) (Baicker, Mullanaithan and Schwarstein, 2015). The difficulty of making price comparisons in the health care sector may also drive the overuse of costly treatments in place of lower-cost substitutes (Carrera et al., 2018).

Despite mounting evidence of such "behavioral hazards" affecting consumers' medical decisions, little is known about how to address them. Provision of information has been shown to increase the take-up of some preventative health products (Delavande, 2008), but can also backfire, as in the case of vaccination promotion (Nyhan et al. 2014; 2015) and information about product safety (Ma, Wang and Khanna, 2017). A better understanding of how health product preferences are formed and updated is needed to facilitate evidence-based and cost-effective choices.

In this paper, we test how three different types of information, posted at the point of sale, affect consumers' choices between brand and generic over-the-counter (OTC) drugs. The OTC drug market is a ripe setting for studying how consumers form preferences over pharmaceutical products. In contrast to prescription drugs, which physicians select for their patients, consumers choose OTC drugs autonomously. To facilitate the comparison of different products, the FDA requires standardized "Drug Facts" labels to be posted on every package, and visible price tags make price comparisons much more straightforward than in the

prescription drug market. Nevertheless, more than half of the sales of familiar household drugs are for branded versions, which cost 40-60% more than their generic equivalents without any treatment or safety advantage. Bronnenberg et al. (2015) argue that the high market shares of branded OTC products reflect consumers' low awareness of the existence and comparability of a generic substitute. Consumers might also perceive a greater degree of uncertainty regarding the safety or quality of the generic product, leading ambiguity-averse individuals to prefer the brand (Muthukrishnan, Wathieu and Xu, 2009). In addition, cognitive effort is required to locate and compute the savings offered by the generic versions, which often have different product names at different retailers.

We conducted a set of temporary labeling interventions at six locations of one national retailer, to test three hypotheses for consumer aversion to generic OTC drugs: (1) lack of awareness of the equivalence of generic drugs to their brand counterparts, (2) inattention to the price difference between brand and generic drugs, and (3) significant uncertainty regarding product quality which might be reduced by information on other customers' purchases. Perceived quality encompasses efficacy as well as other attributes such as safety or taste.

For each hypothesis, we designed product-specific information labels to display relevant facts to consumers. Labels used to test the first hypothesis displayed information on the FDA approval, bioequivalence, and active ingredients of the generic matching the brand. Labels testing the second hypothesis highlighted the brand-generic price difference in percentage terms. To test the third hypothesis, we displayed brand or generic purchase rates for each product, calculated using pre-intervention sales data specific to each product in each store. We also introduced exogenous variation in the posted share (within store and product) by alternating weekly the length of pre-period time used to calculate the posted purchase rates. For two of the tests, we designed labels with two different ways of framing the relevant information in different stores.

We randomly assigned a fixed set of OTC categories to be “treated” with a label attached to the shelf price tag, and each of six treatment stores was randomly assigned to one of five label types (three different types of information, two of which had two framing variations). Using OTC sales data from previous years, we identified six similar stores to use as controls, and use a difference-in-differences approach to measure how consumers respond to the different labels. We estimate treatment effects on our main outcome of interest, the generic share of purchases, as well as the total quantity purchased of brand and generic versions combined. Using household-level sales data, we explore heterogeneity based on customers’ past purchase patterns and test for post-treatment persistence of changes in purchasing behavior.

Our results are as follows. First, information on the comparability of generics has no effect on their purchase rate relative to the national brand. This null effect is surprising, but serves to alleviate the concern that our labels might increase generic purchase rates simply through a salience effect, i.e. by drawing customers’ attention to the presence of a labeled generic product or reducing the time cost of identifying it. Second, we find mixed evidence for the hypothesis of inattention to the price difference between brand and generic drugs. The average generic purchase share of treated products increased when the price difference between brand and generic versions was highlighted as a “savings” relative to the brand, but not when this difference was framed as an additional cost to choosing the brand. Third, we find strong evidence that consumers respond to information on the purchases of other consumers. This is consistent with general uncertainty about product desirability, perhaps encompassing attributes beyond stated equivalence of active ingredients. We find that the generic share of purchases increases by 8.5 percentage points, or 20% relative to the pre-treatment level. This effect is more than 60% as large as the increase associated with a price promotion on the generic store brand, and is particularly strong among brand-loyal customers.

A large theoretical literature examines the ways in which uncertainty regarding product quality affects consumer demand (see Akerlof, 1970; Nelson, 1970; Wiggins and Lane, 1983; and Wolinsky, 1995). A closely related empirical literature analyzes how product quality information affects consumer behavior through branding (Montgomery and Wernerfelt, 1992), mandatory product labeling (Jin and Leslie, 2003), and advertising (Akerberg, 2001; Akerberg, 2003). A smaller but growing literature finds evidence of consumer inattention to non-salient components of costs, such as taxes (Chetty, Looney and Kroft, 2009) and eBay shipping costs (Hossain and Morgan, 2006). We assume that all consumers observe the price of the brand-name product and hypothesize that they may be inattentive to, and may tend to underestimate, the savings offered by the generic product.

Our third test adds to a growing literature on observational learning.<sup>4</sup> There are three main channels through which an individual's demand for a good might be affected by information on other consumers' purchases: (1) goods that are socially visible may confer status, (2) there may be network benefits or "social utility" to using the same product as a peer, and (3) others' usage could serve as signals of a good's quality.<sup>5</sup> While other experimental studies have found positive effects of peer usage disclosure, it is typically difficult to disentangle whether these effects are driven by updated quality priors as opposed to the social channels (status,

---

<sup>4</sup> On the theoretical side, Becker (1991) developed a formal model in which the demand for a good depends positively on its aggregate quantity demanded, i.e. on peer demand. McFadden and Train (1996) formalize consumer learning about a new good's quality through a rational, forward-looking decision process in which they learn from their own experience and from the experiences of their peers. Moretti (2011) derives predictions from a model of social learning for the time path of movie sales after their release. He analyzes how sales diverge over time for movies that perform well or poorly relative to prior expectations and finds a reinforcing pattern: When a movie exceeds expectations in its opening week, consumers update their expectations via input from their peers, leading to greater future sales.

<sup>5</sup> A fourth channel is the behavioral "nudge" that can come from identifying a particular behavior as an injunctive or descriptive norm. This channel is less relevant here for two reasons. First, purchasing generic OTC products is not really considered to be pro-social or an action for the collective good. Second, Cialdini (1984) and the related literature typically find that high existing shares are necessary for these messages to be effective, which is not the case in our context.

network effects).<sup>6</sup> For example, Cai et al. (2009) found that restaurant demand for menu items increases significantly when they are identified as “Top 5” most popular items. But at a restaurant where diners eat “family-style,” the customer might seek to order items that others are likely to enjoy (instead of using the information to update his own preferences for the item). Similarly, in an artificial music market where participants select unknown songs to download, Salganik et al. (2006) found that their choices were influenced by information on each song’s current download count. However, this might be driven by a socially-motivated desire to know the songs that others play.

Since the consumption of pharmaceutical products is a largely private behavior, social status and network effects are arguably absent in our context. Thus, the third channel—learning about quality—is likely to be the only one through which the purchases of other customers would influence one’s purchasing decision. Since most consumers do not buy over-the-counter drugs often, and seldom observe their peers’ purchases, even simple statistics of generic purchase rates may contribute new and useful information, just as physicians may learn about prescription drug quality through the choices of other physicians (Saxell, 2014). However, it is not theoretically clear whether disclosing market shares to customers will lead to an increase in generic share, given that they average 40-50% in the stores that we study. We use in-store customer surveys to complement our field experiment by soliciting priors on the purchase shares of other consumers and assessing self-reported likelihood of being influenced by peer market shares. Findings suggest that brand-buying customers are more likely to be swayed towards the generic by peer-purchase information than generic-buying customers are to be swayed towards the brand, most likely because brand-buyers are less likely to have personal experience with both the brand and generic.

---

<sup>6</sup> An exception is Bursztyn et al. (2013), which separately estimates the influences of “social utility” versus “social learning” (updating quality expectations) in the case of paired peers purchasing financial assets.

The rest of the paper proceeds as follows. Section 2 describes the background on over the counter drug regulations and a simple conceptual framework that embeds our three testable hypotheses. Section 3 describes the retail empirical setting, the experimental design and intervention, and the data used. Section 4 presents the empirical strategy, results and robustness checks, while section 5 concludes.

## **2. Background**

### *2.1 OTC Drugs in the United States*

Generic versions of over-the-counter drugs contain the same active ingredients as their name brand counterparts and are highly regulated. For newer drugs, each manufacturer who wishes to produce a generic version of the drug must obtain their own FDA approval prior to selling it. In the prescription drug market, the FDA tests generics for bioequivalence to the brand, defined as a similar time pattern of active ingredient release and absorption into the blood stream. Of the drugs in our sample, many, but not all, were tested for bioequivalence because they were sold in the prescription market prior to the OTC market. For older drugs, including, for example, acetaminophen (Tylenol) and diphenhydramine (Benadryl), the FDA publishes a “Monograph” specifying regulations for production, packaging, and labeling, but does not actively examine and approve the formulations sold by each manufacturer.

The FDA and clinical studies fail to find differences in safety or efficacy between versions of a drug produced by the original brand patent holder and generic entrants (see Kesselheim et al. 2008).<sup>7</sup> Despite this, the perceptions of consumers and, to a lesser degree, physicians, are that generic drugs are less desirable than the original brand product (Shrank et al. 2011, Shrank et al. 2009). Interestingly, Bronnenberg et al. (2015) find that pharmacists are far more likely to purchase

---

<sup>7</sup> An exception is drugs that have a “narrow therapeutic index (NTI)” meaning that patient response can be sensitive to very small differences in the timing and speed of ingredient absorption. No over-the-counter drugs are considered NTI.



generic over-the-counter drugs than the general population, implying that the generic's drug quality, relative to the brand, is higher than perceived by the average consumer.

