Information Distortion in Label Design in the Over-the-Counter Drug Market

Anastasiia Evdokimova

March 14, 2025

ABSTRACT. This paper examines how over-the-counter drug labels influence consumer perceptions of efficacy, distort decision-making, and shape equilibrium outcomes under counterfactual regulatory scenarios. It addresses a key identification challenge—the unobservability of perceived efficacy under different information structures—by conducting a randomized controlled trial and integrating its findings into a structural model. Using data from a control group and three treatment arms, I construct product-level measures of perceived efficacy beliefs based on pairwise product comparisons. Leveraging control group data supplemented with NielsenIQ data, I estimate a structural demand model that isolates the role of efficacy beliefs while accounting for heterogeneous preferences. I then incorporate updated beliefs to assess equilibrium effects under each information treatment. In equilibrium, the most effective intervention—emphasizing equivalent efficacy—increases substitution between biologically equivalent products by 26%, reduces consumer spending by 12%, but also introduces second-degree price discrimination driven by symptom-label preferences.

KEYWORDS. Consumer Behavior, Information Asymmetry, OTC Drug Markets, triplet embeding.

JEL CODES. D12, I11, L65.

CONTACT. aevdokimova@u.northwestern.edu. I thank David Dranove, Gastón Illanes, and Molly Schnell for their invaluable mentorship and encouragement. I also thank Vivek Bhattacharya, Federico Crippa, Devis Decet, Alessandro Pavan, Robert Porter, Amanda Starc, and Jin Yang, as well as seminar participants in various institutions, for their helpful comments. I am grateful to Getwizer for conducting the experiment. I acknowledge generous financial support from the CISO. Any errors are my own. Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

I. Introduction

Information asymmetries are pervasive in the over-the-counter (OTC) drug market. Pharmaceutical companies have a precise knowledge of drug efficacy, obtained through experiments and clinical trials that establish a direct link between a drug formulation and its efficacy in symptom relief. In contrast, consumers lack access to such rigorous information and instead rely on external cues—such as front-label claims—to assess a drug's efficacy (Harben et al., 2021). To mitigate this asymmetry, the FDA regulates OTC drug labeling, but it remains unclear whether existing regulations effectively prevent firms from distorting consumer perceptions and inflating markups (Berndt et al., 1996; Bronnenberg et al., 2015; Carrera and Villas-Boas, 2023; Crawford and Shum, 2005; Atal et al., 2022).

A striking indication that such distortions persist is the pricing strategy of drugs with identical formulations sold under different labels by the same brand. For example, Excedrin Extra Strength and Excedrin Migraine share identical formulations but are marketed as distinct products with price differences of up to 20%. Notably, 47% of consumers choose the more expensive option despite the availability of a cheaper alternative, suggesting that labeling and pricing strategies exploit information gaps to increase profits. Part of this purchasing behavior may also reflect genuine consumer preferences for certain labels. Since the expected efficacy of a drug in relieving symptoms is unobservable, distinguishing between preference-driven choices and belief-driven misconceptions remains a challenge, crucial for understanding the true impact of information asymmetries in the OTC drug market.

To identify belief-driven misconceptions and assess equilibrium responses to counterfactual regulatory setups aimed at reducing these misconceptions, I distinguish and separately identify consumer beliefs about a drug's efficacy from their preferences. By integrating a randomized controlled trial (RCT) with structural analysis, I evaluate the role of information asymmetries in shaping consumer misconceptions, affecting welfare, and influencing firm pricing strategies in the OTC drug market. This approach enables a comparison of equilibrium outcomes under the current labeling system with those under alternative information provision policies designed to reduce misinformation and enhance market efficiency.

The focus is on FDA-approved OTC drugs for headaches and migraines, which are widely used and have a significant consumer impact (World Health Organization, 2024). Many of these drugs share the same active components, making them particularly well-suited for studying two key consumer misconceptions: the perceived efficacy gap between

brand-name and generic drugs; and the perceived efficacy gap between chemically identical drugs marketed under different labels by the same brand.

Findings reveal that consumers systematically misjudge drug efficacy, leading to suboptimal purchasing decisions that can, in part, be corrected through more transparent information provision that leads to a new welfare-enhancing equilibrium. First, the experiment shows that consumers fail to recognize formulation-equivalent products 42% of the time, confirming the presence of informational distortions. Second, structural model estimates, which incorporate the expected efficacy identified in the experiment alongside NielsenIQ data, highlight the role of misbeliefs in consumer choice. Evidence shows how all information treatments reduce asymmetries and increase substitution between products with both identical and different active components. Third, incorporating a supply model to account for firms' responses to better-informed consumer behavior, counterfactual analyses indicate that the most effective intervention — explicitly stating equivalent efficacy on the front label— in equilibrium increases substitution elasticities by 24%, reduces consumer spending by 12%, and generates a welfare gain of $1.20 per consumer. However, this intervention also enables second-degree price discrimination based on symptom-label preferences, underscoring the need for careful policy trade-offs.

To demonstrate that the observed decision-making patterns, puzzling from a medical efficacy perspective, cannot be fully explained by label preferences, I first assess the extent of efficacy misperceptions under the current labeling system using data from the control group in the RCT. Respondents complete a series of product comparison tasks, where they evaluate triplets of drug labels and select the most similar pair based on perceived efficacy in relieving headaches or migraines. When faced with a triplet consisting of two drugs with identical formulations but different symptom labels and a third drug that shares a symptom label but not the formulation, participants incorrectly choose the label-matching drugs 42% of the time over the biologically equivalent option. Using these triplet comparisons, I construct product-level measures of perceived efficacy (Vankadara et al., 2023), revealing a systematic tendency for consumers to have higher rates of substitution in the perceived efficacy dimension based on label similarity rather than equivalent medical efficacy.

While perceived efficacy is a key determinant of OTC drug choices, other factors also shape purchasing behavior. To reach the main goal of my research, estimate the equilibrium response to counterfactual information treatment scenarios, I isolate the role of efficacy beliefs from other product attributes by estimating a demand model for headache and migraine drugs. A standard mixed-logit model (Berry et al., 1995) fails to disentangle

unobservable expected efficacy from other unobserved factors. To overcome this limitation, I integrate the expected efficacy measures identified in the experiment into the model. Building on Magnolfi et al. (2024), who use crowdsourced product similarity measures, I decompose substitution patterns into observable characteristics, efficacy beliefs, and other unobserved factors. The findings reveal that while misperceptions about efficacy significantly influence decision-making, they do not fully account for consumers' preference for more expensive biologically equivalent products. Brand loyalty and symptom-label preferences remain critical drivers of consumer choice.

