Online Appendix

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

Mariana Carrera and Sofia Villas-Boas

Table of Contents

A. Selection of products and stores for treatment

A.1. Choice of drugs to be treated

We chose the drugs to be treated through a stratified cluster randomization process. We began with the store's categorization of drug classes, or sets of competing products that work in similar ways (e.g. non-sedating allergy drugs includes loratadine, fexofenadine, and cetirizine). In some cases we grouped together a few drug classes that treat the same condition (e.g. non-aspirin pain relievers ibuprofen, acetaminophen, and naproxen) so that products seen as potential substitutes would be either all treated or all untreated. Second, we stratified the cluster randomization within the three broader categories of pain relief, allergy/cold relief, and digestive/stomach. Our intention was to treat a broad range of products treating different types of health conditions, but also to keep clusters of substitutable products intact, either all treated or all untreated. Since drug classes within these categories vary widely in their time since generic entry and price per unit, we do not have perfect balance in the characteristics summarized in Table 1. Table A1 reports p-values for the difference between treated and untreated products within each symptom category.

A.2. Choice of treatment and control stores

In choosing our six treatment stores, we first ruled out stores that were farther than 45 minutes' driving distance from the university, as we would be manually adding labels on the same morning each week at all six stores. The second criterion was high weekly sales quantities, to maximize the number of customers we would reach per label. During visits to each store we confirmed that OTC products were similarly organized and displayed on the shelves.

Five of the six stores selected for treatment include in-store pharmacies. We surveyed the pharmacists in these stores about their interactions with OTC consumers. Pharmacists reported that 5-10 customers per week approach them to ask questions about OTC products. Our household-level sales data indicate that at these stores, the average number of loyalty card holders purchasing an OTC product each week is 190, and loyalty card holders account for 83% of purchases. Thus, we approximate that less than 5% of OTC customers seek the input of pharmacists. We also found through this survey that all pharmacists gave similar answers to customer inquiries regarding the equivalence of generics, emphasizing that they contain the same ingredients as the name-brand products.

For the six treatment stores, we obtained pre-treatment data from 2011 and the first 13 weeks of 2012, which we used to compute mean percent savings and mean generic share at the store-product level to be displayed in the labels for Tests 2 and 3.

To choose six control stores with similar characteristics as our six treatment stores, we used the OTC sales data provided for all stores in the division, from weeks 14 to 23 of 2010 and 2011 combined with weeks 14 to 19 (the pre-treatment period) of 2012. To assess comparability of ex-ante generic share, we used a normalized measure of how generic share in a given store differed from the division-wide average by product-week. For five of the six treatment stores, the average deviation is positive, indicating that these stores had higher generic purchase rates than the average store in the division. We thus chose control stores exhibiting similar patterns. The normalized deviation from the average ranges from -0.02 to 1.49 among our treated stores, from -0.22 to 1.41 among our control stores, and from -1.43 to 3.76 among stores not selected for treatment or control.

As shown in the table below, two treatment stores and two control stores have ex-ante generic purchase shares that are more than one standard deviation larger than the average; three treatment stores and three control stores have generic shares less than one standard deviation above average, and one treatment store and one control stores have generic shares slightly below the average. We also sought to choose control stores with high weekly sales quantities, to reduce noise in generic share. For example, of six potential control stores with ex-ante generic shares roughly comparable to the two treated stores with the highest generic shares, we chose the two with the largest average sale quantities, which were also comparable to the sale quantities of our treated stores. Lastly, we chose five control stores with in-house pharmacies and one control store without one.

Store Assignment	Average generic share, standardized	Average weekly quantity, by product
Treatment: Test 1	1.49	12.79
Treatment: Test 3	1.45	10.40
Control	1.41	8.11
Control	1.08	14.61
Treatment: Test 2	0.78	10.13
Treatment: Test 3	0.74	6.21
Control	0.29	8.31
Control	0.29	11.24
Treatment: Test 2	0.09	5.55
Control	0.05	6.18
Treatment: Test 3	-0.02	6.10
Control	-0.22	13.85

Unassigned stores: Mean	0.15	7.44
Unassigned stores: Range	(-1.43, 3.76)	(2.13, 22.63)

Notes: The average generic share is the average, across all products and pre-intervention weeks, of the difference between the store's and the division-wide mean generic share by product-week, divided by the standard deviation of this store-level average.

We note that choosing specific control stores was necessary for the household-level analysis, since the retailer asked us to select a small number of stores for which to pull that data. For the store-level analysis, our analysis includes robustness checks using the entire set of untreated stores in the division. Also, as long as trends do not differ between the treatment and control stores, the difference-in-differences approach will eliminate any permanent differences across stores.

B. Consumer Survey

B.1. Survey overview

The survey was conducted in person in 3 separate stores in the same division as the treatment and control stores for the experimental analysis. All surveys were conducted on one Saturday. The survey enumerators asked questions verbally and recorded responses on a tablet. A $5 gift card to the retailer was offered as compensation for the respondents' time (approximately 5 minutes). Subjects were surveyed about either Advil/ibuprofen or Tylenol/acetaminophen, based on which one they reported using more often.

First, subjects were asked to report how long ago they purchased the name brand and how long ago they purchased a generic, either from this retailer or from any other retailer's store brand. Respondents who answered "Never" for purchasing a generic were categorized as never having tried the generic. Next, subjects were openly asked "Why do you buy ____ more often?"[1] and "How do you usually choose between brand and generic" and the enumerator coded whether their responses included any of several possible reasons generated in our pilot testing. Then, respondents were randomized to answer 2 of the following 3 questions: "What is your opinion on their taste?" "What is your opinion on how well they work to relieve pain?" and "What is your

[1] For this question, it was assumed that the version they reported purchasing more recently was the version that they purchase more often.

opinion of how safe the product is?" in random order. All questions were multiple choice, with possible answers indicating a preference for the brand, a preference for the generic, a belief that they are equivalent in this dimension, and "I don't know."

The hypothetical choice portion of the survey came next, and began by soliciting perceived price differences between brand and generic. Respondents were shown the tablet screen which first displayed an image of a 24-tablet package of the national brand painkiller alongside a same-sized package of the generic version. The typical price of one of the two was shown, and they were asked to guess the typical price of the other.

The same two packages were shown again, with standard prices shown for both.[2] The subjects were asked which of the two they would choose at these prices, with half of the subjects randomized to imagine "you have a terrible headache or other pain," and the other half told "you need to restock your supply of medicines at home." After making a choice at the regular prices, subjects were randomized to face a similar choice with one of the two items discounted from its typical price. Again, they were asked which one they would choose, holding constant the framing of the choice situation to either headache or restocking. We used the first choice, at the standard set of prices, to categorize subjects as choosing the brand or the generic.

Subjects were ask to think back to their original choice at the standard set of prices, and to guess the fraction of others who would make the same choice as them. With order randomized, they were asked to guess the percentage of pharmacists and the percentage of other shoppers at the same store making the same choice. After each of these guesses, they were told "Suppose the percentage of [other shoppers/pharmacists] who choose the [{brand} or {generic} depending on what subject chose] is only X%, meaning that (1-X)% of pharmacists choose the [{generic} or {brand} depending on what subject did not choose]. If you learned this information, would you still purchase the [{brand} or {generic} depending on what subject chose]?" The value X% was programmed to be 25 percentage points less than the value they had guessed, with a floor of 5%, so that the subject would be responding to a hypothetical signal that the product chosen by the subject is substantially less popular than they had guessed.