## 2.2 Conceptual Framework

Drugs treating symptomatic conditions may be experience goods, meaning individual  $i$  only learns her utility levels  $v_{ib}$  (for the brand version of a given product) and  $v_{igx}$  (for the generic version of that product) after having tried each of them. We assume that since brands precede generics on the market, all buyers of product  $x$  have used the brand version in the past and thus know their private valuation  $v_{ibx}$ .<sup>8</sup> Those who have not tried the generic version do not know their private valuation  $v_{ig}$ , but have an expectation over it,  $E[v_{igx}]$ . We do not assume this expectation is unbiased, since lack of information, prior experience in foreign markets with unregulated or unreliable generics, and differential advertising could create bias. A risk-neutral shopper who has not yet tried the generic will continue to purchase the brand version if:

$$(1) v_{ibx} - E[v_{igx}] > -\alpha_i (p_{bx} - p_{gx})$$

where  $\alpha_i$  represents sensitivity to price,  $p_{bx}$  and  $p_{gx}$  are the prices of the brand and generic, respectively<sup>9</sup>.

Each of our interventions aims to shift either  $E[v_{igx}]$  or  $\alpha_i$ . By displaying information on the active ingredient comparability of brand and generic drugs, we test whether  $E[v_{igx}]$  is based on inaccurate knowledge of this. By displaying the typical price difference in percentage terms, we test whether increasing the salience

---

<sup>8</sup> In consumer surveys we found that 94% of customers reported having purchased the brand version of their preferred headache remedy at least once in the past, whereas only 65% had ever purchased any generic version of it (i.e. not restricted to the specific store-brand of the retailer we studied).

<sup>9</sup> Although we do not explicitly model risk aversion, it is easy to see that if  $v_{ibx}$  is known and  $v_{igx}$  is unknown, then as risk aversion increases, an individual becomes less likely to choose the generic and may continue to purchase the brand, even if the inequality in (1) goes the other direction.

of  $p_{bx} - p_{gx}$  increases the consumer's response to the price difference ( $\alpha_i$ ).<sup>10</sup> By displaying the share of other customers who buy the generic or brand version, we test whether  $E[v_{igx}]$  can be influenced by observational learning.

Observational learning is the process by which consumers update their expectations of the quality of a given good after observing others' decision to purchase it (or not). In our context of choice between two competing versions of a product, observing the share of other customers who buy the generic serves as a signal of the generic's popularity relative to the more expensive brand version. To the extent that this signal exceeds (or falls short of) a customer's prior guess about the share of shoppers who buy the generic, the customer might increase (or decrease) her estimate of the generic's value.

Thus, the effect of posted generic shares could be different on customers who typically buy the brand versus those who typically buy the generic for two reasons. First, these types of customers may have different priors about generic drugs' popularity, and second, those who have never before purchased the generic might have considerably weaker priors about its quality.

To illustrate why customers' differing priors are important, consider a hypothetical case in which the generic purchase share is 50% and, consistent with this share, the customers who see the labels are an even split of would-be generic buyers and would-be brand buyers. We would only expect observational learning to influence the decision of customers who are surprised to learn that their prospective choice is *less* popular than they thought. If would-be brand buyers and would-be generic buyers all underestimate the popularity of generics, then observational learning implies an unambiguously positive effect of labels on generic purchase share. But if, instead, all customers overestimate the share of customers who make the same choice as them, then observational learning could

---

<sup>10</sup> Highlighting the "savings" associated with buying the generic might also increase generic purchases by increasing the perceived deal-value of the product (described as transaction utility by Thaler (1985).

make would-be generic buyers less likely to choose the generic just as it makes would-be brand buyers less likely to choose the brand. If generic and brand buyers are symmetrically biased in their priors, and equally responsive to the true peer purchase share, then the labels might not shift the aggregate generic purchase rate at all.

It is also possible that, regardless of their (possibly biased) priors on peer purchase shares, generic and brand buyers put different weights on this peer purchase shares when updating their expectations of the generic's private utility to them ( $E[v_{igx}]$ ) relative to the brand ( $v_{ibx}$ ). If they are pure experience goods with fixed private utilities, the purchases of other people should be irrelevant to the next choice of all consumers who have previously used both the brand and generic versions of a drug, and thus, know their private values of  $v_{igx}$  and  $v_{ibx}$ . Of course, if the drugs are not perfect experience goods, consumers may still perceive  $v_{igx}$  and  $v_{ibx}$  as uncertain; for example, there may be some risk that the manufacturer has produced a bad batch, or that long-term adverse effects have not been realized. To the extent that such uncertainty exists, the perceived utility from purchasing a drug might be sensitive to peer purchase information even after a consumer has tried it. But we hypothesize it is *less* sensitive to such information after personal experience.

To summarize, if generic shares appear to increase when peer purchase shares are displayed, this could be because (1) generic-buying customers put less weight on the behavior of their peers as a signal of a drug's quality, consistent with the notion that these products are experience goods, and/or (2) brand-buying customers have priors farther away from the posted generic shares. In Section 4.5, we discuss customer surveys we conducted to explore these distinct possibilities.

### **3. Setting and Intervention**

#### *3.1 Retail Setting*

We conducted a four-week intervention in six northern California locations of a national supermarket chain. The retailer offers its own store-label (i.e. in-

house brand) household and food items in addition to over-the-counter drugs. A large share of the locations have in-store pharmacies, where consumers can ask a pharmacist questions about any drugs. However, a pre-treatment survey we conducted suggests that only about 5% of OTC customers seek the input of pharmacists and that pharmacists' typical responses were similar. The shelf layout of OTC products is largely uniform across stores, with store-label versions often, but not always, placed adjacent to their brand counterparts. We were given permission to post labels beneath the price tags of generic versions of products in randomly selected drug classes (see Figure 1 below).<sup>11</sup>



**Figure 1. Example of Label Placement**

### *3.2 Experimental Intervention: Three labeling tests*

For four-weeks, labels were posted beneath the price tags of generic products in treated drug classes in treated stores. The content of the labels differed across three tests corresponding to our three hypotheses, and further across two label versions for tests 2 and 3: label examples are numbered below as tests 1, 2a, 2b, 3a, and 3b. Test 2b (3b) differs from test 2a (2b) only in how the same

---

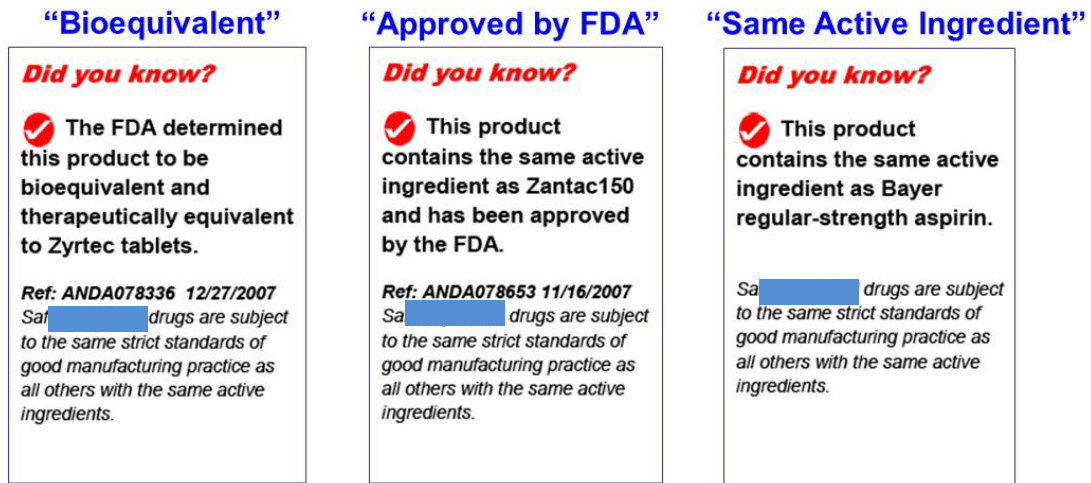
<sup>11</sup> Labels were reposted each week after price tags were updated by the retailer. We thank Yann Pannasie, Raymond Gong, Karen Yao, Caitlin Crooks, Feyisola Shadiya, Brian Mitchell, Roni Hilel, Kyle Kennelly, Jonathan Arenas, Kathy Hua, Fanglin Sun for research assistance in gathering data and implementing the label experiment. We also thank Ishita Arora, Samantha Derrick, Kathy Hua, and Ye Zhong for helping us gather auxiliary data and perform in-store surveys.

information was framed. With the exception of Test 3a, which was conducted in two different stores, each of these tests was conducted in one store.