These findings highlight the role of information misperceptions in consumer choice under current labeling. Since the main goal of my paper is to document equilibrium responses to information treatments, I need to identify beliefs under counterfactual regulatory setups. I do it by including three treatment arms in my RCT, who complete the same product comparison tasks as the control group but with modified product information. The first treatment expands front-label information to indicate other products with identical active ingredients, extending a practice currently used only by generics. The second treatment explicitly states the efficacy equivalence of different products, clearly indicating formulation equivalence (Carrera and Villas-Boas, 2023). The third treatment provides a pre-purchase informational brochure detailing product efficacy and active ingredients—a low-cost intervention that would not require manufacturer cooperation. Findings confirm that label content significantly influences perceived efficacy and that biases are substantially reduced when low-cost information on drug efficacy is provided.

The experiment identifies new expected efficacies under the treatments, which I then incorporate into the demand model to measure actual consumer responses to information provision. Shifts in perceived efficacy directly affect substitution patterns, influencing how consumers choose between products. Treatments that explicitly report formulation equivalence lead to higher substitution elasticities, particularly among biologically identical products within the same brand. The most effective intervention—explicitly reporting products with equivalent efficacy—increases substitution between products with identical formulations by 30% and between branded and generic products by 19%. The stronger effect within brands suggests that consumers were previously less aware of the existence of formulation-equivalent products within brands compared to their awareness of brand-generic alternatives.

Since firm pricing strategies depend on substitution patterns, they may adjust in response to information treatments. To capture these effects, I incorporate supply-side responses

into my analysis using a Nash-Bertrand pricing framework to estimate how firms adjust prices when consumers become more aware of formulation equivalence. When information asymmetries decrease and consumers recognize product equivalence, two opposing forces emerge: on the one hand, greater substitution between identical drugs intensifies price competition, driving prices down. On the other hand, firms exploit persistent label preferences, which now play a relatively larger role in consumer choice, by raising markups on symptom labels associated with higher willingness to pay. The net effect of these opposing forces is a decline in average prices, with prices falling by $1.55 in the counterfactual equilibrium under the second treatment (explicit reporting of equivalent efficacy).

Finally, to estimate consumer gains from increased information transparency, I combine demand and supply responses. Since consumer choices under full information are unobserved, I cannot directly assess how closely treatment choices approximate a full-information scenario (Allcott, 2013). Instead, I evaluate welfare gains using two key measures. First, I estimate equilibrium consumer savings. Under the treatment that explicitly reports formulation equivalence on the label, the new equilibrium results in average savings of $1.09 per consumer, representing 12% of the average price and totaling $6.6 million annually across the 50 largest markets. Consumers save $0.83 when provided with informational brochures, while savings decline to $0.56 when front-label statements only indicate the presence of the same active components. Second, beyond monetary savings, I assess whether information treatments lead consumers to choose drugs with higher perceived efficacy. Results indicate that improved labeling increases average welfare by $1.20 per person, with 10% of this gain attributed to increased consumption of drugs perceived to be more effective.

This paper contributes to the literature on consumer and firm decision-making in the OTC drug market (Bronnenberg et al., 2015; Muthukrishnan et al., 2009; Handel, 2013; Atal et al., 2022; Carrera and Villas-Boas, 2023), departing from these studies in two key dimensions. First, this paper integrates an RCT and a structural model, combining the advantages of these two approaches, which have previously been exploited separately by Atal et al. (2022) and Carrera and Villas-Boas (2023). The experiment identifies baseline and counterfactual expected efficacy without relying on a specific functional form for belief formation, which is an assumption in the structural model of Atal et al. (2022). At the same time, the structural model enables to compute equilibrium response to the counterfactual information provision scenario, going beyond the reduced-form treatment effects estimated by Carrera and Villas-Boas (2023), who were the first to propose and test labeling interventions in a retailer-level experiment to reduce information uncertainty in the OTC market. Second, it broadens the

scope of research on misconceptions about drug efficacy in the OTC market. While existing literature primarily focuses on the branded vs. generic puzzle — where branded products dominate market share despite being more expensive than generics — this study extends the analysis beyond that comparison.It documents the existence of a broadly defined misbelief in consumer perceptions of drug efficacy, examining deviations from medically established efficacy and analyzing how these misconceptions adjust in response to new information.

More broadly, this paper contributes to the literature on information distortion in consumer markets by examining a regulated information provision environment where consumers make decisions based on incomplete or biased information. Prior research has documented various forms of "irrational" choices resulting from product quality distortions (Ellison and Wolitzky, 2012; Ellison and Ellison, 2011; Houde, 2018). This paper builds on studies that explore how consumers obtain information from product labels and how firms strategically use labeling to extract higher markups (Barahona et al., 2023; Kiesel and Villas-Boas, 2013; Zhu et al., 2015). It extends these analyses to an additional market segment and employs a novel methodological approach to better understand the consequences of labeling-driven information distortions. Furthermore, this study contributes to the broader discussion on market inefficiencies arising from incomplete quality information, as examined in prior research (Dranove et al., 2003; Jin and Leslie, 2003; Greenstone et al., 2006; Dranove and Jin, 2010; Roe et al., 2014; Houde, 2018; Vatter, 2021).

Methodologically, this paper contributes to the discussion on integrating RCTs with structural approaches to generate more effective policy recommendations (Todd and Wolpin, 2023). It demonstrates how this integration can be implemented and highlights that computing a new equilibrium under each counterfactual provides insights beyond reduced-form treatment effects. This approach is motivated by the increasing use of advanced machine learning tools to identify unobservable choice patterns (Magnolfi et al., 2024). Unlike previous methods that rely on aggregate substitution measures, my approach explicitly defines and estimates a specific but unobservable parameter—quality (efficacy)—using machine learning techniques. This enables a clearer distinction between consumer preferences and unobservable product characteristics, enhancing the methodology's applicability to counterfactual analysis.

The paper is structured as follows: Section II outlines the institutional background, data, survey, and experiment design. Section II.C presents reduced-form evidence motivating the structural model. Section IV introduces the demand model, with results discussed in Section V. Section VI covers the supply-side setup, while Section VII integrates experimental

findings into the demand and supply models, analyzing their effects on decision-making, savings, and welfare.

II. Setting and data

II.A. OTC drug market

This paper examines the OTC drug market, where consumers make purchasing decisions without a prescription or official diagnosis. As a result, self-diagnosis and self-treatment are common: individuals independently assess their symptoms and select medications they believe will provide relief. However, most consumers lack precise knowledge of the relationship between a drug's chemical and biological properties and its effectiveness, leading them to navigate the market with imperfect information. Consequently, they rely on front-label information (Harben et al., 2021) and past purchasing experiences to guide their choices. The supply of OTC drugs is regulated by the FDA, which determines both market availability, FDA-approved active ingredients, and packaging standards.[1]

II.B. Data

My analysis focuses on pain medications that alleviate migraine and headache symptoms. I focus on this market since it is widespread [2] and provides me with two key advantages: (1) the ability to have different symptom labels within the same consideration set — there is a non-zero chance that both migraine and headache medications are considered by migraine and headache sufferers; (2) variation in symptom labeling despite identical active components — active ingredients approved for migraine relief are also approved for headache relief. As of 2017, the year of my analysis, the FDA-approved active ingredients marketed for migraine relief are: a) the Migraine formula" (acetaminophen + aspirin + caffeine), and b) ibuprofen. For headache relief, the approved active ingredients are: a) the Migraine formula," b) ibuprofen, c) naproxen sodium, d) acetaminophen, and e) aspirin.