The last hypothetical choice section of the survey asked subjects to choose between two sodas. The choices was either between Coke and Pepsi, or between regular Coke, Diet Coke, and Coke

[2] A randomized subset saw the prices in dollar terms as well as in percentage terms, with either the brand price followed by "X% more than generic" or the generic price followed by "Y% less than brand," but this variation was found to have no effect on the choice.

Zero. Similar to the questions posed previously about the painkiller, respondents were asked to guess what share of other shoppers at this store would make the same choice as them, and then whether they would still make the same choice if they learned that a smaller share of other shoppers make that choice.

The survey then asked subjects whether they or any family member work in the healthcare-related occupations of nurse, physician or pharmacist. They were also asked whether more of their friends purchase brand or generic over the counter drugs and whether more of their friends purchase the different types of soda, with "I don't know" being a possible answer for each question. Finally, they were asked to confidentially enter responses categorizing their income, education, and race or ethnicity.

B.2 Main Survey Findings

Data was collected from 298 respondents. 34.9% were categorized as never having tried the generic, 10.4% had tried the generic but purchased the brand most recently, and 55% had purchased the generic more recently or equally recently as the brand.

Of people who report never having purchased a generic or store-brand version of their painkiller of choice, almost half believe that the national brand is superior either in efficacy (47%), safety (25%), or taste (21%). Another 22% report uncertainty about how the generic compares to the brand in either efficacy, safety or taste.[3] Of those who have tried the generic, smaller shares believed that either the brand or generic is superior in efficacy (11.6%), safety (10.3%), or taste (11.4%). Furthermore, as we would expect given that these subjects have tried both products, significantly fewer reported uncertainty in any of these comparisons (18.6% vs. 35.6%, $p<.001$).

Price guesses for the brand or generic were used to elicit beliefs about the price difference between the two. Results slightly differed here between the two painkillers. For a package of ibuprofen (national brand Advil) with 24 tablets, respondents underestimated the $2.20 difference in standard retail price by $0.35 on average ($p<.001$). For acetaminophen (national brand Tylenol) in a 100 count package, the standard price difference is $3.00 and the average guess was significantly larger at $4.36 ($p<.001$). This difference in results could be explained by people underestimating how price differences correlate with package size: in packages of larger quantity, the price difference between generics and brands tends to be smaller in percentage

[3] 34% total reported uncertainty on one of these three dimensions by choosing "Don't Know" or "Refuse to Answer: " 22% of respondents who have never tried the generic indicated uncertainty but no belief of superiority.

terms. The price guesses indicate that respondents expect generics to cost 34-37% of the comparable brand price, regardless of the package size. If we look collectively at respondents asked about 24-count ibuprofen and 100-count acetaminophen, we cannot reject the null hypothesis that perceived price differentials match the averaged true price differential (p=0.34).[4] Furthermore, we find no evidence that people who purchased the brand version most recently, or people who have never purchased the generic, make different estimates of the price differential.

65% of respondents chose the generic painkiller in the hypothetical choice at standard prices. However, there was a statistically significant effect of whether this hypothetical choice was framed as a situation of "headache or pain" (59% chose generic) versus "restocking" (70% chose generic, p=.034). When asked which they would choose if the other product was discounted (typical discounted price shown for either the brand or generic, depending on initial choice), 32% of those who had chosen the brand and 22% of those who had chosen the generic said that they would switch.

Guesses about Pharmacist and Shopper Choices

Priors about the share of consumers who buy the brand or generic were diffuse: 50% was the modal answer among those who chose the generic as well as those who chose the brand, accounting for 24% of all responses. The rest of the distribution was widespread, ranging from 5% to 100% (mean guess of the share of other customers choosing the generic = 55%, standard deviation = 20). As expected, we see 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). Interestingly, the guesses of those who choose the brand themselves are closer, on average, to the true proportions. That is, we find no evidence that the beliefs of brand-buying consumers are less accurate than those of generic-buying consumers.

We also see significant order effects, however, among the participants who choose the brand for themselves: If they were first asked to guess what share of pharmacists make the same choice as them, before being asked to guess the corresponding share of consumers, they guessed significantly higher shares of customers would choose the generic. This suggests that thinking about pharmacists' choices leads brand-buying consumers to focus on active ingredient

---

[4] In addition, we find that asking people to estimate the generic's price as a "percent less than" the brand price yields comparable estimates as soliciting a guess in dollar terms, but asking people to guess the brand drug's price as a "percent more than" the generic leads to lower implied price differentials.

comparability more than they would otherwise. Generic-buying consumers, by contrast, did not guess differently based on the order of these questions.

We find that customers who chose the generic were significantly more likely to say they would stick with this choice even if they learned that a smaller percentage of other customers made the same choice as them (91% vs. 81%, $p=.008$). Customers who initially chose the brand were more than twice as likely (19.4% vs. 8.7%) to say that they would "Probably not" or "Definitely not" still purchase the same product.

Taken together, these results suggest that informing people of the true generic shares among the store's customers could increase generic purchase rates because those who typically choose the brand are more likely to be swayed by this information than those who typically choose the generic, but not because those who typically choose the brand are further off with their initial priors.

Appendix C. Survey Questions

C.1 Consumer Survey

Respondents were first asked whether they use Advil/ibuprofen or Tylenol/acematiminophen more regularly, and then given a survey focused on the brand vs. generic choice of the painkiller they use. Below, the questions for the Advil/ibuprofen survey are shown.

1. When was the last time you purchased Advil®-brand ibuprofen, either at [*retailer*] or another retailer?

- ○ Less than 3 months ago
- ○ 3-6 months ago
- ○ 6 months- 1 year ago
- ○ More than 1 year ago
- ○ Never

2. When was the last time you purchased any generic (store-brand) ibuprofen, either at [*retailer*] or another retailer?

- ○ Less than 3 months ago
- ○ 3-6 months ago
- ○ 6 months- 1 year ago
- ○ More than 1 year ago
- ○ Never

3. Why do you buy __________ more often?

__[free response]____

Randomized to receive 2 of the following 3 questions, for the purpose of reducing time burden.

4. If you have tried both, what is your opinion on how well they work to relieve pain?

○ Brand-name works better ○ They work the same ○ Generic works better

5. If you have tried both, what is your opinion on their taste or ease of swallowing?

○ I prefer brand-name ○ They seem the same ○ I prefer generic

6. What is your opinion of how safe the product is?

○ Brand-name is safer ○ They are equally safe ○ Generic is safer

7. Why do you think ______________ is more safe?

____

____

9. Here is a picture of two packages of ibuprofen. One of these packages has the typical price shown beneath it. Please make your best guess of the price of the other one.

[*randomize whether brand or generic has price shown*]

$5.49	**$__.___**
or	
$__.___	**$3.29**

[*randomize the framing of the Question 10*]

10a. *[Framing 1]* Suppose you have a terrible headache. Which are you more likely to purchase?
10b. *[Framing 2]* Suppose you need to restock your supply of medicines at home. Which would you prefer to purchase?