**Test 1.** To test the hypothesis that consumers lack basic information on brand/generic drug comparability, we created labels that described the similarity between brand and generic products as specifically as possible. Each treated drug product received one of the three labels displayed in Figure 2. The strongest statement we used was: “*The FDA determined this product to be therapeutically equivalent and bioequivalent to [corresponding brand product],*” taken verbatim from the FDA approval letter, for drugs with such approval letters available on the FDA website. The second statement used was “*This product contains the same active ingredient as [corresponding brand product] and has been approved by the FDA,*” shown with the reference number and date of FDA approval. This label appeared on products for which we found notices of FDA approval, but either no electronically available letter, or a letter that did not include any statement about bioequivalence. The third statement, which was posted for older-generation drugs whose manufacturers need not seek explicit approval from the FDA prior to marketing a generic, was “*This product contains the same active ingredient as [corresponding brand product].*”<sup>12</sup>

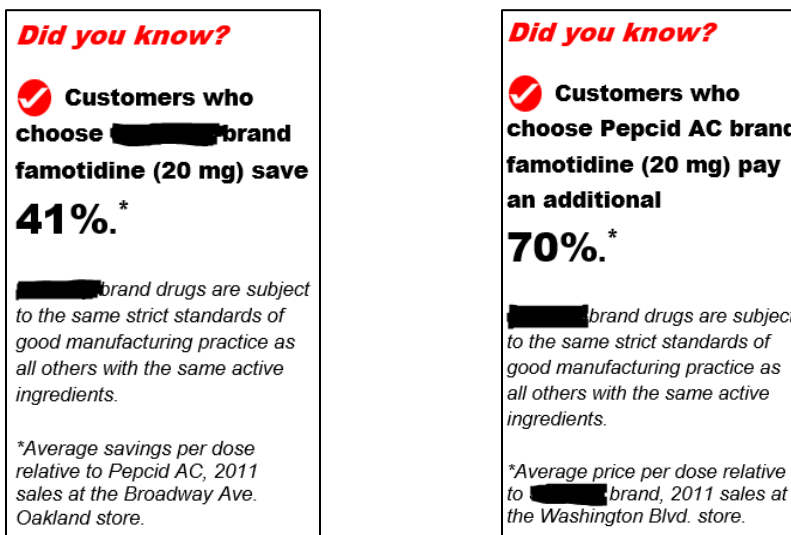
---

<sup>12</sup> As described in Section 2, for older-generation drugs such as acetaminophen or aspirin, rules regarding the production of the drug are reported in an FDA monograph, and new manufacturers are not required to apply for approval to market their own versions of such drugs. Thus, we could not use a statement as strong as the others for these products.



**Figure 2. Labels highlighting comparability/quality**

*Test 2a.* To test for inattention to price differences, we posted labels stating “Customers who choose this product save X%” with a footnote specifying that the savings was relative to the specified brand product per dose. X ranged from 14% to 68% in the products labeled and an example of one such label is below.



**Figure 3. Labels highlighting the price difference.**

[On left: Test 2a label example; On right: Test 2b label example]

**Test 2b.** In another store, we highlighted the price differences in a different way, by stating “*Customers who choose [corresponding brand product] pay Y% more than the generic alternative.*” In this type of label, the price difference is framed as a loss rather than a gain. Also, for the same brand and generic prices, Y will be a larger number than X, because the generic price is a smaller denominator. For these reasons, we hypothesized that Test 2b would have a stronger effect than Test 2a. Note, however, that the label was placed below the generic product, as we were not permitted to place labels below branded products.

**Test 3a.** To test for observational learning, we posted labels stating “*X% of customers in this store choose this product instead of [corresponding brand product].*” The values of this share were calculated for each product and each store, using either the previous year’s sales data (Jan-Dec 2011) or the first three months of the current year (Jan-March 2012). To obtain quasi-exogenous variation in the value of the share displayed, holding constant the product and the store, we alternated which method of calculation was used in each store’s labels, each week. An example of one such label is below.



**Figure 4. Labels with Generic purchase rate of Peers**

[On left: Test 3a label example; On right: Test 3b label example]

**Test 3b.** An alternate way to frame the information displayed in Test 3a is to report the share of customers who buy the brand product, e.g. “*Y% of customers in this store choose [corresponding brand product] instead of this product.*” If the mere act of bringing attention to the purchase of a specific product leads consumers to buy it, or if the statement is read as an implicit endorsement of a particular product, then Test 3b could have a different effect than Test 3a. If, instead, both labels only affect purchases insofar as they shift customers’ beliefs about what others buy, then Tests 3a and 3b should have the same effect on consumer purchases.

### *3.3 Data and Summary Statistics*

We use two types of data: first, a panel data set of store-level weekly sales for OTC drugs at the UPC level for all 56 stores in the geographic division of our treated stores, where a week starts on a Wednesday and ends on a Tuesday at store closing. Our main pre-period, for which we have store-level data available for all these stores, is the 6 weeks prior to treatment: weeks 14 to 19 of 2012 (the experiment was conducted during weeks 20-23). We also have data covering all stores for the same set of weeks in 2010 and 2011, which we use to conduct placebo tests.

Second, we use a transaction-level dataset with household identifiers, based on customer loyalty cards. This dataset includes the six treated stores and six stores that were pre-selected as a similar set of comparison stores, and spans the period from 2011 week 23 (one year preceding our treatment period) to 2012 week 38 (15 weeks after our treatment ended). These data allow us to investigate how customers shopping during the treatment period changed their behavior relative to their prior purchase patterns, whether different types of shoppers responded differently to the experimental treatments, and whether any changes persisted beyond the intervention.



### *OTC Product Classes*

Our analysis includes the twelve largest OTC drug classes that offer generic (store-label) versions for the majority of the products, broadly grouped into the categories of pain relief, allergy/cold symptoms relief, and digestive/stomach relief. A drug class may include competing products that work in a similar way (e.g. non-sedating antihistamines) or one active ingredient alone (e.g. acetaminophen/Tylenol). We excluded children's and infants' drugs to focus on products a shopper was likely choosing for own consumption. Since our goal was to study brand-generic substitution rather than therapeutic (across product) substitution, we further combined classes that are commonly interchanged (for example, ibuprofen, acetaminophen, and naproxen were grouped as "non-aspirin pain relief," and proton pump inhibitors and H2 acid blockers were grouped as "acid reflux relief"). We then randomly chose four of the following eight class groups to be treated: oral allergy, nasal allergy, acid reflux relief, laxatives, non-aspirin non-nighttime pain relief, aspirin pain relief, nighttime pain relief, and cough/cold/flu products. Summary statistics for the treated and untreated products across the three categories of pain relief, allergy/cold symptoms relief, and digestive/stomach relief appear in Appendix Table A1.

### *Store-level data*

The data contain the gross revenue, net revenue (net of promotions) and the total quantity sold of a particular UPC in a given week in a given store, with which we calculate each item's gross price and net price.<sup>13</sup> Prices may be adjusted by the retailer weekly.<sup>14</sup> For the rest of the paper, we use "price" to refer to the net or

---

<sup>13</sup> The revenue is obtained as two columns (net and gross) in the raw data that are equal to each other if the product was not on promotion during a certain week in a certain store. Those two revenue columns will differ if there are promotions: the net column will feature a smaller dollar value than the gross column. If we divide both revenue variables by the quantity sold, we obtain the gross shelf price and the average promotional price.

<sup>14</sup> These adjustments are done at the same time in all stores, between Tuesday's closing and Wednesday's opening. Thus, we define weeks in our data as beginning on Wednesday and ending on Tuesday.

promotional price, given that this is the price faced by the majority of the customers.<sup>15</sup>

We collapse our data to the level of the active ingredient and dosage combination rather than the UPC. For example, if 500mg acetaminophen is sold in quantities of 12, 30, and 100, and in gelcaps as well as tablets, we combine all of these UPCs into one observation, but another strength acetaminophen would be grouped separately. We use the term “product” to refer to a set of brand and generic formulations with the same active ingredient at the same dosage level. Thus, we refer to product “generic share” as the share of quantity (i.e. packages) purchased, within this set, for the store’s private label (generic) versions. We compute an average price per unit (i.e. per pill) for the brand and generic versions of each product by dividing the sum of net price paid by the total unit quantity sold in each store-week, and also the price of the generic version as a share of the unit price of the brand version. Lastly, we create indicator variables for “brand on sale” and “generic on sale” which equal 1 if any of the UPC’s for the brand or generic versions of the product are offered at a promotional price.

Table 1 reports descriptive statistics across the stores assigned to Test 1, Test 2, Test 3, and the control group, during the pre-treatment period, for both treated products and untreated products. As we would expect, there are no significant differences in prices across these stores. Sales quantities, however, tend to be lower in the stores assigned to Tests 2 and 3 than in the Test 1 store and the control stores, for both treated and untreated products. The generic share of treated products purchased is smaller in the Test 2 stores than in the control stores and this difference is marginally significant ( $p=0.081$ ; see Appendix Table A2 for standard errors, clustered by product, of the difference between each set of treated stores and

---

<sup>15</sup> A loyalty card for this retailer is required to obtain the promotional price. Since our household level dataset is obtained through loyalty card purchases, we can compute the share of purchases that are made using a loyalty card, and thus, at the promotional price. Across the six treated stores, this share varies from 79% to 86% with a mean of 83%.

control stores). Apart from this, and the sales quantity differences driven by store size, the differences between stores receiving treatments and control stores are not statistically significant in the pre-treatment period.

#### *Household-level data*

Longitudinal purchases at the household level are available through purchases made via loyalty cards. At this retailer, discounts posted throughout the store are only available via loyalty card and their use is frequent, 89% of OTC drug purchases. In our household-level dataset, we include households who made a purchase of any OTC drug in our sample during the pre-treatment or treatment period, and who also have prior purchases (linked by loyalty card) of any OTC drug in our sample from week 48 of 2011 to week 12 of 2012. Note that the prior purchases are not required to be of the same type of drug they are purchasing during the treatment or pre-treatment period, to avoid further reducing the size of this sample. We create the following control variables: total number of OTC purchases during the period from 2011 week 48 to 2012 week 12, the share of those purchases that were for a generic version of a product, an indicator for having purchased product  $j$  (brand or generic version), and an indicator for having purchased any generic version of product  $j$  where  $j$  is the product they are observed to purchase during the pre-treatment or treatment period. We also calculate the household's average percentage discount on total purchases at the store, as a proxy for price sensitivity.