I utilize three primary data sources: (1) NielsenIQ Retail Scanner (RMS) and NielsenIQ Consumer Panel (Homescan) data from the Kilts Center at the University of Chicago, which offer detailed insights into prices, quantities, product characteristics, and consumer

[1] https://www.fda.gov/drugs/how-drugs-are-developed-and-approved/over-counter-otc-nonprescription-drugs

[2] Headache disorders being among the most prevalent neurological disorders, affecting 48.9% of the global population (Ahmed, 2015).

demographics; (2) a randomized controlled trial (RCT) comprising a control group and three treatment arms, designed to identify perceived efficacy; and (3) the Medical Expenditures Panel Survey (MEPS), which is used to quantify the outside option.

II.B.1. NielsenIQ data

The data on products, markets, and socio-demographic characteristics are sourced from the NielsenIQ Retail Scanner (RMS) and NielsenIQ Consumer Panel (Homescan) databases, provided by the Kilts Center at the University of Chicago. The NielsenIQ Retail Scanner data offers detailed information on products at the week/store level, including Universal Product Code (UPC), prices, quantities, and product characteristics. The NielsenIQ Consumer Panel data includes purchase information from a set of panelists at the UPC, store, and day level, along with demographic details about the panelists.

I selected 2017 for my analysis as the most recent year unaffected by health-related shocks.[3]Products are defined by brand, active components, symptom label, and size, consolidating multiple UPCs. Since NielsenIQ Scanner data lacks symptom labels, I identified them by matching common consumer names to observed labels and active ingredients. My analysis covers 350 products from seven major brands—Advil, Tylenol, Bayer Aspirin, Aleve, Excedrin, and Motrin—along with equivalent generic products. Generics account for 2% of sales in the selected markets (see online appendix 14), with an average product price of $9.66 (Table 1). A market is defined by month, retailer, and Nielsen Designated

Statistic	Mean	St. Dev.	Min	Max
Price (Dollars Per Pack)	9.66	5.04	1.50	30.59
Share Within Market	0.01	0.01	0.01	0.25
Outside good share	0.30	0.08	0.13	0.4

TABLE 1: Basic Statistics: Average Across All Markets and Products

Marketing Area (DMA), which are geographic regions in the U.S. where local advertising is uniform. Prices are aggregated at the product and market levels by taking the average within DMA. This does not limit price variation across markets: in my data the average coefficient of variation of prices for purchased products across stores within a market is 0.07, and the average variation in the average number of unique products purchased is 0.12.

[3]2018 is excluded due to potential disruptions from Affordable Care Act changes and later years due to (post)pandemic trends

Moreover aggregating to the DMA level is common (DellaVigna and Gentzkow, 2019). To narrow the focus, the analysis is restricted to the 50 largest retailers in DMAs, tracked monthly, resulting in a total of 600 markets. Seventeen retailers serve these markets (see online Appendix Table 15). In May, a randomly selected month, the number of products each retailer offers ranges from 10 to 98, and the number of markets served ranges from 1 to 19. Based on my market definition, the maximum number of active components is 6, and all markets include all possible active components. This is discussed in more detail in Section II.C.

Homescan data supplements the RNS data with consumer demographics, including age, gender, household income, and household size.

II.B.2. Medical expenditures panel survey

The Medical Expenditure Panel Survey (MEPS), provided by (Blewett et al., 2023), combined with the RMS data, is used to compute the outside options. I define an outside option as a combination of headache and migraine prescription medications as well as all non-prescription medications that do not require FDA approval (such as homeopathic remedies). MEPS provides data on the region/month-level quantity of medications prescribed for headache or migraine pain and filled by consumers.

To obtain the outside option at the market level rather than the region-month level, I create weights to redistribute region-month volumes across markets. These weights are defined as the income and age composition of each market, weighted equally. This approach ensures that the outside option reflects demographic differences across markets, with an average share of 30% (Table 1).

II.C. Descriptive evidence

This subsection examines how front-label designs influence consumer choices and firm behavior using NielsenIQ data. The findings highlight two key insights. First, brands sell biologically equivalent products under different labels with no second-degree price discrimination based on medical conditions. Second, 47% of store-week purchases involve consumers selecting a more expensive drug despite the availability of a cheaper, biologically identical alternative under a different label.

II.C.1. Pricing strategy

Based on the FDA Label Repository[4] when possible brands produce biologically identical products under different symptom labels. To examine whether this phenomenon constitutes a market failure, I first evaluate pricing strategies. As a case study, I focus on two drugs: Excedrin Extra Strength and Excedrin Migraine, and test the hypothesis of whether identical products that vary only by label also vary in price. To test this, I conduct the following regression for each package size available in the market at a week (t) and store (s) level:

$$\text{price}_{s,t}^{\text{extra strength}} = \beta_1 \cdot \text{price}_{s,t}^{\text{migraine}} + \gamma_s + \delta_t + \epsilon_{s,t} \tag{1}$$

where γ_s - store fixed effects,δ_t - week fixed effects. Based on the results (Table 2), β_1 is statistically different from 1. Therefore, I can claim that Excedrin Migraine and Excedrin Extra Strength at the weekly level are priced differently.

	Price Extra Strength								
	20 geltab	24 caplet	30 caplet	50 caplet	80 geltab	100 caplet	200 caplet	250 caplet	300 caplet
Price Migraine	0.3952	0.4740	0.3416	0.0186	0.6202	0.4889	0.2206	0.0708	0.0342
	(0.0042)	(0.0039)	(0.0098)	(0.0074)	(0.0078)	(0.0058)	(0.0146)	(0.0175)	(0.0339)
Fixed-effects									
Store	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	296,988	533,548	29,066	39,074	121,809	265,748	37,194	13,578	10,449
R^2	0.89431	0.89653	0.87805	0.52289	0.78754	0.69236	0.77362	0.85698	0.11951

Clustered on store level standard-errors in parentheses

TABLE 2: Correlation Between Excedrin Extra Strength and Excedrin Migraine Prices, Weekly

To evaluate the presence of second-degree price discrimination based on front-label attributes, I categorized all observed week-store-level prices of identical products differing only by symptom label into three groups: (1) identical prices ($p_m = p_{es}$), (2) migraine-labeled products priced higher ($p_m > p_{es}$), and (3) migraine-labeled products priced lower ($p_m < p_{es}$). I then calculated the share of weeks, at the store level, in which each pricing strategy was applied.