◯ Advil, $5.49

◯ [redacted] ibuprofen, $3.29

11a. Please guess the percentage of **shoppers at this store** who make the same choice as you.

_______ %

11b. Please guess the percentage of **pharmacists** who make the same choice as you.

_______ %

To those who chose the brand:

12a. Suppose the percentage of <u>shoppers</u> who choose the *[store]*-label ibuprofen is _____ *[10 or 25 percentage points larger than guess indicates, at random, capped at 95%]*, meaning that only ____ choose Advil-brand ibuprofen. If you learned this information, would you still purchase the brand?

○ Definitely not ○ Probably not ○ Probably yes ○ Definitely yes

13a. Suppose the percentage of <u>pharmacists</u> who choose the generic ibuprofen is _____ *[10 or 25 percentage points larger than guess indicates, at random, capped at 95%],* meaning that only ____ choose Advil-brand ibuprofen. If you learned this information, would you still purchase the brand?

○ Definitely not ○ Probably not ○ Probably yes ○ Definitely yes

To those who chose the generic:

12b. Suppose the percentage of <u>shoppers</u> who choose Advil-brand ibuprofen is _____ *[10 or 25 percentage points larger than guess indicates, at random, capped at 95%]*, meaning that only ____ choose *[store]*-label ibuprofen. If you learned this information, would you still purchase the generic?

○ Definitely not ○ Probably not ○ Probably yes ○ Definitely yes

12b. Suppose the percentage of <u>pharmacists</u> who choose Advil brand ibuprofen is _____ *[10 or 25 percentage points larger than guess indicates, at random, capped at 95%]*, meaning that only ____ choose *[store]*-label ibuprofen. If you learned this information, would you still purchase the generic?

○ Definitely not ○ Probably not ○ Probably yes ○ Definitely yes

13. Do more of your close friends purchase brand or generic over-the-counter painkillers?

○ More buy Brand drugs ○ More buy Generic ○ An even mix. ○ I don't know what kind they buy.

14a. If offered a free soda, which would you choose?

○ Coca-Cola ○ Pepsi ○ None

[*or*] 14b. If offered a free soda, which would you choose?

○ Coca-Cola ○ Diet Coke ○ Coke Zero ○ None

[Questions 11a, 12, and 13 repeated for the same soft drink choices shown to each respondent.]

15. Do you work as a health-related professional? (e.g. as a physician, nurse, or pharmacist?)

○ Yes ○ No

16. Do you have a family member who works as a health-related professional?

○ Yes ○ No

Questions 17-19 solicited self-reported highest level of education, earnings, and race.

C.2. Pharmacist Survey

Survey Questions for Pharmacist Managers at In-store Pharmacies

1. About how many customers, on the average day, solicit pharmacist advice regarding the choice or use of over-the-counter (OTC) drugs?

   Choose one: 0-5 5-10 10-15 15-20 20 or more.

2. Of all the customers soliciting pharmacist advice about OTC drugs, what percentage of them express concern about costs, or a desire for an inexpensive drug?

   Choose one: 0-10% 10-20% 20-30% 30-40% 40-50% 50% or more.

3. Of all the customers soliciting pharmacist advice about OTC drugs, how many ask specifically about the quality or comparability of generic OTC drugs?

   Choose one: 0-20% 20-40% 40%-60% 60%-80% 80% or more

4. If a customer specifically asks about the quality or comparability of generic OTC drugs, what is the pharmacist's usual response?

   Please write down response: ____

   ____

5. Would this response be the same regardless of the category of over-the-counter drugs being considered?

   Choose one: Yes No

6. In your view, what are the primary reasons why a customer might choose to purchase a national-brand OTC drug instead of a lower-priced generic alternative?
   (List in order of importance.)

   1.____

   2.____

   3.____

7. If a customer solicits pharmacist advice regarding the choice between various OTC drugs, but does not specifically ask about generic OTC drugs, how often does the pharmacist point to the generic OTC medications as a low-cost, but equally effective option?

   Choose one: Never/Rarely Sometimes Usually Always

D. Appendix Figures and Tables

We assigned a fixed set of OTC drug classes to be treated using the process below. The same drug classes were treated at all stores receiving treatment (see flow diagram for store choice in Fig. A2). Treatment consisted of informational labels posted at the point of purchase.