## **4. Empirical Specifications and Results**

### *4.1 Effects of Labeling Interventions: Store-level analysis*

In the store-level analysis, we measure the effects of the labeling interventions on OTC drug purchases at the store-week-product level. The two outcomes of interest are the generic share of each product sold, computed as  $gs =$

$q_{gen} / (q_{gen} + q_{brand})$ , and the total quantity sold  $Q = (q_{gen} + q_{brand})$ .<sup>16</sup> Using a differences-in-differences approach, we estimate the effect of our treatments by comparing the change in the sales of treated OTC products from the six-week pre-treatment period to the four-week treatment period, in the treatment stores versus the control stores. We also implement a triple difference-in-differences that compares the difference-in-difference of treated products to that of untreated products.

To illustrate the approach, we first report the pure difference-in-differences in generic share for all of our labeling interventions, pooled together as the “treated” set of stores, in Table 2. This table shows the generic shares, averaged across weeks, products and stores, for the treatment period and the pre-treatment period, in treatment stores and control stores, among both treated and untreated products. The number of observations in each cell represents the number of product-store-week observations in the treated or untreated group of products, stores, and period.

The top panel, corresponding to treated products, shows that in the pre-treatment period, mean generic shares were 48.8% in the control stores and 45.5% in the treatment stores, an insignificant difference. From the pre-treatment period to the treatment period, average generic shares of treated products increased by 4.6 percentage points in the treatment stores and decreased by 1.2 percentage points in the control stores. The increase of average generic share in the treatment stores was marginally significant, as was the difference-in-differences ( $DD_{t.p.}$ ) estimate of 5.8 percentage point increase in treated products within treated stores pooled together, relative to control stores.

The bottom panel shows the parallel comparisons for untreated products. Among these products, the change in generic share from the pretreatment period to the treatment period was -0.014 in treatment stores and 0.001 in control stores,

---

<sup>16</sup> Note that we focus on the number of packages purchased as *quantity*, rather than the number of daily doses.

leading to a negative but not statistically significant difference-in-differences ( $DD_{u.p.}$ ) estimate of -1.5 percentage points.

Lastly, the table shows the triple-differences estimate, following Chetty, Looney and Kroft (2009), which is the difference between  $DD_{t.p.}$  and  $DD_{u.p.}$ . For the three interventions pooled together, the estimate is an increase of 0.073 in the generic share of purchases and marginally statistically significant. This measure is not our primary focus, however, for two reasons: First, each of the three labeling tests could have a different impact, and second, it is plausible that by highlighting different aspects of store-label generic drugs, the treatments could have spillover effects on untreated products.<sup>17</sup>

In the tables that follow, we separately report the second difference estimates for treated products and untreated products, for generic share and total quantity sold, for each of our three labeling interventions. We add store-level and product-level fixed effects and estimate coefficients for three treatment dummies:  $T_1$ ,  $T_2$ , and  $T_3$ , which are interactions of the treatment time dummy  $t_t$  and indicators for store  $s$  being one of the stores treated with Test 1, Test 2, or Test 3 labels, respectively. The equation estimated in the odd numbered columns of Table 3 is:

$$(2) \quad Y_{jst} = \beta_1 T_1 + \beta_2 T_2 + \beta_3 T_3 + t_t + \delta_j + \delta_s + \varepsilon_{ist}$$

where  $Y_{jst}$  denotes the quantity ( $Q_{ist}$ ) or generic share ( $gs_{ist}$ ) of purchases of product  $i$  in store  $s$  in time  $t$ ,  $\delta_s$  denotes store fixed effects to control for store-specific constant factors,  $\delta_j$  denotes product fixed effects, and  $t_t$  is a treatment time dummy that is equal to one during the treatment month and equal to zero during the pre-treatment month. The coefficients on  $T_1$ ,  $T_2$ , and  $T_3$  can be interpreted as average treatment-specific changes between the pre-treatment month and the treatment month, relative to changes over this same time period in the control stores.

---

<sup>17</sup> The spillover effects could be either positive, if positive information about labeled products leads customers to infer similar positive information about unlabeled products, or negative, if customers infer that the labeled products were chosen based on having more positive attributes than the unlabeled products.

In a second specification, shown in even number columns, we add controls for whether any generic versions, brand versions, or both versions of product  $i$  were on sale in store  $s$  during week  $t$ . Also, for the specifications where generic share is the outcome, we weight each product's observations by the total quantity sold during the period 2011wk1-2012wk12.

Table 3 reports the results of these regressions. Only Test 3 had a positive and statistically significant effect on the generic share of treated products, a 6.8 percentage point increase in generic share in the preferred specification of column 2. The estimated effect of Test 2 on generic share is somewhat smaller and statistically insignificant, and the estimated effect of Test 1 appears close to zero. For untreated products, Columns 3 and 4 show that the estimated effects of all three tests are very close to zero, suggesting that the labels did not influence the purchasing of products that did not receive labels.<sup>18</sup> Columns 5-8 show the same specifications estimated for the outcome of total item quantity for treated and untreated products. Results show that Test 2 had a statistically significant effect, raising the average weekly quantity purchased by more than 1 relative to a baseline average quantity of 10.3. The magnitudes of the effects of Test 1 and Test 3 on quantity are similar but statistically insignificant on treated products, and the estimates in Columns 7-8 suggest no significant effects on the quantity of untreated products.

Note that in Table 3, we estimate the average effect of Test 1 across three different comparability statements used on the information labels, and the average effects of Test 2 and Test 3 across the two framing variations that were used in different treated stores. In Table 4, we disaggregate these results. Among the different comparability statements used on information labels, none had a significant effect on purchases. The disaggregation of Test 2 shows that one of the

---

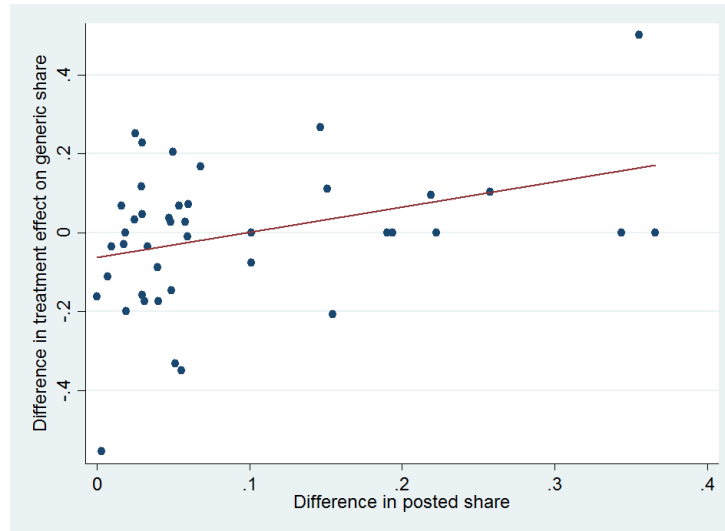
<sup>18</sup> In theory, there could be spillover effects from the labels, to other generic products in the same store.

methods of framing the price difference (“save x%” relative to the brand) has larger point estimates for both generic purchase share and quantity purchased, but these effects are only marginally significant ( $p < 0.10$ ). The other framing method (“pay an additional y%” relative to the generic), by contrast, had small and insignificant point estimates for both outcomes.

Varying the framing of Test 3 did not change the impact of posted peer purchases on generic share. The point estimates for the two framing variations are similar and both significant at the 5% level. This evidence is consistent with the labels working through observation learning. If the results instead indicated that the statement “45% of customers buy the generic” has a stronger effect on generic share than the statement “55% of customers buy the brand,” then we might infer that part of the impact was driven either by pointing out the existence of the generic product, or through an implicit perceived encouragement to buy it. The finding that these statements have equivalent effects on generic share, instead, supports the interpretation that these labels provide an informative signal used by consumers to update their expectations of the quality of the generic product: in a binary choice situation, 45% of customers buy the generic product conveys the same information as 55% of customers buy the brand product.

Since we also randomly varied whether the generic purchase shares displayed to consumers were calculated using 2011 sales data or data from the first quarter of 2012, we have exogenous variation in the share displayed. Figure 5 graphically shows the correlation between the difference in posted share and the week-by-week differences in treatment effects for the Test 3 treatment stores and products. Each point in the scatterplot corresponds to one product (active ingredient) in one of the three stores that were treated with Test 3 labels. Note that we are underpowered in this analysis, due to the fact that the items with greatest differences in store-level purchase shares of 2011 and early 2012 tend to be items that are purchased less frequently, driving greater variance in their generic purchase

shares. The best fit linear prediction has a positive slope, suggesting that every 10 percentage point increase in the posted generic share is associated with an additional 6 percentage point increase in the predicted generic purchase rate.