As shown in Figure 1, for 100-caplet packages, the median store maintains identical pricing for 75% of the weeks, with this share dropping to 50% for stores stocking the product year-round. The remaining weeks are nearly evenly split between cases where migraine-labeled products are priced higher and cases where extra-strength variants are

[4]https://labels.fda.gov/

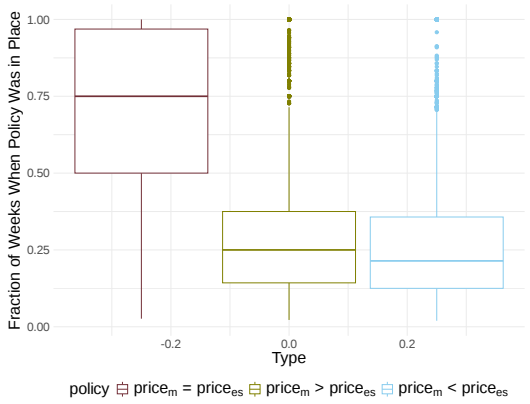


Figure 1: Pricing policy for 100 Caplets: Excedrin Migraine and Excedrin Extra Strength

more expensive. Pricing variation is even more pronounced for smaller packages (see Online Appendix, Figure 35). These findings suggest no systematic second-degree price discrimination based on observable front-label attributes.

II.C.2. Consumers' "wrong" choices

Price differences alone may not reduce consumer welfare if consumers are informed and, without strong label preferences, opt for the cheaper drug when available. For now, assuming no strict label preferences among consumers, I can identify whether a transaction was a "wrong" choice—i.e. when a more expensive alternative was purchased despite a cheaper, biologically identical product of the same brand being available.

For this analysis, I calculated the fraction of "wrong" choices within each store/week and then plotted a histogram for all stores/weeks in 2017. As shown in Figure 2, under the previously discussed assumptions, the average fraction of "wrong" purchases is 47%. The prevalence and high frequency of these "wrong" choices suggest that the existence of different labels for biologically equivalent drugs of the same brand, combined with their pricing strategies, poses a potential threat to consumer welfare.

This serves as a first approximation of the potential inefficiencies caused by marketing biologically equivalent products of the same brand under different labels. The approach may lead to overestimation of misinformed choices. This occurs because the analysis assumes that consumers do not have symptom label preferences, such as placebo effects associated with labeling. If such preferences exist, some consumers may be making deliberate choices rather than mistakes, inflating the measured fraction of "wrong" decisions. Or, conversely, underestimation of the issue. If a consumer selects the cheaper drug, it is unclear whether

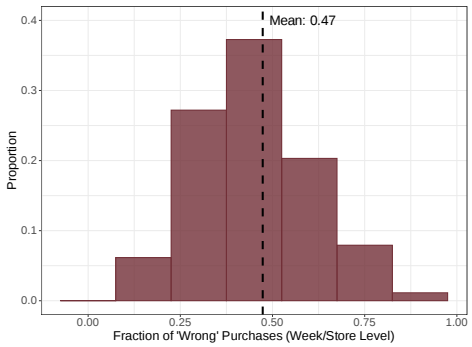


Figure 2: Histogram of Wrong Choices, Week/Store Level

this decision results from an informed choice or mere luck. Without distinguishing between the two, the estimated fraction of misinformed choices may be lower than the true extent of the problem. My structural model (Section IV) addresses the issue of separate identification, while the RCT results from the control group provide evidence of the existence of relative efficacy misconceptions III.

III. Randomized control trial

As highlighted in the previous section, selecting a more expensive product from a biologically identical pair does not, on its own, serve as direct evidence of market failure. This is because it remains infeasible to clearly distinguish between preference-driven choices and belief-driven misconceptions as underlying factors. To address this challenge, I conduct a randomized controlled trial (RCT) on individuals aged 18 and older, representative of active OTC drug shoppers in the U.S. The survey, administered in winter 2024 by the marketing firm Getwizer, includes a baseline group and three treatment arms, totaling 1,120 respondents. Participants were stratified into four groups using covariate-adjusted randomization (Bugni et al., 2019), ensuring balanced comparisons across treatments.

The RCT serves multiple objectives. It measures perceived efficacy both before and after exposure to different information treatments, separately assessing perceptions of migraine and headache relief. Additionally, it identifies the probability of entering the market with the belief that a drug is necessary for symptom relief. This probability will be incorporated into my structural model (IV) to allow for heterogeneity in the decision-making of migraine and headache relief shoppers.

III.A. Data

III.A.1. Baseline perceived efficacy

Given the challenges of assuming that individuals can precisely quantify the efficacy of each drug (Künstle et al., 2022) and that these scales are comparable across individuals, I instead rely on a relative pairwise comparison measure of expected efficacy. Since perceived efficacy may vary depending on whether an individual seeks relief for headaches or migraines, participants are randomly assigned to one of two groups: one evaluating headache relief and the other migraine relief.

Participants evaluate drug effectiveness through pairwise comparisons within triplets, allowing for the measurement of relative perceptions of efficacy. An example of this question format is shown in Figure 3.[5]

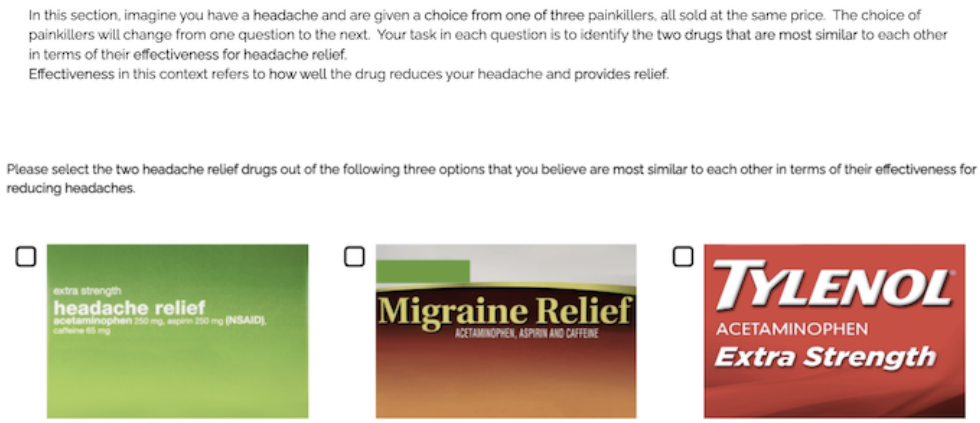


Figure 3: Example of the Question

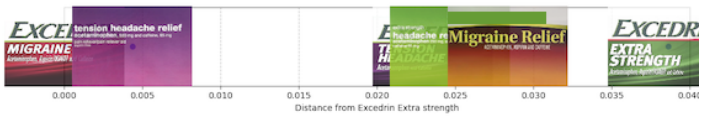
Each respondent answered 15 questions of this type.[6] I then aggregate responses on relative efficacy perceptions at the population level using soft ordinal triplet embedding (Vankadara et al., 2023; Magnolfi et al., 2024), an algorithm that translates respondents' binary choices into coordinates within a relative perceived efficacy space. A detailed discussion of this algorithm is provided in Online Appendix IX.A.2.