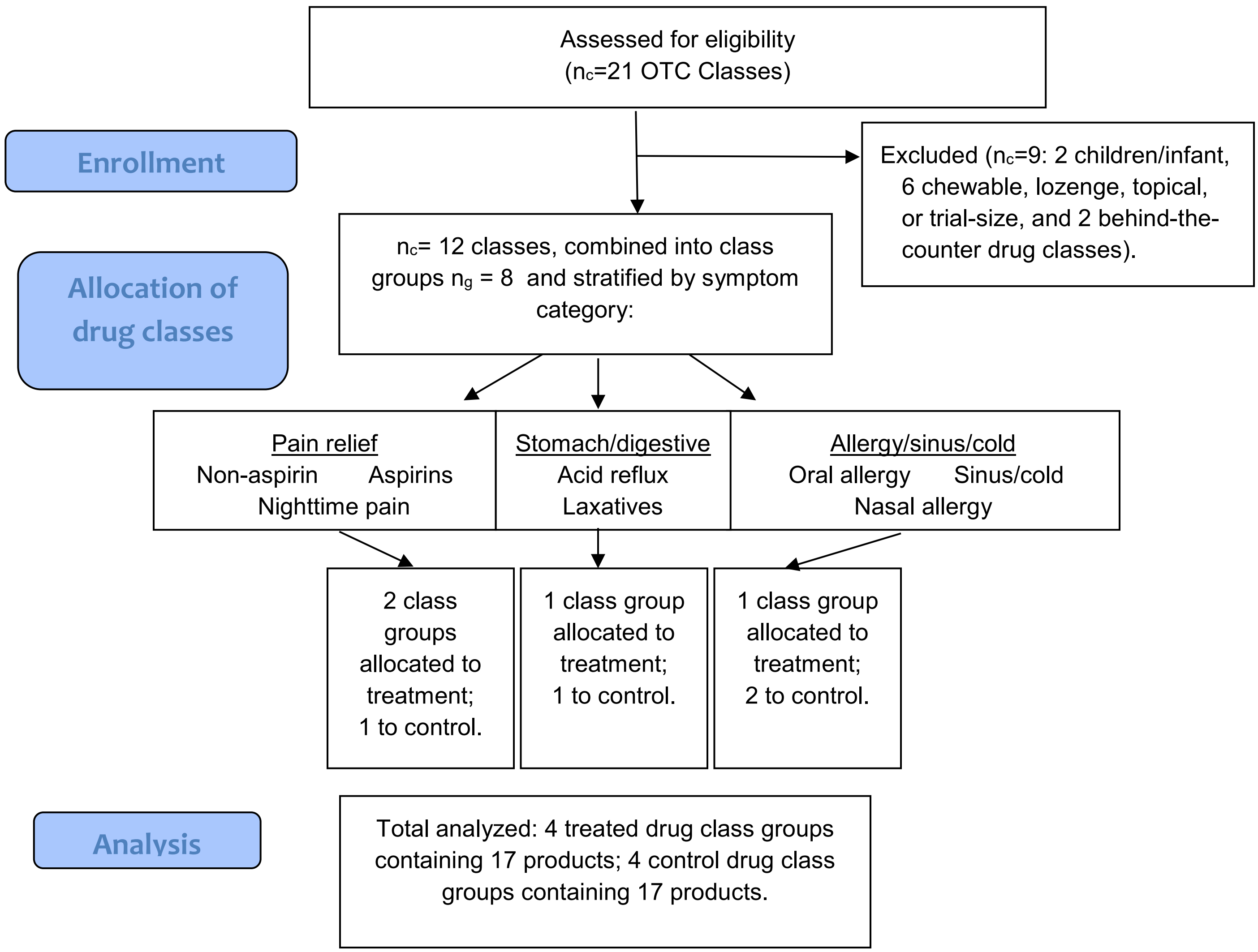

Figure A1. Flow Diagram, Over the Counter Drug (OTC) Experimental Design

Six treatment stores were chosen by convenience and randomly assigned to one of five types of labeling treatments, using the process shown below. The five different labeling treatments included three types of information, two of which had two framing variations. Using OTC sales data from the previous year, we identified six similar stores to use as control stores.

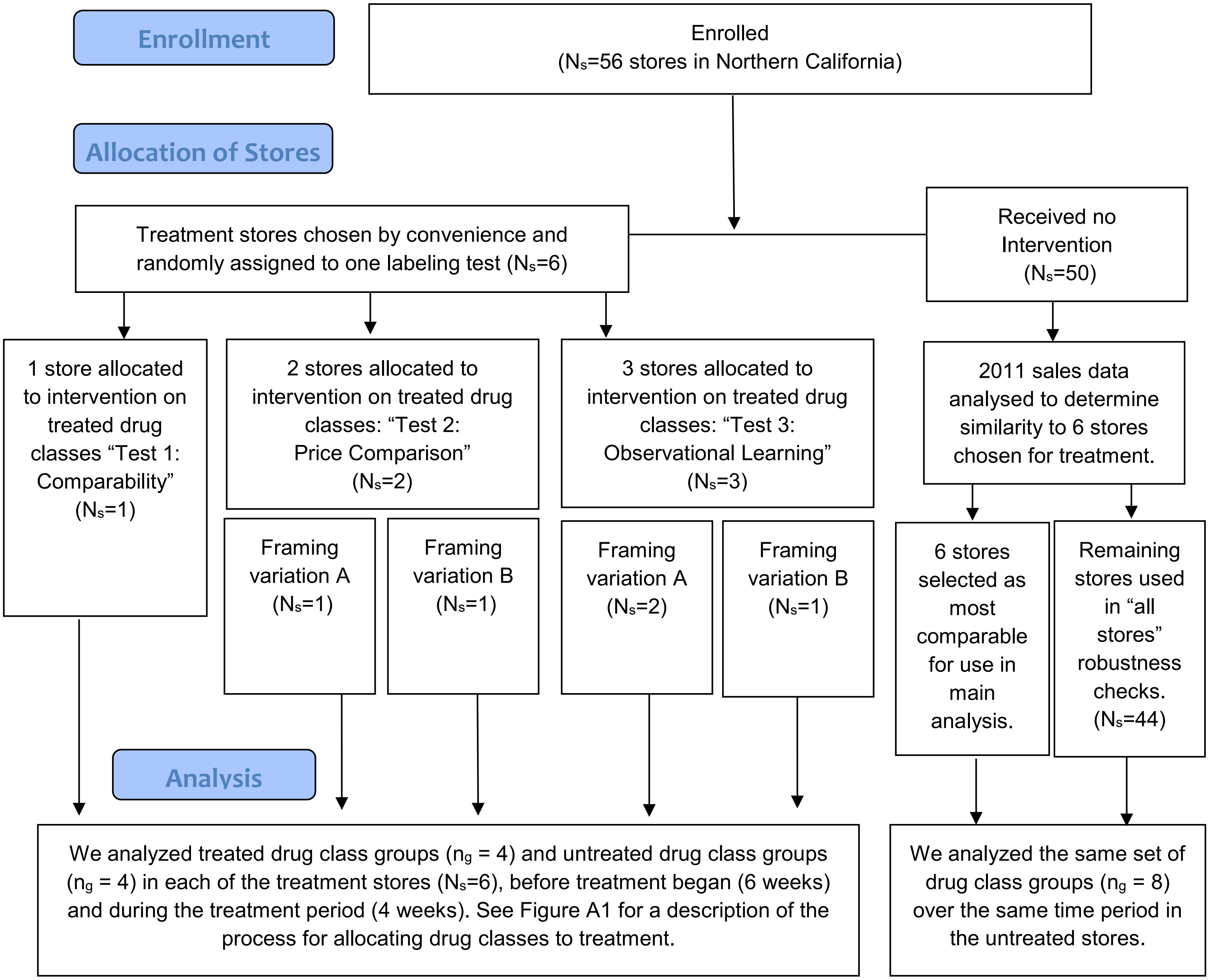

Figure A2. Flow Diagram, Over the Counter Drug (OTC) Experimental Design

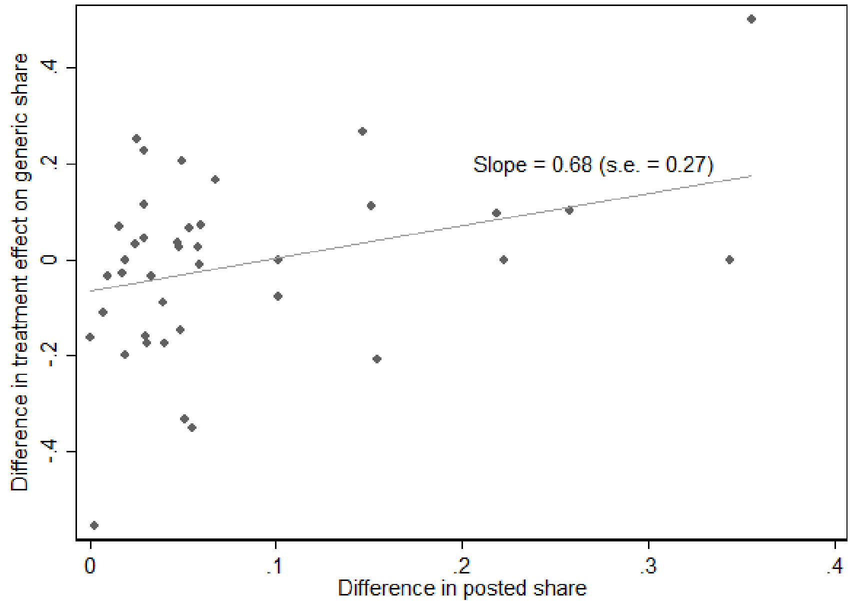

Figure A3. Effect of posted shares of prior customers on current generic shares

Notes: As part of the treatment for stores posting shares of prior customers choosing the generic, the posted amount was varied across the weeks of treatment. In weeks 1 and 3, the posted share was calculated using a different prior year of purchases than in weeks 2 and 4. The points plotted each represent one treated product in one of the stores receiving this treatment. For each product, we calculated the difference between the estimated treatment effect on generic share between weeks 1&3 and weeks 2&4, and plotted this difference against the difference in the posted share of prior customers who had purchased the generic. The linear fit plotted weights each observation by quantity and the regression is shown in Table A