**Figure 5. Effect of posted share of peers on current generic shares**

#### 4.2 Event-Study Model

The difference-in-differences model estimated in the previous section is identified under an assumption of parallel trends, i.e. assuming that sales in treated stores and untreated stores were not trending differently prior to the labeling treatment. To assess this, we estimate the following event study model including the prior weeks for which we have store-level data for all stores. For each of the three labeling interventions ( $m = 1, 2,$  and  $3$ ) and for treated and control products separately, we estimate:

$$(3) \quad GenericShare_{ist} = \sum_{t=-3}^1 \beta_t D_{sm} + \alpha_i + \alpha_s + \alpha_t + \epsilon_{ist}$$

where  $D_{sm}$  is a dummy variable equaling one if store  $s$  received labeling intervention  $m$ ,  $\alpha_i$  are product fixed effects and  $\alpha_s$  are store fixed effects. Fixed effects for biweek-of-sample  $t$  represent time periods in two-week intervals relative

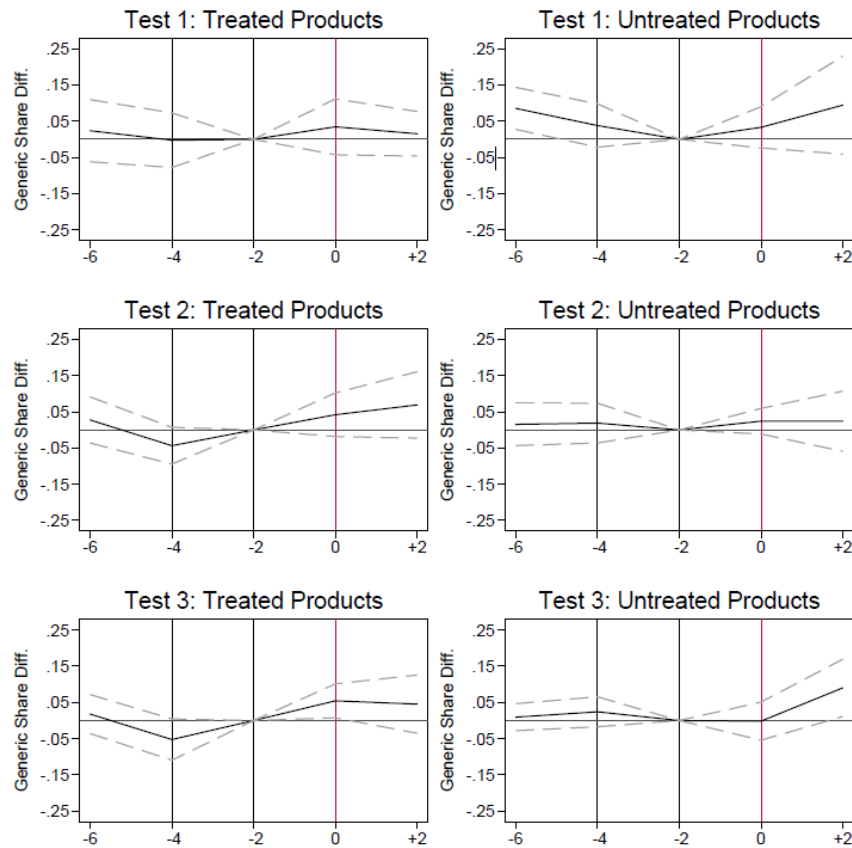


to the first two weeks of the treatment (i.e.,  $t=0$  for the first two weeks of labeling treatment, and  $t=1$  for the last two weeks of treatment.). The  $\beta_t$  vector contains the coefficients of interest, capturing the difference between the treated stores and the untreated stores during each specific period relative to the excluded biweek of sample,  $t=-1$ . If the generic shares of products are trending similarly in treated stores and control stores before the treatment periods, there should be no trend in the  $\beta_T$  coefficients during the pre-treatment period and their values should be statistically indistinguishable from zero.

Figure 6 plots the estimates we obtain from equation (3), with the  $\hat{\beta}_t$  plotted in black and the 95 percent confidence intervals, based on standard errors clustered at the product (active ingredient) level, plotted in gray dotted lines. Vertical lines separate the sample into two weeks subperiods during the pre-treatment and treatment periods; our store-level dataset does not extend beyond the treatment period. The omitted period is the two weeks prior to the start of treatment. In the periods before the labeling treatment, none of the  $\beta_t$  estimates are statistically different from zero at the 95% significance level, consistent with the assumption of pre-existing parallel trends for all treatment stores relative to the controls.

#### *4.3 Customer-level analysis*

An important concern arises when the analysis is restricted to store-level totals: We cannot confirm whether existing buyers of OTC products within a store are shifting their purchases towards generics, or if instead the labels attract the attention of different customers, who may already be buyers of generics at other retailers. Household-level data allows us to control for the past shopping choices of the customers observed during the treatment period. This enables us to rule out that the effects are driven by the composition of shoppers changing, and furthermore to test whether the labels have different effects on consumers who have previously purchased generic versus branded OTC products.



**Figure 6. Event-Study Analysis of Generic Share**

An initial look at this data reveals strong habit persistence in the choice of brand versus generic formulations. For example, of the 27,085 household-drug combinations that we observe with two or more purchases of the same drug, 51% only ever buy the brand.<sup>19</sup> Of the remaining 49%, 80% buy the generic in the first purchase that we observe, and this group continues choosing the generic at a rate of 83% in future purchases.

For this analysis, we focus on the purchases made by individuals whose loyalty cards show a prior purchase of at least one OTC drug from week 48 of 2011 to week 12 of 2012, our past purchases observation period. Each individual

<sup>19</sup> This pattern persists when increasing the minimum number of purchases. With a minimum of 5 purchases of the same product, 46% have only bought the brand, and with a minimum of 10 purchases, 42% have only bought the brand.

purchase of a treated or untreated product is an observation, and the dependent variable, *generic*, is equal to one if the purchase made was for a generic version. We estimate a linear probability difference-in-differences model, similar to the store-level equation (1) in Columns 1 and 4 of Table 5, for treated and untreated products. The equation is:

$$(4) \text{Generic}_{ijst} = XB + \beta_1 T_1 + \beta_2 T_2 + \beta_3 T_3 + t_t + \delta_j + \delta_s + \varepsilon_{ijst}$$

where  $\text{Generic}_{ijst}$  is equal to 1 if the store-brand version of product  $j$  is purchased by customer  $i$  in store  $s$  in time week  $t$ . Controls in  $X$  always include dummies for whether brand, generic, or both versions of the product were on sale during a given week in store  $s$ , and add household level controls in subsequent columns. As in the store-level analysis,  $t_t$  identifies the four-week treatment period;  $T_1$ ,  $T_2$ , and  $T_3$  are the treatment period interactions with the stores used in tests 1-3;  $\delta_j$  are product fixed effects and  $\delta_s$  are store fixed effects.

Table 5 shows the results of this estimation, which qualitatively matches those of the store-level analysis. In the first column (before adding household-level control variables), the point estimate of the effect of Test 1 is near zero and insignificant, the estimate for Test 2 shows a 4.9 percentage point increase in the probability of choosing a generic OTC drug over its brand-name counterpart, and the estimate for Test 3 represents a 7.9 percentage point increase. Both the Test 2 and Test 3 effects are statistically significant at the 5% and 1% levels, respectively. Column 4 shows that none of the three tests had significant effects on the probability of generic choice for untreated products, with point estimates remarkably close to zero.

Next, in columns 2 and 5, we add household-level controls based on the past purchases observation period: the generic share of OTC purchases, any previous purchase of the product now being purchased, any previous purchase of the generic version of this product, and the past average percentage discount obtained on total spending, a proxy for price sensitivity. These covariates have significant

explanatory power, raising the R-squared of the regression from .12 to .36 for treated products. If the type of customer making purchases changed between the pre-test period and the labeling test period, differently in treated stores versus control stores, we would expect the estimated treatment effects to shrink between columns 1 and 2. When controlling for past purchasing behavior, the estimated effect of Test 2 shrinks from a 4.9 to a 3.2 percentage point increase in generic share and is no longer statistically significant. For Test 3, the effect is a 6.1 percentage point increase in generic share and remains highly statistically significant. This suggests that different composition of shoppers does not explain a significant portion of Test 3's impact.

In columns 3 and 6, we test for heterogeneity of treatment effects based on prior purchasing experience. Given that symptom-treating over-the-counter drugs are experience goods, customers who have already purchased this store's generic version of a certain drug may be much less responsive to the information provided in certain labeling tests. To test this, we add an interaction term between each test and an indicator for having previously purchased the generic version of the same product.<sup>20</sup> The baseline effect of Test 3, now representing the effect only on customers who have never before purchased the generic version, increases to 0.082, a percentage point increase of 8.2. Since the baseline probability that a customer of this type will choose the generic is 32 percent, this is a 25% increase in likelihood of choosing the generic.<sup>21</sup> The point estimate of the interaction between Test 3 and the dummy variable for making a prior generic purchase is large and negative. Combined with the fact that customers who have previously purchased the generic version of product *j* chose the generic 82 percent of the time,

---

<sup>20</sup> We also add interactions between having previously purchased the generic and the treatment time period, and between having previously purchased the generic and indicators for the Test 1, Test 2, and Test 3 stores.

<sup>21</sup> The given probability is the average for all customers with no prior purchases of the generic version of the product purchasing treated products in the pre-test period across all six treated stores.

this large and negative point estimate suggests a smaller relative increase, only 2%, in generic purchasing for customers not previously tied to the brand, with the caveat that this coefficient is not statistically significant. The results for Tests 1 and 2 indicate no significant effect for either type of customer, and results in columns 4-6 show no significant effects of any treatment on untreated products.

Appendix Table A3 shows household-level results disaggregated by labeling statement, as done previously using store-level data in Table 4, and also includes the set of households with no observed prior purchases of OTC drugs. In the latter subsample, we find that Test 1, which showed comparability statements noting FDA approval of the generic and its confirmed bioequivalence, when applicable, had a large positive effect on generic purchase probability. This suggests that customers who have never purchased an OTC drug at this retailer may be uninformed about the comparability of store-label and national brands.