One way to illustrate triplet embedding results is by computing the relative distance

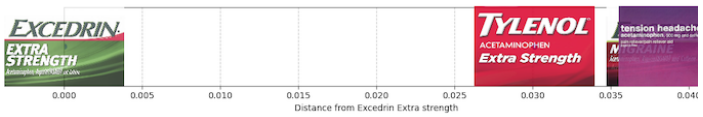
[5] I ask respondents about effectiveness rather than efficacy, as the latter is a medical term referring to clinical trial results, whereas effectiveness reflects an individual's perception of a drug's ability to relieve symptoms. In this context, expected efficacy and effectiveness are used interchangeably.

[6] The number of questions per respondent was determined based on a pilot survey, aiming to balance the trade-off between increasing information for the researcher and minimizing respondent fatigue.

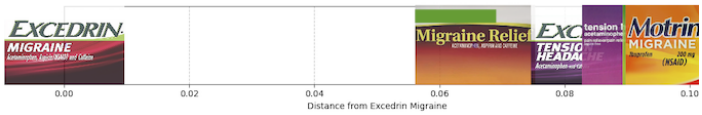
between a given product and all others. Figure 4a shows the distances from Excedrin Migraine to its closest substitutes, while Figure 4b presents the corresponding distances for Excedrin Extra Strength, based on responses from individuals selecting products for migraine relief. Respondents struggle to identify biologically equivalent products—those with the same active ingredients and efficacy—as their closest substitutes. Additionally, the results highlight the contextual influence of symptom labels on consumer decision-making. Individuals tend to first group products with synonymous symptom labels before differentiating between branded vs. generic versions or biologically equivalent products within the same brand. The positioning of drugs in the relative perceived efficacy space suggests that current information provision in the OTC market fosters biased beliefs about drug efficacy.



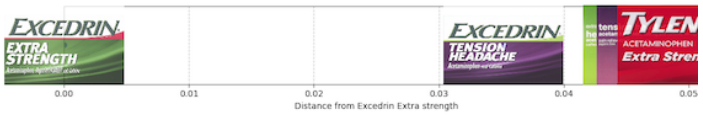
(a) Excedrin Migraine; Relieve Migraine



(b) Excedrin Extra Strength; Relieve Migraine



(c) Excedrin Migraine; Relieve Headache



(d) Excedrin Extra Strength; Relieve Headache

Figure 4: Distance from Excedrin Products to Other Drugs for Migraine and Headache Relief

Notably, triplet embedding offers key advantages over second-choice data and conjoint analysis. Unlike the second-choice approach, which only identifies relative substitution patterns (e.g., "A is the closest substitute to B"), triplet embedding also quantifies the substitution distance between objects, providing a more detailed substitution structure. Compared to conjoint analysis, triplet embedding does not rely on assumptions about the independence or additivity of decision-making factors (Luce and Tukey, 1964; Knoblauch and Maloney, 2012). Online Appendix IX.F, presents additional discussion on the substitution patterns within the perceived efficacy dimension and presents a full space.

III.A.2. Experiment

A key focus of my counterfactual analysis is the change in belief about a drug's efficacy after an individual receives new information. To avoid making structural assumptions about

the relationship between the beliefs and the information provided, I conduct a randomized controlled trial (RCT) with three treatment arms. Each treatment arm consists of 224 participants, selected through covariate-adaptive randomization to match the control group (Bugni et al., 2019).

The first treatment involves modifying the current strategy used by generic pharmaceutical companies to encourage consumers to choose generics over branded versions (Carrera and Villas-Boas, 2023). Typically, generics emphasize that they contain the same active ingredients as branded products (Figure 5a). However, given the low proportion of generic purchases, this approach seems ineffective. To enhance it, I added similar information to branded product labels, indicating that they, too, contain the same active ingredients (Figure 5b). This labeling is applied uniformly across all products with the same active molecule, ensuring that consumers are consistently informed about product equivalence.

The second treatment builds on findings from (Carrera and Villas-Boas, 2023), which show that information about the similarity in effectiveness based on other consumers' experiences is more successful in educating consumers about biological equivalence than information about active ingredients. In this treatment, the front label explicitly conveys equivalent efficacy for products containing the same molecule. As shown in Figure 6a, this approach helps bridge potential gaps in consumer knowledge by highlighting biochemical similarity, even if differences in branding or labeling initially obscure it.

From a policy perspective, both treatments are costly and require regulatory enforcement. Therefore, I conducted a third, more cost-effective information treatment that does not require such enforcement. In addition it allows me to compare the difference from direct education of the consumers versus information provision. In this treatment, I provided consumers with information about the effectiveness and active components of products before they chose (Figure 6b). To avoid confusing consumers with information about both migraine and headache medications, which might imply they are different symptoms, I focused solely on information about migraine drugs. This approach ensures that consumers spend several minutes reviewing the material, reducing the likelihood of them skipping it.

In the treatments that highlight biological equivalence by indicating the same active components and same efficacy on the front label, the triplet questions remain the same, but the drugs' front label is changed. For the treatment that provides respondents with the brochure, the only change compared to the baseline survey is that respondents are shown a brochure with relevant information before the triplet questions are asked.[7] For each

[7]The survey design controls for the respondents' attention to the provided information.



(a) Generic Actual Market



(b) "Same Active Components" Treatment

Figure 5: Comparison of Front Label Examples



(a) "Same Effectiveness"


Here are some common OTC medications used to treat migraine, plus their active ingredients:

Brand or Product Name	Active Ingredients (plus amounts per tablet or capsule)
Aleve®	Naproxen sodium 220 mg
Advil®	Ibuprofen 200 mg
Bayer	Enteric-coated aspirin 325 mg
Anacin®	Aspirin 400 mg, caffeine 32 mg
Excedrin® Migraine	Aspirin 250 mg, acetaminophen 250 mg, caffeine 65 mg*
Excedrin® Tension Headache	Acetaminophen 500 mg, caffeine 65 mg*
TYLENOL® Regular Strength	Acetaminophen 325 mg
TYLENOL® Extra Strength	Acetaminophen 500 mg
TYLENOL® 8 HR Arthritis Pain	Acetaminophen 650 mg
Motrin® IB	Ibuprofen 200 mg

(b) "Migraine Educational Material"

Figure 6: Second and Third Treatments

information treatment design, I compute new values of the relative expected efficacy.[8]

III.A.3. Probability of entering the market with a certain symptom belief

The probability of entering a market with the belief that a drug is needed to address migraine symptoms is computed based on the survey responses.[9] To account for potential socio-demographic differences in migraine prevalence, I follow the medical literature and categorizerespondents into four groups based on median income and age (above and below 65). The results of this classification are presented in Table 3. The computed probabilities align with the medical literature, which indicates that lower-income populations and younger individuals experience higher rates of migraine pain and/or self-diagnosis (Stewart et al., 1992).

Using these probabilities, I estimate the likelihood with which each individual in the socio-demographic sample enters the market believing they need a drug to relieve migraine symptoms or headaches.