Appendix Table A1. Summary Statistics of OTC Product Classes

	Untreated products	Difference (Treated - Untreated)
Panel A: Pain relief		
Brand price per unit	0.18	0.04
	(0.02)	(0.06)
Generic price, as a share of brand price	0.55	0.05
	(0.06)	(0.08)
Weekly quantity sold per product	30.83	-23.28
	(9.53)	(9.78)
Generic share	0.43	0.05
	(0.11)	(0.13)
N (product x store x week observations)		648
N (unique products, untreated)		3
N (unique products, treated)		6
Panel B: Allergy/cold symptoms relief		
Brand price per unit	0.37	0.34
	(0.15)	(0.21)
Generic price, as a share of brand price	0.48	0.13
	(0.11)	(0.12)
Weekly quantity sold per product	6.18	13.25
	(1.96)	(4.53)
Generic share	0.58	-0.15
	(0.07)	(0.10)
N (observations)		576
N (unique products, untreated)		4
N (unique products, treated)		4
Panel C: Digestive/stomach symptoms relief		
Brand price per unit	0.52	0.00
	(0.01)	(0.01)
Generic price, as a share of brand price	0.58	0.08
	(0.01)	(0.01)
Weekly quantity sold per product	4.39	2.03
	(0.49)	(0.82)
Generic share	0.39	0.01
	(0.02)	(0.03)
N (observations)		576
N (unique products, untreated)		5
N (unique products, treated)		3

Notes: Pre-treatment period of six weeks. Standard errors, clustered by product, are shown beneath the means. We cannot share these statistics broken down by category, nor identify the specific products that were treated, because it is considered proprietary information by the retailer.

Appendix Table A2. Descriptive Statistics, Pre-treatment period

Panel A: Treated Products

		Differences relative to Control stores		
	Control stores	Test 1 stores	Test 2 stores	Test 3 stores
Brand price per unit	0.44	-0.00	0.00	-0.00
	(0.08)	(0.00)	(0.00)	(0.00)
Generic price, as a share of brand price	0.62	-0.00	0.00	0.01
	(0.04)	(0.01)	(0.01)	(0.01)
Weekly quantity sold per product	12.06	2.34	-3.09	-3.18
	(2.70)	(1.23)	(0.87)	(1.48)
Generic share	0.47	0.01	-0.07	-0.04
	(0.05)	(0.03)	(0.04)	(0.03)

Panel B: Untreated Products

		Differences relative to Control stores		
	Control stores	Test 1 stores	Test 2 stores	Test 3 stores
Brand price per unit	0.38	0.03	0.01	0.00
	(0.08)	(0.03)	(0.01)	(0.01)
Generic price, as a share of brand price	0.54	-0.02	-0.01	0.02
	(0.04)	(0.03)	(0.02)	(0.01)
Weekly quantity sold per product	12.06	5.26	-2.17	-2.17
	(4.14)	(2.48)	(1.16)	(0.60)
Generic share	0.47	0.01	-0.064	0.03
	(0.06)	(0.03)	(0.034)	(0.03)

Notes: Standard errors, clustered at the drug class-by-store level, are in parentheses. Sample (N=936 for treated products, N=864 for untreated products) includes 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, pill type or inactive ingredients). Prices are in dollars per unit, inclusive of discounts, averaged over the different UPCs sold for each active ingredient, weighted by purchase share, 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.

Appendix Table A3. Difference in generic share purchased

	(1)	(2)
Difference in share presented	0.84	0.68
as % choosing generic	(0.36)	(0.23)
Constant	-0.077	-0.064
	(0.053)	(0.034)
Weighted by quantity	No	Yes
N	39	39
R^2	0.159	0.101

Notes: Sample includes treated products within 3 stores treated with Test 3 labels, which showed the share of prior customers choosing the generic of that product. The posted share was varied across the weeks of treatment. In weeks 1 and 3, the posted share was calculated using a different prior year of purchases than in weeks 2 and 4 (either 2011, or first 12 weeks of 2012, depending on the store). For each treated product within each treated store, we calculated the difference between the estimated treatment effect on generic share between weeks 1&3 and weeks 2&4, and regressed this difference on the difference in the shares posted between these periods. Clustered standard errors at the drug class-by-store level are in parentheses.

Appendix Table A4: Household-Level Transition Probabilities between Brand and Generic

Panel A.	Second observed purchase: Brand	Second observed purchase: Generic	N (first purchases)	Percent
First observed purchase: Brand	6,489 87.3%	947 12.7%	7,434 100%	54.5%
First observed purchase: Generic	947 15.3%	5,257 84.7%	6,206 100%	45.5%
N (second purchases)	7,437 54.5%	6,203 45.5%	13,640 100%	100%

Panel B.	Third observed purchase: Brand	Third observed purchase: Generic	N (second purchases)	Percent
Second observed purchase: Brand	2571 90.30%	274 9.60%	2,845 100%	54.2%
Second observed purchase: Generic	286 11.9%	2,114 88.1%	2400 100.0%	45.8%
N (third purchases)	2,857 54.5%	2,388 45.5%	5,245 100%	100%

Notes: Household-level dataset used; observations from prior to the pre-intervention period only. Panel A includes all household-drug combinations with at least two purchases during this period. Panel B includes all household-drug combinations with at least three purchases during this period.

Appendix Table A5: Household level results by previous OTC purchases and label content, framing