#### *4.4 Post-treatment effects on customers exposed to labels*

To explore whether the information provided in the labels leads to persistent changes in purchasing behavior, we turn to purchase data from the period following label removal, which we divide into three consecutive four-week intervals. Our access to household identifiers is crucial for this analysis, because OTC products are purchased infrequently: only twenty percent of the shoppers in any subsequent four-week period made a purchase of any OTC drug during the treatment period. Thus, it is unsurprising that testing for a different change in generic purchase rates between treatment and control stores, from the pre-treatment period to the post-treatment period, yields no significant effects (Appendix Table A4, odd-numbered columns), because the majority of customers shopping in the post-treatment period may not have seen the labels at all.<sup>22</sup>

---

<sup>22</sup> The estimated equation for the regression results in Table A4 matches equation (4) except  $PT_1$ ,  $PT_2$  and  $PT_3$  represent interactions between treated stores, and a post-labeling period replaces the treat-time dummy:

$$Generic_{ijst} = XB + \beta_1 PT_1 + \beta_2 PT_2 + \beta_3 PT_3 + \delta_t + \delta_j + \delta_s + post_t + \varepsilon_{ijst}$$

To test for persistent changes among shoppers who were present in the OTC drug aisles during the four-week labeling period, we adopt another difference-in-differences approach. We compare pre- to post-intervention changes in generic purchase rates between customers who were or were not exposed to the labels, using any purchase during the treatment period (of a treated or untreated product) as a proxy for label exposure. We estimate this model within treated stores alone:

$$\begin{aligned}
 (5) \text{ Generic}_{ijst} &= XB + \beta_1 PT_1 + \beta_2 PT_2 + \beta_3 PT_3 + \beta_4 PT_1 * Exposed \\
 &+ \beta_5 PT_2 * Exposed + \beta_6 PT_3 * Exposed \\
 &+ \beta_7 T1Store * Exposed + \beta_8 T2Store * Exposed \\
 &+ \beta_9 T3Store * Exposed + \delta_j + \delta_s + \varepsilon_{ijst}
 \end{aligned}$$

Results for the treated stores are shown in Table 6. Estimated effects are somewhat noisy, but suggest that the test 3 labels continued to influence the choice of products among customers who had made purchases during the label period. Appendix Table A4 (even numbered columns) shows qualitatively similar results of a triple-difference specification that includes control stores, similarly defining exposed customers in control stores as those who purchased any OTC product during the treatment time period.

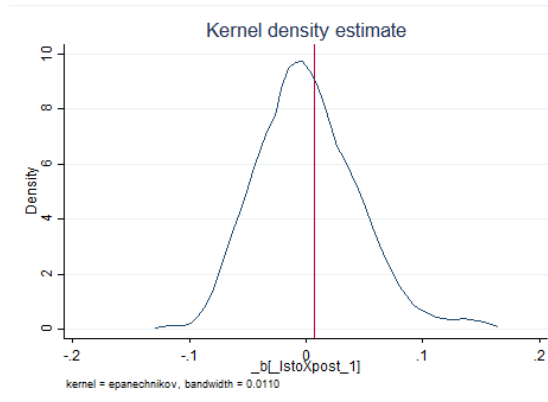
Note that this approach assumes the labels did not change the composition of shoppers making an OTC purchase at a given time, i.e. being “exposed” to the labels is quasi-random. The fact that we find no significant change in the quantity of purchases during the labeling period, for Test 3, is consistent with the assumption that making a purchase is primarily driven by the need to purchase an OTC drug during the treatment time period. Also, in Appendix Table A5, we find no evidence for exposed customers having different propensities to choose generic products prior to the treatment, compared against customers at the same store who did not make an OTC purchase during the treatment period.

#### *4.4 Robustness*

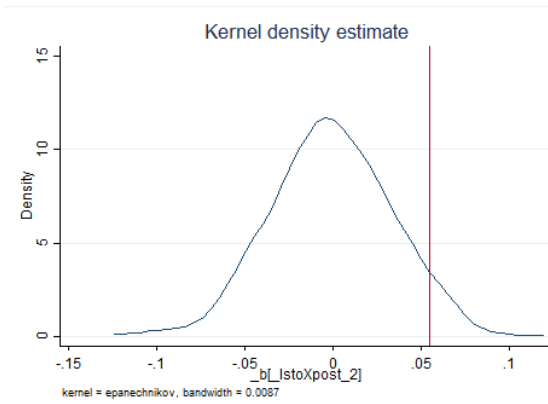
To assess the robustness of our store-level results, we tested for serial correlation in store-product observations at the week level, and could not reject the null hypothesis of no serial correlation in generic share. Although the test detected serial correlation in quantity purchased, our results using quantity as an outcome do not change when correcting for serial correlation, and our results for generic share do not change when including quantity as a regressor (Appendix Table A6).

We also show that results are not dependent on our choice of control stores: Appendix Table A7 shows that we obtain similar results when using all other stores from the geographic division (N=50) as untreated. To verify that our results for generic share are not driven by short-term seasonal trends specific to the stores that we treated, we conducted a falsification test. We used sales data from 2010 and 2011 to test for placebo difference-in-differences effect on generic share in the stores and weeks we treated in 2012, relative to untreated stores, but during the prior years. Results are shown in Appendix Table A8 using our six designated control stores, and in Table A9 using all untreated division stores as the control group. There are no significant placebo effects in the treated weeks and stores in years when there was no treatment.

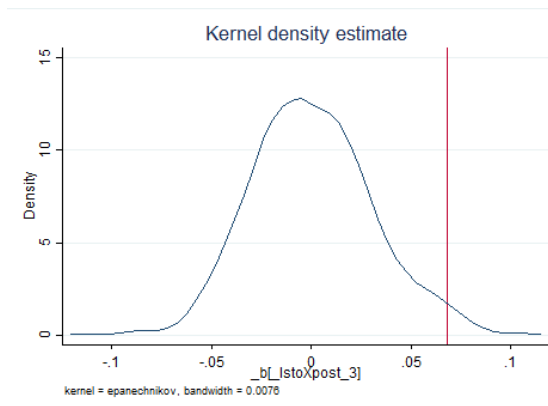
Finally, we used a randomization inference test to nonparametrically estimate p-values for the main estimates of our three treatment effects. We used the data from the pre-treatment and treatment periods in all 56 stores. In each of 500 permutations, we randomly drew one store to be assigned Treatment 1, two stores to be assigned Treatment 2, three stores to be assigned Treatment 3, and six stores to be used as control stores. The regression based on Equation (2) was re-run for each of these draws, and we plot in the three panels of Figure 7 how the point estimates for the three treatment effects compare to the point estimates obtained over 500 draws. These results show that for Treatment 3, only 4.2% of randomly selected permutations yield a treatment effect as large as the one we obtain.



Panel A: Labeling Treatment 1



Panel B: Labeling Treatment 2



Panel C: Labeling Treatment 3

Figure 7. Distribution of Point Estimates Drawn from Randomly Chosen Assignments.

Note: Point estimate from the actual treatment/control assignment shown as a red vertical line.



#### *4.5. Mechanisms for the effect of information on peer purchases*

As discussed in Section 2.2, a simple model of observational learning, in which shoppers place a constant weight on the choices of others, would predict that the posted share should have opposite effects on people whose priors of their peer shoppers' generic purchase shares are above, versus below, the posted shares. This could lead to a null net effect of the labels on generic share when the current share of generic customers is around 50% and priors are either unbiased or symmetrically biased towards what one personally buys. There are two reasons why posted shares close to 50% could, however, lead to a net increase in the generic purchase share: (1) generic-buying customers may put less weight on the choices of others as a signal of drug quality, because they have already tried both the brand and generic, or (2) brand-buying customers may have priors of the generic purchase share that are further from the true purchase ratios than the priors of generic-buying customers, on average.

To explore these two possible explanations, we surveyed 298 customers in three of the retailer's locations.<sup>23</sup> Customer participants were first asked to make a hypothetical choice between the brand and generic versions of an over-the-counter painkiller, with prices shown. They were asked to guess what share of other shoppers at the store would make the same choice that they did. Then, respondents were asked to consider hypothetical information stating that the share of other shoppers making the same choice they made was smaller than they had guessed, and asked whether they would still choose the brand (or generic) if this information were true, or if they would consider switching to the generic (or brand) product.

Before presenting the findings related to the two possible explanations described above, we note that the survey responses generally support the assumptions of our conceptual framework: 94% of respondents have purchased the

---

<sup>23</sup> Details are provided in Appendix B. We avoided stores that were treated during the treatment period to reduce the probability of surveying customers who had seen the posted labels.

brand version of the painkiller at some point in the past, confirming that most shoppers have personal experience with the brand-name product. By comparison, only 65% of respondents have ever purchased a generic version of the painkiller. Among consumers who have tried both the brand and the generic, 81% make a hypothetical choice of buying the generic at the typical list prices shown to them, similar to the patterns observed among repeat consumers in the household-level dataset. Of those who have never tried the generic, only 37% make the hypothetical choice of the generic in the survey. Consistent with being less likely to have tried the generic, more brand-choosing respondents than generic-choosing respondents answered “Don’t Know” or refused to answer whether they believe that the brand works better than the generic at relieving pain, that the generic works better than the brand, or that they work equally well (16% vs. 6% of generic-choosing customers,  $p=0.018$ ; see Appendix Table A10). These responses are consistent with the hypothesis that many brand buyers have imprecise priors regarding the efficacy of generic drugs, relative to the brand.

The survey responses also reveal diffuse priors regarding the share of consumers who buy the brand or generic: 50% is a modal answer, accounting for 24% of responses, and the remaining responses range from 5 to 95%.<sup>24</sup> We find that consumers who choose the brand, on average, believe that fewer consumers choose the generic (49%) than consumers who choose the generic themselves (58%,  $p<.001$ ). The guesses of those who choose the brand are actually closer, on average, to the true proportions in sales data. That is, we find no evidence that the beliefs of brand-buying consumers are less accurate than those of generic-buying consumers.

To assess how the predisposition for “observational learning” might differ between generic and brand buyers, we analyzed responses about how likely one

---

<sup>24</sup> Although the survey did not allow for non-response on this question, survey enumerators noted that many respondents wanted to skip this question, stating “I have no idea.” Such respondents, when pressed to make a guess, would typically answer “half.”

would be to change their choice (i.e. consider switching from the brand to the generic, or consider switching from the generic to the brand) if they learned that the share of customers making the same choice as them was smaller than what they guessed. Of consumers who had chosen the brand version, 20% said they would “probably” or “definitely” buy the generic if this was the case. By contrast, only 8.9% of generic buyers said they would “probably” or “definitely” buy the brand if they learned that buying the generic was less common than they thought ( $p < .01$ ).<sup>25</sup> With the caveat that these questions were framed as hypotheticals, we interpret this as suggestive evidence that brand-buying customers are more likely to be swayed towards the generic by information on the purchases of other customers than generic-buying customers are to be swayed towards the brand. This is consistent with over-the-counter drugs being an experience good; those who have already tried a product are less uncertain about its therapeutic value.