[8] I discuss the computation of this value in more detail in the Online Appendix IX.A.3. New expected efficacy

	Age below 65	Age 65 +
Income ≤ 75	0.53	0.36
Income ≥ 75	0.45	0.32

TABLE 3: Probability to Enter the Market with the Belief That You Need Migraine Relief

III.B. Descriptive evidence

III.B.1. Wrong choices in perceived efficacy dimension

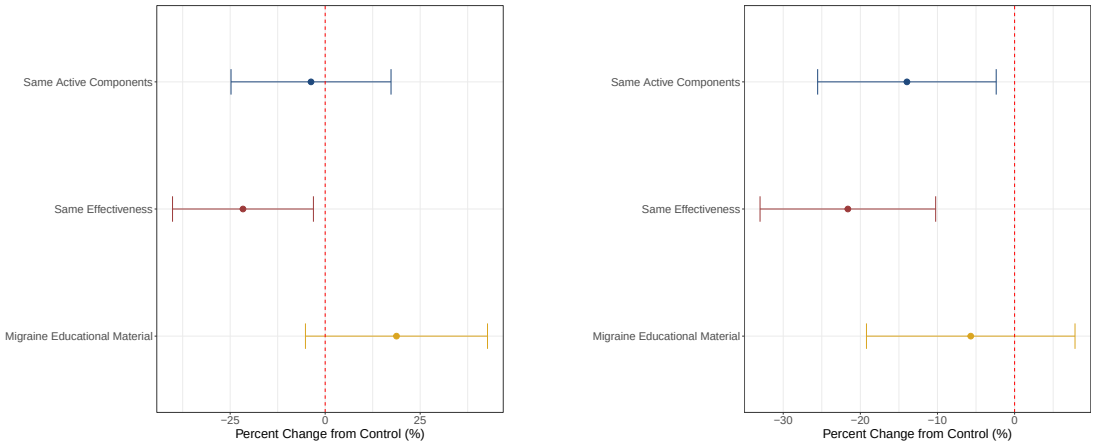
To demonstrate that the "wrong" choices discussed in Section II.C are not solely driven by strong label preferences, I analyze responses from triplet-based comparisons. In this analysis, I compute "wrong" choices based solely on perceived efficacy, focusing on cases where two biologically equivalent drugs have different symptom labels, while a third drug shares a label with one of them but contains different active ingredients. I then calculate the fraction of cases in which participants perceive the drug with the matching label—rather than the biologically equivalent option—as the most similar in perceived efficacy. On average, this probability is 42%.

To evaluate whether my information treatments reduce these misconceptions, I calculate the change in the fraction of "wrong" choices based on the medical definition of efficacy for each treatment relative to the control group. Specifically, I distinguish between two cases: first, when individuals select two branded products with different active ingredients, ignoring a generic version of one as the third option in the triplet; and second, when individuals select products with different active ingredients but identical symptom labels, disregarding a third product that is of the same brand, contains the same active components, but has a different label.

Figure 7 presents the average percent change in the fraction of correct choices across different treatments relative to the control group, with standard errors computed at the 10% significance level due to the small sample size and high variance in percent of "wrong" choices. Among the interventions, the front-label statement indicating equal effectiveness had the strongest impact in reducing both types of misconceptions. It decreased the likelihood of failing to select biologically identical products with different symptom labels or failing to choose a generic version by approximately 22% compared to the control. The second most effective treatment was the one explicitly stating which products share the same

presented in Online Appendix IX.F

[9]Detailed computation is discussed in the Online Appendix IX.A.1.



(a) Brand/Generic

(b) Different label biologically equivalent

Figure 7: Change in the Fraction of Correct Choices Across Different Treatments Compared to Control Group

active ingredients. This intervention resulted in a roughly 4% reduction in failing to select a generic version and a 14% reduction in failing to recognize biologically identical products with different symptom labels. In contrast, the least effective treatment was the educational material, which showed no improvement—or even an increase—in the probability of making "wrong" choices.

This pattern suggests that conveying efficacy equivalence through a direct front-label statement requires minimal cognitive effort, enabling consumers to immediately recognize product similarities. In contrast, treatments highlighting identical active ingredients demand greater cognitive processing, as individuals must actively associate ingredients with effectiveness. Educational materials require the highest cognitive effort, as respondents must first absorb detailed information and later recall and apply it when making a decision. Additionally, it is possible that presenting respondents with a list of products containing different active ingredients inadvertently led them to perceive those products as closer substitutes.[10] Notably, respondents were more likely to correctly match products with different labels within the same brand. This suggests that brand-generic equivalence is relatively well understood, whereas intra-brand equivalence remains less familiar to consumers. Additional evidence in the Online Appendix (37) indicates that individuals with higher ex-ante information acquisition costs were the most responsive to the information treatments.

[10]This hypothesis is explored further in the section on changes in substitutability within the perceived efficacy dimension (V.C).

III.B.2. Mechanism

To test the role that information acquisition plays in driving misperceptions, and to establish that providing low-cost information directly targets the root cause of these misconceptions, I examine whether the likelihood of making a "wrong" choice is systematically linked to lower information acquisition costs. To formally assess this relationship, I estimate the following regression model:

$$\begin{aligned}\mathbf{Pr}(\text{wrong choice"})_i = &\beta_0 + \beta_1 \cdot \text{female}_i + \beta_2 \cdot \text{age}_i + \beta_3 \cdot \text{age}_i^2 + \beta_4 \cdot \log(\text{income})_i + \\ &\beta_5 \cdot \text{high school education} + \sum_k \gamma_k \cdot X_i^k + \epsilon_i \end{aligned} \tag{2}$$

where X_i^k represents socio-demographic controls. Due to the small sample size and the high correlation between education and income, I retain education as a control variable instead of both.

Dependent Variable:	**Pr**("wrong choice")		
Model:	(1)	(2)	(3)
age	-0.0181	-0.0243	-0.0206
	(0.0077)	(0.0084)	(0.0076)
age^2	0.0002	0.0002	0.0002
	(0.0000)	(0.0000)	(0.0000)
high school education and higher	-0.0969	-0.1313	-0.0724
	(0.0557)	(0.0648)	(0.0527)
Group allocation	Yes	Yes	Yes
region	Yes	Yes	Yes
industry		Yes	
frequency purchase			Yes
Observations	224	210	224

mean: 42% wrong choices
Heteroskedasticity-robust standard-errors in parentheses

TABLE 4: Impact of Information Acquisition Costs on Wrong Choices

Table 4 indicates that individuals around age 45 have the lowest probability of making medically "wrong" choices. Furthermore, having at least a high school education reduces this probability by approximately 10 percentage points, representing 23% of the mean probability of making wrong choices. These findings remain consistent across all specifications. A likely explanation is that individuals with lower information acquisition costs—such as middle-aged individuals or those with higher education levels (Maćkowiak et al., 2023)—are

better equipped to process information about active ingredients and their efficacy. Drug labels convey both zero-cost information (e.g., linking products to symptoms) and costly information (e.g., understanding active ingredients). Those with lower acquisition costs are more likely to engage with the latter, leading to better-informed decisions and a reduced likelihood of making wrong choices, as confirmed by the regression results.