	Treated Products				Untreated Products			
	HH with no previous OTC purchases		HH with 1+ previous OTC purchases		HH with no previous OTC purchases		HH with 1+ previous OTC purchases	
Y = Generic purchased	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Test 1: Comparability Statement	0.083		-0.011		0.046	0.047	-0.009	-0.009
	(0.032)		(0.026)		(0.050)	(0.050)	(0.027)	(0.027)
	{0.28}		{0.73}		{0.88}	{0.87}	{0.76}	{0.78}
	[-0.118, 0.331]		[-0.088, 0.043]		[-0.893, 0.955]	[-0.745, 1.136]	[-0.404, 0.333]	[-0.451, 0.332]
a. "Same active ingredient"		0.046		-0.013				
		(0.043)		(0.029)				
		{0.54}		{0.72}				
		[-0.605, 0.471]		[-0.349, 0.266]				
b. "… and approved by the FDA"		0.086		-0.037				
		(0.063)		(0.029)				
		{0.56}		{0.33}				
		[-1.092, 1.273]		[-0.609, 0.471]				
c. "FDA determined bioequivalence"		0.152		0.022				
		(0.042)		(0.031)				
		{0.23}		{0.66}				
		[-0.230, 0.757]		[-0.357, 0.476]				
Test 2: Price Comparison	0.026		0.065		0.020		-0.000	
	(0.021)		(0.031)		(0.042)		(0.023)	
	{0.22}		{0.06}		{0.63}		{0.98}	
	[-0.041, 0.116]		[-0.004, 0.161]		[-0.164, 0.110]		[-0.128, 0.037]	
a. Framing: Save X%		0.049		0.086		0.027		-0.004
		(0.026)		(0.040)		(0.060)		(0.032)
		{0.06}		{0.06}		{0.64}		{0.89}
		[-0.207, 0.243]		[-0.240, 0.398]		[-0.366, 0.926]		[-0.437, 0.331]
b. Framing: Pay Y% More		-0.016		0.024		0.006		0.007
		(0.046)		(0.046)		(0.019)		(0.025)
		{0.83}		{0.89}		{0.78}		{0.75}
		[-0.283, 0.466]		[-0.274, 0.218]		[-0.103, 0.212]		[-0.123, 0.416]
Test 3: Observational Learning	0.073		0.074		0.051		-0.003	
	(0.021)		(0.015)		(0.016)		(0.020)	
	{0.004}		{0.002}		{0.03}		{0.87}	
	[0.031, 0.135]		[0.038, 0.109]		[0.015, 0.123]		[-0.059, 0.041]	
a. Framing: X% choose generic		0.071		0.075		0.055		-0.002
		(0.028)		(0.014)		(0.022)		(0.024)
		{0.06}		{0.002}		{0.12}		{0.93}
		[-0.006, 0.183]		[0.037, 0.107]		[-0.060, 0.187]		[-0.106, 0.044]
b. Framing: Y% choose brand		0.077		0.073		0.043		-0.005
		(0.033)		(0.039)		(0.016)		(0.021)
		{0.07}		{0.25}		{0.14}		{0.78}
		[-0.343, 0.329]		[-0.230, 0.160]		[-0.265, 0.152]		[-0.358, 0.277]
N	6082	6082	8809	8809	6355	6355	8256	8256
N, drug class X store clusters	48	48	48	48	48	48	48	48
Dependent Variable Mean	0.44	0.44	0.45	0.45	0.44	0.44	0.48	0.48
Tests of equality, *p*-values								
H_0: Test 1 = Test 2	0.24		0.06		0.8		0.85	
H_0: Test 1 = Test 3	0.83		0.01		0.94		0.86	
H_0: Test 2 = Test 3	0.16		0.78		0.54		0.94	
H_0: Test 1 = Test 2 = Test 3	0.38		0.04		0.89		0.97	
H_0: Test 1a = Test 1b = Test 1c		0.20		0.22				
H_0: Test 2a = Test 2b		0.20		0.40		0.72		0.81
H_0: Test 3a = Test 3b		0.90		0.97		0.67		0.93

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. "Test 1," "Test 2," and "Test 3" treatment indicators are interactions for treated store and treatment time period. 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. The specifications in Columns 3 and 7 match Columns 1 and 4 in Table 5 but are replicated here to ease comparison between households with and without previous OTC purchases during the year prior to the start of the labeling tests. Standard errors are clustered at the drug class-by-store level and are not adjusted for multiple hypothesis testing. Braces and square brackets below contain p-values and 95% confidence intervals from a wild cluster bootstrap procedure (Stata boottest, 2000 replications, Webb weights), which is also used for the tests of coefficient equality. Significance stars are omitted.

Appendix Table A6: Post-treatment effects, Difference-i

Y = Generic purchased	Treated Products				
	Weeks 24-27		Weeks 28-31		Weeks
	(1)	(2)	(3)	(4)	(5)
Test 1 x Post x Shopped During		0.056		-0.051	
		(0.55)		(0.45)	
		[-0.133, 0.286]		[-0.182, 0.107]	
Test 2 x Post x Shopped During		-0.042		-0.051	
		(0.67)		(0.35)	
		[-0.195, 0.074]		[-0.161, 0.069]	
Test 3 x Post x Shopped During		0.056		0.011	
		(0.17)		(0.83)	
		[-0.029, 0.145]		[-0.119, 0.127]	
Test 1 x Post	0.036	0.018	0.045	0.064	0.008
	(0.32)	(0.70)	(0.14)	(0.10)	(0.80)
	[-0.040, 0.120]	[-0.095, 0.117]	[-0.020, 0.106]	[-0.018, 0.132]	[-0.051, 0.084]
Test 2 x Post	-0.036	-0.023	0.010	0.028	-0.006
	(0.12)	(0.40)	(0.66)	(0.41)	(0.89)
	[-0.085, 0.013]	[-0.081, 0.043]	[-0.042, 0.063]	[-0.054, 0.100]	[-0.094, 0.081]
Test 3 x Post	0.019	0.001	0.026	0.024	0.012
	(0.52)	(0.98)	(0.39)	(0.56)	(0.60)
	[-0.044, 0.085]	[-0.083, 0.073]	[-0.037, 0.093]	[-0.067, 0.124]	[-0.041, 0.055]
Post x Shopped During		0.005		0.035	
(i.e. exposed)		(0.85)		(0.18)	
		[-0.045, 0.059]		[-0.022, 0.081]	
Test 1 stores x Shopped During		-0.009		-0.008	
		(0.86)		(0.87)	
		[-0.131, 0.095]		[-0.126, 0.098]	
Test 2 stores x Shopped During		0.024		0.022	
		(0.72)		(0.73)	
		[-0.098, 0.148]		[-0.093, 0.159]	
Test 3 stores x Shopped During		-0.005		-0.002	
		(0.92)		(0.97)	
		[-0.101, 0.103]		[-0.105, 0.096]	
Shopped During Treatment Time		-0.029		-0.025	
		(0.06)		(0.08)	
		[-0.058, 0.001]		[-0.054, 0.005]	
N	8137	8137	8538	8538	9807
N, drug class X store clusters	0.45	0.45	0.45	0.45	0.45
Dependent variable mean	13	13	13	13	13

Notes: Linear probability models for the choice of a generic. Observations represent individual purchases of any treated or untreated dı
treated with labeling test X and the specified post-treatment time period, capturing the difference-in-differences in generic purchase sha
untreated stores. "Shopped During" is an indicator for the individual having made any purchase of a treated or untreated OTC drug duı
"Exposed" indicator in Table 6, for treated store purchases. The interactions "Test X x Shopped During" capture the average differenc
Parentheses show p-values and square brackets contain 95% confidence intervals from a wild cluster bootstrap procedure (Stata bootte