## 5. Conclusion

Unlike prescription drugs, OTC drugs are purchased by consumers with unrestricted choice over competing products and direct access to prices and standardized drug facts. Nevertheless, entrenched brand preferences among near substitutes may result from biased beliefs or uncertainty about the differences between brand and generic drugs. We implemented a labeling experiment at six locations of a national retailer to test three hypotheses for consumer aversion to generic OTC drugs: (1) lack of information regarding their similarity to the brand, (2) inattention to the price difference, and (3) biased beliefs or uncertainty about the generic that can be addressed with information on their peers’ purchases.

For the first and second hypotheses, we found only limited evidence. Labels providing comparability information had no overall impact, although the subset of them that noted FDA-confirmed bioequivalence of generics to the brand increased

---

<sup>25</sup> In both cases, X% was engineered to be a smaller percentage than what they gave as their prior, indicating that more people than they expect choose the *opposite* product as they chose. Details of this process are in the Appendix.

the generic purchase rates of consumers who had never before made an OTC drug purchase at this retailer. Labels that displayed price differences as “savings” in percentage terms increased total generic sales, but did not sway customers who had not yet tried the generic.

By contrast, we found strong evidence supporting the third hypothesis. Labels showing the share of customers at this store who buy the generic (or brand) version of each product increased generic purchase rates by 7 percentage points overall, and 8.8 percentage points among customers who had not yet tried the generic. This effect is more than 60% as large as the average increase associated with price promotions, suggesting it is equivalent to reducing the generic’s price by an additional 3.4 percent,<sup>26</sup> and is particularly strong among brand-loyal customers.

The fact that the effect was similar regardless of whether the labels provided the share buying the generic or the share buying the brand suggests that the effect was driven by observational learning rather than an implicit encouragement to choose the generic. Given that OTC drugs are consumed privately, individuals cannot easily observe the drugs that their peers purchase. Our survey evidence suggests that customers who have never yet tried the generic may be more responsive to the information that other customers frequently choose it. We conclude that some brand-loyal consumers are wary of the quality or safety of generic products but respond to the information that other customers find them acceptable.

These findings have implications for policies aimed at promoting evidence-based and cost-effective choices. Public health interventions and employer wellness initiatives using peer promotion may be more effective than simply advertising facts. Given the growing emphasis in U.S. healthcare on giving patients an active role in the choice of their treatments, this paper provides a first stab at a better understanding what types of new information can shift their choices.

---

<sup>26</sup> A price promotion of the generic reflects, on average, a 5.5% price reduction.

## References

Ackerberg, Daniel A., 2001. "Empirically Distinguishing Informative and Prestige Effects of Advertising," *The RAND Journal of Economics*, vol. 32: 316- 333.

Ackerberg, Daniel A., 2003. "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination," *International Economic Review*, vol. 44: 1007-1040.

Akerlof, George A, 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, MIT Press, vol. 84(3), pages 488-500, August.

Bronnenberg, Bart, Jean-Pierre Dubé, Matthew Gentzkow and Jesse Shapiro (2015) "Do pharmacists buy Bayer? Informed shoppers and the brand premium" *Quarterly Journal of Economics*.

Bursztyn, Leonardo, Ederer, Florian P., Ferman, Bruno and Yuchtman, Noam (2013) "Understanding Peer Effects in Financial Decisions: Evidence from a Field Experiment." Available at SSRN: <http://ssrn.com/abstract=2101391>.

Cai, Hongbin, Yuyu Chen, and Hanming Fang. 2009. "Observational Learning: Evidence from a Randomized Natural Field Experiment." *American Economic Review*, 99(3): 864-82.

Carrera, Mariana, Dana Goldman, Geoffrey Joyce, and Neeraj Sood 2018. "Do Physicians Respond to the Costs and Cost-Sensitivity of their Patients?" *American Economic Journal- Economic Policy*, 10 (1), 113-152.

Chetty, Raj, Looney, Adam, and Kory Kroft, 2009. "Salience and Taxation: Theory and Evidence" *American Economic Review* 99(4): 1145-1177.

Hossain, T. and Morgan, J., 2006: "...Plus Shipping and Handling : Revenue (Non) Equivalence in Field Experiments on eBay," *B.E. Journals in Economic Analysis and Policy: Advances in Economic Analysis and Policy*, 6, 2, 1-27.

Jin, Ginger Z., and Phillip Leslie, 2003 "The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Cards," *Quarterly Journal of Economics*, vol. 118: 409-451.

Kesselheim AS, Misono AS, Lee JL, et al. Clinical Equivalence of Generic and Brand-Name Drugs Used in Cardiovascular Disease: A Systematic Review and Meta-analysis. *JAMA*. 2008;300(21):2514-2526. doi:10.1001/jama.2008.758.

Ma, Juan, Zhaoning Wang and Tarun Khanna, “Why Advertising Safety Isn't Safe? Reminder Effect and Consumers' Negative Response to Information about Product Quality” (January 26, 2017). INSEAD Working Paper No. 2017/08/STR/EMI. Available at SSRN: <https://ssrn.com/abstract=2906179>

McFadden, D.L. and Train, K. (1996). Learning from self and others. *Journal of Political Economy*, 104(4), 684-703.

Montgomery, Cynthia A & Wernerfelt, Birger, 1992. "Risk Reduction and Umbrella Branding," *The Journal of Business*, University of Chicago Press, vol. 65(1), pages 31-50, January.

Moretti, E. (2011). Social learning and peer effects in consumption: evidence from movie sales. *The Review of Economic Studies*, 78(1), 356-393.

Saxell, Tanja. “Private Experience and Observational Learning in Pharmaceutical Demand.” Working Paper <https://sites.google.com/site/tsaxell/> 2014.

Shrank, William H., et al. "Physician perceptions about generic drugs." *The Annals of pharmacotherapy* 45.1 (2011): 31-38.

Shrank, William H., et al. "Patients' perceptions of generic medications." *Health Affairs* 28.2 (2009): 546-556.

Train K.E. *Discrete Choice Methods with Simulation*. Cambridge: University Press, 2002.

Wiggins, Stephen. and W.J. Lane, 1983. “Quality Uncertainty, Search, and Advertising,” *American Economic Review*, vol. 73: 881-894.

Wolinsky, Asher, 1995 “Competition in Markets for Credence Goods,” *Journal of Institutional and Theoretical Economics*, vol. 151, 117-131.

**Table 1. Descriptive Statistics, Pre-treatment period**

*Panel A: Treated Products*

	Control stores	Test 1 stores	Test 2 stores	Test 3 stores	All
Brand price per unit	0.41	0.41	0.42	0.41	0.41
	0.01	0.03	0.02	0.02	0.01
Generic price, as a share of brand price	0.60	0.58	0.60	0.58	0.59
	0.01	0.02	0.02	0.01	0.01
Weekly quantity sold per product	11.30	13.48	8.37	8.36	10.26
	0.56	1.20	0.67	0.47	0.34
Generic share	0.49	0.51	0.41	0.46	0.47
	0.01	0.03	0.02	0.02	0.01

*Panel B: Untreated Products*

	Control stores	Test 1 stores	Test 2 stores	Test 3 stores	All
Brand price per unit <sup>a</sup>	0.42	0.43	0.40	0.42	0.42
	0.01	0.04	0.03	0.02	0.01
Generic price, as a share of brand price <sup>a</sup>	0.52	0.51	0.52	0.55	0.53
	0.01	0.03	0.02	0.02	0.01
Weekly quantity sold per product	9.60	13.85	8.28	7.98	9.33
	0.54	1.87	0.82	0.77	0.39
Generic share	0.43	0.44	0.39	0.46	0.43
	0.01	0.26	0.32	0.33	0.01

<sup>a</sup>Liquid and nasal products are excluded from the price calculations since units could not be determined. **Notes:** This table shows means and standard errors across all product-week-store observations for pre-treatment weeks 2012wk14-2012wk19. "Weekly quantity" is the number of packages sold per product (same active ingredient but may vary in units (count of doses), brand, flavor, and pill type). Prices are in dollars per unit (count), inclusive of discounts, averaged over the different UPCs sold for each active ingredient and then averaged across the different products in the treated and untreated groups. "Generic price as share of brand price" is the per-unit price of generic formulations divided by the per-unit price of brand formulations. "Generic share" is the number of generic packages of each product divided by the total number of packages sold for each product, by week.

**Table 2: Difference-in-Differences in Generic Share, Pooled Treatments***Panel A: Treated Products*

	Control Stores	Treatment Stores		Differences
Pre-treatment means	0.488	0.455		-0.033
s.e.	0.053	0.043		0.024
N	485	484		969
Treated Time means	0.476	0.501		0.025
s.e.	0.057	0.048		0.028
N	319	322		641
Difference over time	-0.012	0.046	DD <sub>t.p.</sub> =	0.058
s.e.	0.018	0.025		0.020
N	804	806		1610

*Panel B: Untreated Products*

	Control Stores	Treatment Stores		Differences
Pre-treatment	0.426	0.433		0.007
s.e.	0.045	0.038		0.020
N	592	576		1168
Treated Time means	0.426	0.419		-0.008
s.e.	0.047	0.042		0.026
N	390	374		764
Difference over time	0.001	-0.014	DD <sub>u.p.</sub> =	-0.015
s.e.	0.024	0.021		0.032
N	982	950		1932
Triple difference			DDD =	0.073
s.e.				0.037
N				3542

---

Notes: Stores treated with Test 1, Test 2, and Test 3 are all included as the pooled treatment stores. Six treated stores and six control stores are included. Pre-treatment time is weeks 14-19 in 2012; treated time is weeks 20-23 in 2012. Standard errors are clustered by product.