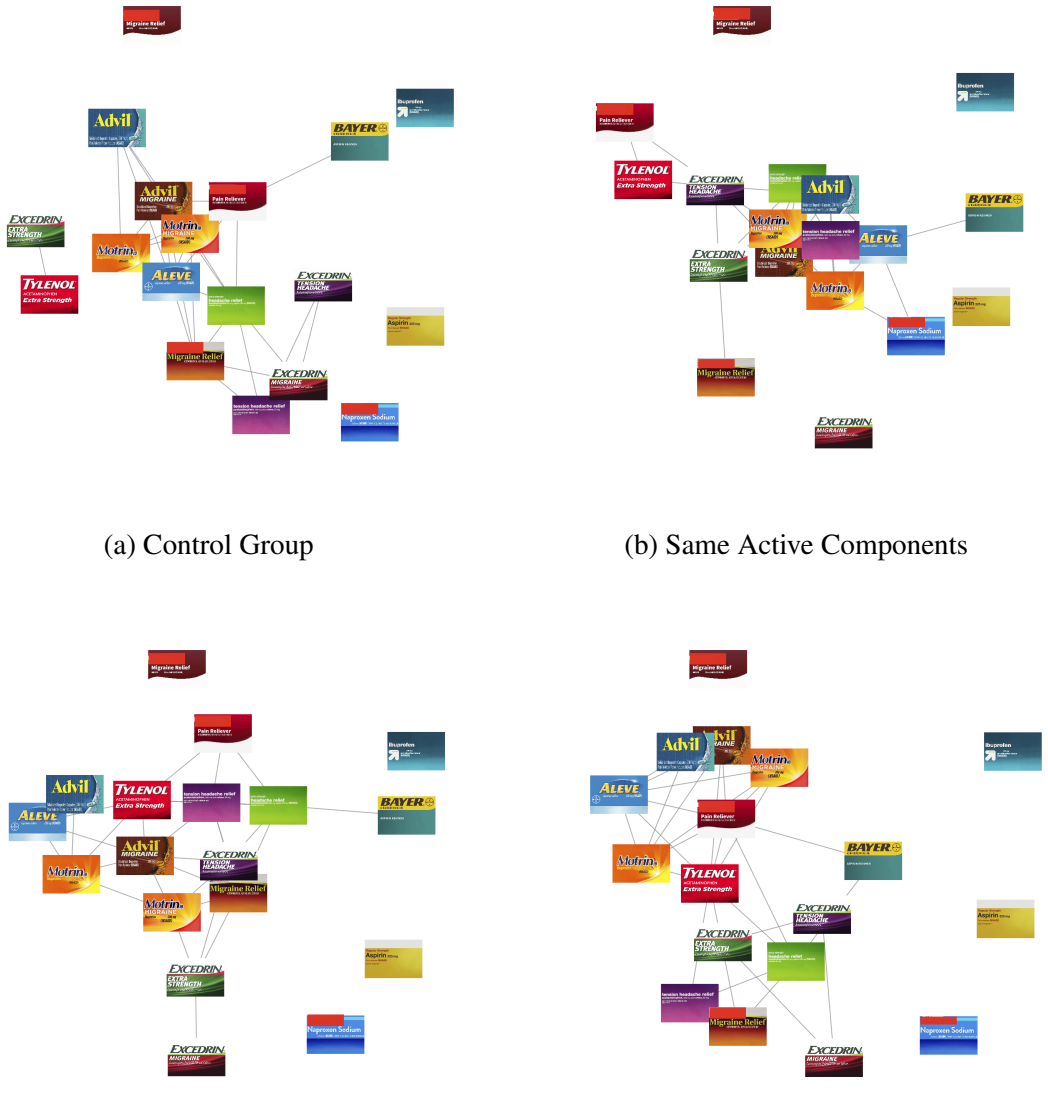
III.B.3. Change in substitution in efficacy dimension

While information treatments reduce the probability of making a "wrong" choice in the triplet comparisons, triplet embedding allows for a broader evaluation by examining substitution patterns among products in the perceived efficacy dimension and how these patterns evolve in counterfactual information scenarios. Under a framework defined by medical efficacy, chemically identical products of the same brand but with different labels, as well as branded and generic versions, should be closer substitutes in the perceived efficacy space than other products.

To assess these substitution patterns, I use the relative efficacy spaces generated by the triplet embedding algorithm to construct a network graph, where each node represents a product and edges indicate perceived similarity. The pairwise Euclidean distance between products in a multidimensional perceived efficacy space serves as the basis for determining substitutability. Rather than connecting all product pairs, I apply a thresholding approach to filter out weaker relationships, ensuring clarity and interpretability within the network. Specifically, only the 25% closest product pairs are connected by edges, highlighting the strongest substitutable relationships. These clusters represent groups of highly substitutable products.

The network graphs in Figure 8 illustrate several key patterns that reinforce the previously discussed changes in wrong choices and provide additional insights into how substitution patterns evolve across treatments. First, in the control group, substitution clusters are primarily defined by symptom labels rather than medical equivalence. For instance, Excedrin Extra Strength forms a cluster with Tylenol Extra Strength, despite having different active ingredients. Moreover, in most cases, generic and branded versions of the same active ingredient do not belong to the same cluster and are not directly connected. However, when participants are provided with explicit information about active ingredients, some branded and generic products with the same active components become connected, forming stronger substitution relationships. Additionally, Excedrin Extra Strength no longer clusters with Tylenol Extra Strength and instead connects with its generic equivalent, both labeled for

headache and migraine relief. The "Same Effectiveness" treatment, while not as effective in establishing new brand-generic connections, significantly increases substitutability among biologically identical products with different labels, making them more likely to be perceived as close substitutes.



(a) Control Group

(b) Same Active Components

(c) Same Efficacy

(d) Migraine Education Material

Figure 8: Distance from Excedrin Products to Other Drugs for Migraine and Headache Relief

Meanwhile, the Migraine Education treatment supports the wrong choice results: al-

though there are some improvements in forming medically valid connections, the primary effect of this treatment is a general increase in overall substitutability across all products, reducing differentiation between them. These network graphs suggest that if we assume perceived efficacy is the primary driver of substitution patterns, then all information treatments, compared to the control, lead to an increase in substitutability between products with similar medical efficacy, albeit to different degrees. The "Same Active Ingredients" treatment is the most effective at improving brand-generic substitutability, while the "Same Effectiveness" treatment primarily enhances substitutability among biologically equivalent products with different labels.