n-Differences between Treated and Untreated Stores

	Untreated Products					
32-36	Weeks 24-27		Weeks 28-31		Weeks 32-36	
(6)	(7)	(8)	(9)	(10)	(11)	(12)
-0.005		0.053		0.044		0.058
(0.93)		(0.17)		(0.34)		(0.20)
[-0.136, 0.155]		[-0.042, 0.157]		[-0.093, 0.208]		[-0.050, 0.193]
-0.031		-0.007		-0.054		-0.004
(0.58)		(0.91)		(0.44)		(0.92)
[-0.157, 0.098]		[-0.106, 0.141]		[-0.224, 0.072]		[-0.134, 0.089]
-0.006		0.014		0.015		0.066
(0.86)		(0.70)		(0.71)		(0.13)
[-0.088, 0.067]		[-0.101, 0.112]		[-0.112, 0.088]		[-0.020, 0.130]
0.011	0.005	-0.014	0.045	0.029	0.032	0.013
(0.76)	(0.85)	(0.60)	(0.04)	(0.22)	(0.08)	(0.63)
[-0.071, 0.088]	[-0.059, 0.054]	[-0.114, 0.028]	[0.003, 0.115]	[-0.067, 0.094]	[-0.010, 0.083]	[-0.069, 0.090]
0.005	0.027	0.026	0.020	0.038	0.017	0.015
(0.88)	(0.50)	(0.68)	(0.41)	(0.25)	(0.40)	(0.60)
[-0.081, 0.086]	[-0.050, 0.078]	[-0.071, 0.084]	[-0.038, 0.096]	[-0.025, 0.119]	[-0.022, 0.077]	[-0.039, 0.098]
0.015	0.036	0.030	0.040	0.035	0.011	-0.012
(0.51)	(0.14)	(0.29)	(0.04)	(0.15)	(0.63)	(0.62)
[-0.030, 0.065]	[-0.030, 0.108]	[-0.048, 0.111]	[0.005, 0.093]	[-0.008, 0.122]	[-0.063, 0.081]	[-0.091, 0.063]
0.038		-0.073		-0.042		-0.049
(0.11)		(0.00)		(0.12)		(0.01)
[-0.012, 0.082]		[-0.112, -0.051]		[-0.083, 0.026]		[-0.094, -0.020]
-0.005		-0.020		-0.019		-0.021
(0.93)		(0.65)		(0.64)		(0.65)
[-0.126, 0.098]		[-0.125, 0.092]		[-0.125, 0.095]		[-0.127, 0.092]
0.021		-0.005		-0.003		-0.001
(0.74)		(0.88)		(0.93)		(0.98)
[-0.094, 0.155]		[-0.090, 0.081]		[-0.083, 0.079]		[-0.083, 0.086]
-0.000		-0.016		-0.010		-0.016
(0.99)		(0.74)		(0.82)		(0.70)
[-0.098, 0.100]		[-0.092, 0.127]		[-0.083, 0.128]		[-0.087, 0.120]
-0.031		-0.001		-0.000		-0.003
(0.05)		(0.95)		(0.99)		(0.85)
[-0.058, -0.001]		[-0.054, 0.025]		[-0.047, 0.023]		[-0.049, 0.019]
9807	8028	8028	8763	8763	10392	10392
0.45	0.48	0.48	0.48	0.48	0.48	0.48
13	12	12	12	12	12	12

rug during either the specified post-period or the pre-treatment period. For each test, "Test X x Post" is an interaction for a store
are overall between the pre-treatment period and the specified post-treatment period for the stores that received treatment X versus
ring the treatment time period, indicating their presence in the OTC aisles of the store. "Shopped During" is equivalent to the
e between generic purchase rates between exposed and unexposed customers within Test X store(s) in the pre-treatment period.
st, 2000 replications, Webb weights). Significance stars are omitted.

Appendix Table A7: Robustness Checks for Treatment Effects on Quantity

	Serial Correlation Correction				Poisson Model			
	Treated Products		Untreated Products		Treated Products		Untreated Products	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Test 1: Comparability Statement	1.21	1.21	0.46	0.46	0.096	0.096	0.040	0.040
	(0.85)	(0.85)	(0.87)	(0.87)	(0.061)	(0.061)	(0.10)	(0.10)
Test 2: Price Comparison	1.28		0.64		0.10		0.056	
	(0.43)		(0.26)		(0.061)		(0.046)	
Framing: Save X%		1.68		0.53		0.13		0.048
		(0.63)		(0.62)		(0.068)		(0.060)
Framing: Pay Y% More		0.86		0.76		0.046		0.075
		(0.50)		(0.43)		(0.090)		(0.048)
Test 3: Observational Learning	1.30		0.69		0.11		0.055	
	(0.92)		(0.18)		(0.050)		(0.037)	
Framing: X% choose generic		1.49		0.62		0.13		0.048
		(0.79)		(0.21)		(0.056)		(0.043)
Framing: Y% choose brand		0.92		0.83		0.075		0.075
		(1.19)		(0.41)		(0.063)		(0.057)
N	1556	1556	1438	1438	1556	1556	1438	1438
Dependent variable mean	10.9	10.9	11.6	11.6	10.9	10.9	11.6	11.6

Notes: The odd numbered columns match the specifications of Table 3 and the even numbered columns match the specifications of Table 4, with Driscoll-Kraay standard errors shown for the serial correlation correction using Stata xtscc command. The Poisson model is implemented with conditional fixed effects at the store-by-product level using xtpoisson. "Generic on promotion" and "Brand on promotion" are included, as are store, product, and time period fixed effects, in all specifications. Robust standard errors are in parentheses and the Poisson point estimates are interpreted as percent effects. Significance stars omitted.

Appendix Table A8: Treatment Effects on Generic Share in Treated (2012) and Placebo Years (2011, 2010)
Wild cluster bootstrapped p-values and 95% confidence intervals

Panel A. Six selected control stores

Y = Generic share	Treated Products			Untreated Products		
	2012	2011	2010	2012	2011	2010
Test 1: Comparability	0.009	-0.060	-0.005	-0.008	0.046	0.064
Statement	(0.86)	(0.19)	(0.83)	(0.85)	(0.62)	(0.46)
	[-0.129, 0.075]	[-0.190, 0.055]	[-0.098, 0.069]	[-0.200, 0.069]	[-0.318, 0.103]	[-0.043, 0.156]
Test 2: Price comparison	0.047	-0.051	0.031	0.008	-0.009	0.042
	(0.08)	(0.11)	(0.39)	(0.78)	(0.66)	(0.49)
	[-0.007, 0.124]	[-0.100, 0.020]	[-0.049, 0.088]	[-0.084, 0.051]	[-0.059, 0.055]	[-0.055, 0.134]
Test 3: Observational	0.060	-0.006	0.025	0.018	0.017	-0.056
	(0.02)	(0.83)	(0.24)	(0.46)	(0.61)	(0.13)
	[0.008, 0.109]	[-0.070, 0.045]	[-0.019, 0.067]	[-0.029, 0.065]	[-0.060, 0.079]	[-0.144, 0.018]
Weighted by quantity	Yes	Yes	Yes	Yes	Yes	Yes
N	1533	1400	1284	1389	1394	1272
N, clusters	48	48	44	48	48	44
Dependent variable mean	0.45	0.44	0.40	0.46	0.44	0.40

Panel B. All 50 district stores included as control stores

Y = Generic share	Treated Products			Untreated Products		
	2012	2011	2010	2012	2011	2010
Test 1: Comparability	-0.002	-0.022	-0.006	-0.007	0.032	0.087
Statement	(0.952)	(0.424)	(0.828)	(0.879)	(0.747)	(0.104)
	[-0.211, 0.278]	[-0.158, 0.097]	[-0.108, 0.053]	[-0.362, 0.435]	[-0.425, 0.548]	[-0.314, 0.481]
Test 2: Price comparison	0.037	-0.009	0.031	0.009	-0.027	0.068
	(0.126)	(0.680)	(0.454)	(0.706)	(0.206)	(0.181)
	[-0.026, 0.134]	[-0.049, 0.062]	[-0.055, 0.082]	[-0.083, 0.046]	[-0.104, 0.053]	[-0.035, 0.157]
Test 3: Observational	0.047	0.036	0.025	0.019	0.000	-0.020
Learning	(0.056)	(0.152)	(0.203)	(0.345)	(0.999)	(0.488)
	[-0.002, 0.087]	[-0.029, 0.074]	[-0.022, 0.058]	[-0.027, 0.060]	[-0.106, 0.069]	[-0.101, 0.036]
Weighted by quantity	Yes	Yes	Yes	Yes	Yes	Yes
N	6820	6329	6149	6114	6164	5983
N, clusters	224	224	220	224	224	220
Dependent variable mean	0.43	0.43	0.40	0.43	0.43	0.39

Notes: Observations are at the week-store-drug level. Test 1 was conducted at one store, test 2 was conducted at two stores, and test 3 was conducted at three stores. Store, product, and time period fixed effects are included in all specifications, as well as indicators for generic, brand, or both types of products being on price promotion. Parentheses contain p-values and brackets contain 95% confidence intervals based on wild-cluster bootstrapping (Stata boottest) with drug class-by-store clusters. Since one of the control stores did not exist in 2010, the number of clusters is smaller that year. Significance stars omitted.