**Table 3: Treatment Effects of Each Labeling Intervention, Store-level data**

	Generic share				Quantity			
	Treated Products		Untreated Products		Treated Products		Untreated Products	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Test 1: Comparability Statement	-0.037 (0.032)	0.0067 (0.026)	-0.017 (0.045)	0.0038 (0.024)	1.22 (0.77)	0.91 (0.74)	0.092 (1.54)	0.17 (1.53)
Test 2: Price comparison	0.053 (0.038)	0.055 (0.033)	-0.037 (0.044)	0.0060 (0.021)	1.29 (0.55)	1.05 (0.47)	0.29 (0.41)	0.23 (0.41)
Test 3: Observational Learning	0.073 (0.028)	0.068 (0.027)	0.000076 (0.033)	0.0071 (0.020)	1.21 (0.80)	1.24 (0.81)	0.39 (0.32)	0.37 (0.29)
Generic on promotion		0.11 (0.023)		0.085 (0.026)		0.86 (0.38)		1.01 (0.37)
Brand on promotion		-0.13 (0.041)		-0.15 (0.039)		1.64 (0.41)		0.73 (0.55)
Both on promotion		-0.030 (0.032)		0.013 (0.049)		0.68 (0.56)		2.60 (1.05)
Weighted by quantity	No	Yes	No	Yes	--	--	--	--
N	1610	1610	1932	1932	1680	1680	2040	2040
Dependent variable mean	0.47	0.47	0.43	0.43	10.3	10.3	9.33	9.33

Notes: Observations are at the week-store-drug level. Quantity is total products purchased of both brand and generic versions. Generic share is the share of generic purchases within that quantity. Test 1 was conducted at one store, test 2 was conducted at two stores, and test 3 was conducted at three stores. All tests were conducted in the same six week period. In six control stores, no labeling tests were conducted. "Generic on promotion" and "Brand on promotion" are dummy variables indicating that any of the applicable products are on a price promotion during that week in that store. Store, product, and week fixed effects are included in all specifications. Standard errors clustered at the product level are in parentheses; significance stars are omitted.

**Table 4: Treatment Effects by Label Content and Framing**

	Generic share		Quantity	
	Treated Products	Untreated Products	Treated Products	Untreated Products
	(1)	(2)	(3)	(4)
<b>Test 1: Comparability Statement</b>				
"Same active ingredient"	-0.000 (0.031)		2.368 (1.463)	
"... and approved by the FDA"	-0.006 (0.056)		-0.239 (1.191)	
"FDA determined bioequivalence"	0.042 (0.027)		-0.766 (1.205)	
No Label (untreated)		0.004 (0.024)		0.173 (1.532)
<b>Test 2: Price Comparison</b>				
Framing: Save X%	0.075 (0.040)	-0.028 (0.036)	1.546 (0.748)	0.007 (0.555)
Framing: Pay Y% More	0.019 (0.049)	0.038 (0.024)	0.549 (0.617)	0.447 (0.430)
<b>Test 3: Observational Learning</b>				
Framing: X% choose generic	0.064 (0.031)	0.019 (0.020)	1.408 (0.841)	0.295 (0.367)
Framing: Y% choose brand	0.081 (0.032)	-0.029 (0.032)	0.910 (0.901)	0.527 (0.451)
Weighted by quantity	Yes	Yes	--	--
N	1610	1932	1680	2040
Dependent variable mean	0.47	0.43	10.3	9.33

Notes: Observations are at the week-store-drug level. Quantity is total products purchased of both brand and generic versions. Generic share is the share of generic purchases within that quantity. The label statements for Test 1 were varied at the level of the product, based on generic product's FDA status. The framing of the information presented in Tests 2 and 3 was varied at the store level. Each framing variation for Tests 2 and 3 was tested at one store, with the exception of the first framing of Test 3 ("X% choose generic") which was tested at two stores. Controls for price promotions are included as in Table 3, and the generic share regressions are weighted by purchase quantity. Standard errors clustered at the product level are in parentheses.

**Table 5: Treatment Effects and Heterogeneity at the customer level**

<i>Y = Generic purchased</i>	Treated Products			Untreated Products		
	(1)	(2)	(3)	(4)	(5)	(6)
Test 1: Comparability Statement	-0.007 (0.025)	-0.012 (0.022)	0.002 (0.032)	0.005 (0.024)	0.000 (0.025)	-0.009 (0.027)
Test 2: Price comparison	0.049 (0.017)	0.032 (0.019)	0.027 (0.028)	-0.009 (0.031)	-0.013 (0.025)	-0.031 (0.027)
Test 3: Observational Learning	0.079 (0.023)	0.061 (0.019)	0.088 (0.023)	0.004 (0.024)	0.019 (0.027)	0.008 (0.033)
Test 1 x Previous Use of Generic			-0.035 (0.091)			0.023 (0.058)
Test 2 x Previous Use of Generic			0.045 (0.086)			0.060 (0.062)
Test 3 x Previous Use of Generic			-0.071 (0.037)			0.026 (0.040)
Previous Use of Generic		0.443 (0.025)	0.463 (0.031)		0.390 (0.032)	0.395 (0.036)
Generic share of all OTC drugs previously purchased		0.313 (0.012)	0.313 (0.012)		0.303 (0.018)	0.302 (0.018)
Previous use of purchased product (brand or generic version)		-0.205 (0.025)	-0.205 (0.025)		-0.177 (0.034)	-0.177 (0.035)
N	9166	9166	9166	9747	9747	9747
R-sq	0.118	0.357	0.358	0.115	0.341	0.342
Dependent variable mean	0.45	0.45		0.46	0.46	
		No Previous Use of Generic	0.31			0.34
		Previous Use of Generic	0.82			0.76

Notes: Linear probability models for the choice of a generic. Observations represent each individual purchase of a treated or untreated drug in the pre-treatment or treatment period. The sample includes households with at least one prior purchase of an OTC product prior to the start of the pre-treatment time period. "Test 1," "Test 2," and "Test 3" treatment indicators are interactions for treated store and treatment time period. "Previous Use of Generic" is an indicator for the household having previously purchased the generic version of the OTC drug currently being purchased. Store and product fixed effects, controls for price promotions, and treatment period dummy are included. Models 2, 3, 5, and 6 include household-level controls: total number of prior OTC drug purchases and past average percentage discount on total purchases at the retailer. Models 3 and 6 include interaction terms between treatment time and previous use of generic, and between previous use of generic and indicators for stores receiving test 1, test 2, and test 3. Standard errors, clustered at the product level, are in parentheses. Significance stars are omitted.

**Table 6: Post-treatment effects on choice of generic**

*Y = Generic purchased*

	Treated Products within Treated Stores			Untreated Products within Treated Stores		
	Weeks 24-27 (1)	Weeks 28-31 (2)	Weeks 32-36 (3)	Weeks 24-27 (4)	Weeks 28-31 (5)	Weeks 32-36 (6)
Test 1 x Post x Exposed	0.046 (0.077)	-0.051 (0.068)	-0.0078 (0.058)	-0.0037 (0.040)	-0.0034 (0.034)	0.0026 (0.044)
Test 2 x Post x Exposed	-0.056 (0.059)	-0.012 (0.050)	-0.0011 (0.043)	-0.087 (0.040)	-0.048 (0.049)	-0.050 (0.060)
Test 3 x Post x Exposed	0.10 (0.041)	0.087 (0.051)	0.059 (0.045)	-0.069 (0.036)	-0.019 (0.028)	0.00059 (0.029)
Baseline effect: Post period						
Test 1 x Post	0.0097 (0.047)	0.069 (0.044)	0.028 (0.033)	-0.0081 (0.032)	0.0037 (0.038)	-0.015 (0.027)
Test 2 x Post	-0.00069 (0.032)	0.042 (0.039)	0.027 (0.032)	0.032 (0.031)	0.020 (0.036)	-0.00083 (0.037)
Test 3 x Post	-0.00019 (0.038)	0.022 (0.037)	0.026 (0.032)	0.040 (0.029)	0.023 (0.020)	-0.0029 (0.023)
Baseline effect: Exposed						
Test 1 stores x Exposed	-0.015 (0.039)	-0.0085 (0.040)	-0.013 (0.040)	-0.039 (0.021)	-0.033 (0.020)	-0.028 (0.026)
Test 2 stores x Exposed	-0.0068 (0.053)	0.00066 (0.053)	-0.0064 (0.054)	0.0039 (0.030)	0.0072 (0.031)	0.0067 (0.033)
Test 3 stores x Exposed	-0.057 (0.039)	-0.047 (0.036)	-0.051 (0.037)	-0.00081 (0.029)	0.0051 (0.029)	0.0038 (0.029)
N	3443	3733	4251	4033	4419	5289
Dependent variable mean	0.45	0.45	0.45	0.47	0.47	0.47

Notes: Linear probability models for the choice of a generic. Observations represent each individual purchase of a treated or untreated drug during the specified post-period combined with the pre-treatment period. For each test, "Test X x Post" is an interaction for a store treated with labeling test X and the specified post-treatment time period, capturing the difference in generic purchase share overall between the pre-treatment period and the specified post-treatment period for that set of stores. "Exposed" is an indicator for the individual having made any purchase of a (treated or untreated) OTC drug during the treatment time period, indicating their presence in the OTC aisles of the store. The interactions "Test X x Exposed capture the average difference between generic purchase rates among exposed and unexposed customers within Test X store(s) in the pre-treatment period. Standard errors in parentheses. Significance stars omitted.