Appendix Table A9: Treatment Effects on Total Quantity in Treated (2012) and Placebo Years (2011, 2010)
Wild cluster bootstrapped p-values and 95% confidence intervals

Panel A. Six selected control stores

Y = Quantity sold	Treated Products			Untreated Products		
	2012	2011	2010	2012	2011	2010
Test 1: Comparability	1.09	0.17	-0.40	0.40	-0.42	-1.90
Statement	(0.33)	(0.96)	(0.67)	(0.84)	(0.58)	(0.18)
	[-1.38, 4.87]	[-8.03, 5.36]	[-2.30, 2.61]	[-1.75, 4.31]	[-2.53, 1.49]	[-5.87, 0.84]
Test 2: Price comparison	1.18	1.23	1.26	0.58	-0.09	-0.19
	(0.18)	(0.37)	(0.11)	(0.42)	(0.85)	(0.73)
	[-0.54, 2.97]	[-1.59, 4.05]	[-0.25, 2.99]	[-0.99, 1.88]	[-1.16, 0.91]	[-1.34, 0.87]
Test 3: Observational	1.30	0.96	1.30	0.64	0.13	-0.27
Learning	(0.10)	(0.48)	(0.10)	(0.14)	(0.77)	(0.64)
	[-0.19, 2.84]	[-1.73, 3.72]	[-0.18, 2.85]	[-0.22, 1.52]	[-0.72, 0.99]	[-1.39, 0.93]
N	1556	1440	1320	1438	1440	1318
N, clusters	48	48	44	48	48	44
Dependent variable mean	10.94	10.33	10.74	11.60	11.41	12.03

Panel B. All 50 district stores included as control stores

Y = Generic share	Treated Products			Untreated Products		
	2012	2011	2010	2012	2011	2010
Test 1: Comparability	0.87	-0.21	-1.40	0.06	-0.50	-1.86
Statement	(0.394)	(0.954)	(0.161)	(0.965)	(0.479)	(0.200)
	[-2.49, 4.91]	[-7.93, 15.98]	[-2.71, 1.62]	[-6.85, 4.07]	[-3.98, 1.04]	[-6.63, 9.21]
Test 2: Price comparison	1.06	0.81	0.31	0.26	0.01	-0.05
	(0.087)	(0.310)	(0.476)	(0.687)	(0.984)	(0.869)
	[-0.23, 2.51]	[-0.95, 2.41]	[-0.57, 1.42]	[-1.28, 1.48]	[-1.17, 0.76]	[-0.97, 0.76]
Test 3: Observational	1.15	0.50	0.23	0.26	0.24	-0.21
Learning	(0.006)	(0.580)	(0.620)	(0.451)	(0.284)	(0.662)
	[0.44, 1.88]	[-1.43, 2.31]	[-0.66, 1.19]	[-0.52, 1.05]	[-0.21, 0.68]	[-1.17, 0.96]
N	7225	6708	6558	6623	6713	6557
N, clusters	224	224	220	224	224	220
Dependent variable mean	9.12	8.53	8.54	9.23	9.18	9.24

Notes: Observations are at the week-store-drug level. Quantity is total products purchased of both brand and generic versions. Test 1 was conducted at one store, test 2 was conducted at two stores, and test 3 was conducted at three stores. Store, product, and time period fixed effects are included in all specifications, as well as indicators for generic, brand, or both types of products being on price promotion. Parentheses and brackets contain *p*- values and 95% confidence intervals based on wild-cluster bootstrapping (Stata *boottest*) with drug class-by-store clusters. Since one of the control stores did not exist in 2010, the number of clusters is smaller that year. Significance stars omitted.

Appendix Table A10: Consumer Survey Responses

Panel A.	Overall	Grouped by past reported purchases	
		Never tried generic	Tried generic
What is your opinion on how well they work to relieve pain?			
Brand-name works better	22%	46%	9%
Generic works better	2%	0%	3%
They work the same	67%	38%	82%
Don't know / Refused to answer	10%	15%	6%
N	200	71	129
What is your opinion of how safe the product is?			
Brand-name is safer	15%	25%	10%
Generic is safer	1%	0%	1%
They are equally safe	76%	60%	84%
Don't know / Refused to answer	9%	15%	6%
N	191	65	126
What is your opinion on their taste?			
I prefer brand-name	12%	21%	8%
I prefer generic	2%	0%	4%
They are the same to me	58%	37%	69%
Don't know / Refused to answer	27%	41%	20%
N	202	70	132

Panel B.			
Y = Chose generic in hypothetical choice presented by survey.	Overall Mean	Never tried generic Mean	Tried generic Mean
Situation framing: Terrible headache	59%	32%	74%
N	134	47	87
Situation framing: Restocking supply	70%	40%	87%
N	165	57	107
Pooled across both situation frames	65%	37%	81%
N	298	104	194

Survey respondents were classifed on their past use of brand and generic based on their responses to questions 1 and 2 on the survey shown in Appendix C.1. Then, they were each asked a randomly selected two of the three questions shown. Percentages of each subgroup selecting each answer are shown.

Appendix Table A11: Consumer Survey Responses

	Categorization based on hypothetical choice	
	Brand choosers	Generic choosers
What is your opinion on how well they work to relieve pain?		
Brand-name works better	40%	13%
Generic works better	0%	3%
They work the same	43%	78%
Don't know / Refused to answer	16%	6%
N	67	133
What is your opinion of how safe the product is?		
Brand-name is safer	29%	7%
Generic is safer	0%	1%
They are equally safe	61%	84%
Don't know / Refused to answer	11%	8%
N	66	125
What is your opinion on their taste?		
I prefer brand-name	14%	8%
I prefer generic	0%	4%
They are the same to me	49%	63%
Don't know / Refused to answer	31%	25%
N	71	131

Survey respondents were categorized as brand choosers or generic choosers after being asked to make a hypothetical choice between the brand and generic versions of the painkiller they typically use, either Tylenol/acetaminophen or Advil/ibuprofen, with typical price and package quantity shown (see question 10 of the survey shown in Appendix C.1). Each respondent was also asked a randomly selected two of the three questions shown. Percentages of each subgroup selecting each answer are shown.