IV. Demand on the OTC drug market

Consumers are often assumed to base drug purchases on product characteristics, including perceived quality (Atal et al., 2022; Bronnenberg et al., 2015; Carrera and Villas-Boas, 2023; Crawford and Shum, 2005). For non-prescription drugs, which primarily alleviate specific symptoms but have ex-ante unobservable efficacy, I define quality as consumers' beliefs about a drug's ability to relieve their symptoms. Expected efficacy, being inherently unobservable, varies both across individuals and within individuals based on their symptom perceptions.

The previous section demonstrated how efficacy misconceptions shape consumer choices. Building on this, I show how survey data can be used to isolate the role of expected efficacy in decision-making. Under specific assumptions, standard mixed-logit estimation tools can be applied by leveraging iterated expectations and a full probability model. I conclude by discussing the estimation process and the instruments used to address potential endogeneity concerns.

Incorporating beliefs about efficacy as a parameter in the decision-making problem is essential for computing equilibrium responses to information treatments, as it serves as a key parameter in my counterfactual analysis. Failure to identify expected efficacy not only prevents disentangling efficacy from unexplained product-specific fixed effects but also risks conflating it with preferences for label and active component attributes. This conflation arises because consumers may rely on these attributes when assessing a product's effectiveness in alleviating symptoms [11].

[11] Consumers use the front label as a guide to determine which products offer higher efficacy in alleviating symptoms compared to others

IV.A. Set up

I adapt the utility function from Crawford and Shum (2005) to capture the unique characteristics of the non-prescription drug market. While their model treats expected efficacy as an indicator of drug quality, it is tailored for the prescription drug market, where uncertainty arises after a physician selects a drug, and patients learn about efficacy through prescription experience. In contrast, uncertainty in the non-prescription market occurs at the choice stage, as consumers must independently match their symptoms to the appropriate drug. Without physician guidance (Catlin and Brass, 2018; Woźniak-Holecka et al., 2012), their decisions rely primarily on front-label information (Harben et al., 2021) and prior market experience. In my model, I do not take a stance on how consumers form perceived efficacy beliefs instead I recover them from the RCT. Every consumer seeking to alleviate symptoms (s) and entering a market (t) considers the following attributes when making a purchase decision for good (j): $\mathcal{C} : (l_j, a_j, g_j, p_{j,t}, w_j, q_{i,j,s}, \psi_i,)$, where: (1) l_j - symptom label; (2) a_j - active components; (3) g_j - indicator whether the drug is a branded or generic; (4) $p_{j,t}$ - price; (5) w_j - size of the package; (6) $q_{i,j,s} = \mathbf{E}\left[\textit{Efficacy}_{i,j} | s_i\right]$ - symptom (s_i) specific expected efficacy; (7) ψ_i - non-observable by researcher factors that drive choices, such as previous experience, recommendations from friends, etc...

Incorporating these parameters in the standard mixed-logit demand model (Berry et al., 1995; Nevo, 2000) in a market (t) utility of a consumer (i) who believes they are having a symptom/disease (s) from a possible set of symptoms $s \in \mathcal{S}$ for a product (j) has a following form:

$$
\begin{aligned}
u_{i,s,j,t} = & \underbrace{\hat{\delta}_{j,t} + \gamma_j \cdot p_{j,t}}_{\text{mean utility}} + \underbrace{\sum_{k,r} p_{j,k} \cdot z_{i,r} \cdot \gamma^{obs}_{k,r} + \gamma_s \cdot p_j \cdot \mathbf{I}[s_i = s]}_{\text{consumer characteristic specific price sensitivity}} \\
& + \underbrace{\overbrace{\beta^{eff.}_s \cdot \mathbf{E}\left[\textit{Efficacy}_{i,j} \mid s\right] \cdot \mathbf{I}[s_i = s]}^{\text{beliefs}} + \overbrace{\sum_{k,r} x_{j,k} \cdot z_{i,r} \cdot \beta^{obs}_{k,r} + \sum_k x_{j,k} \cdot \mathbf{I}[s_i = s] \cdot \beta^{obs}_{s,k}}^{\text{preferences}}}_{\text{product attributes - observed consumer characteristics specific utility}} \\
& + \sum_k x_{j,k} \cdot v_{i,k} \cdot \beta^{unobs}_{k,r} + \epsilon_{i,s,j,t}
\end{aligned}
\tag{3}
$$

Where:

$$
\hat{\delta}_{j,t} = \sum_k x_{j,k} \cdot \bar{\beta}_k + \tilde{\delta}_{j,dma} + \delta_{dma,month} + \xi_{j,t} \tag{4}
$$

Where first component $\hat{\delta}_{j,t}$ is the combination of product, market[12] fixed effects and $\epsilon_{i,j,m,s}$ is a consumer specific log normal idiosyncratic shock. The second element of the utility function $\gamma_j \cdot p_{j,t}$, corresponds to the disutility derived from paying price $p_{j,t}$ for product j and $\gamma_j \sim \mathcal{N}(\alpha, \sigma)$. $\beta_s^{eff.}$ are symptom-specific sensitivity to beliefs about how effective the drug is. Additionally, preference for package size is also a random variable. This allows for flexible substitution patterns across sizes of packages. Based on this formulation, expected efficacy can be treated as an individual-symptom-level product characteristic. In the standard approach, when quality is unobserved, it is infeasible to separately identify this characteristic from the "unobservable" product fixed effects or the error term, in situations when it is heterogeneous.

IV.B. Aggregation

Based on my market definition, a consumer can enter a market believing they need to alleviate migraine or headache symptoms. I can use the law of total probability and rewrite 3, treating symptom as one of the consumer's characteristics:

$$
\begin{aligned}
u_{i,j,t} = & \underbrace{\hat{\delta}_{j,t} + \gamma_i \cdot p_{j,t}}_{\text{non symptom-specific mean utility}} \\
& + \underbrace{\sum_{k,r} p_{j,k} \cdot z_{i,r} \cdot \gamma_{k,r}^{obs} + \gamma_m \cdot p_j \cdot \mathbf{Pr}_i(s_i = m) + \gamma_h \cdot p_j \cdot \mathbf{Pr}_i(s_i = h)}_{\text{consumer characteristic-specific price sensitivity}} \\
& + \underbrace{\beta_m^{eff.} \cdot \mathbf{E}\left[\textit{Efficacy}_{i,j} \mid m\right] \cdot \mathbf{Pr}_i(s_i = m) + \beta_h^{eff.} \cdot \mathbf{E}\left[\textit{Efficacy}_{i,j} \mid h\right] \cdot \mathbf{Pr}_i(s_i = h)}_{\text{symptom-specific beliefs}} \\
& + \underbrace{\sum_k x_{j,k} \cdot \mathbf{Pr}_i(s_i = m) \cdot \beta_{s,k}^{obs} + \sum_k x_{j,k} \cdot \mathbf{Pr}_i(s_i = h) \cdot \beta_{s,k}^{obs} + \sum_k x_{j,k} \cdot v_{i,k} \cdot \beta_k^{unobs}}_{\text{preferences}} \\
& + \underbrace{\epsilon_{i,j,t}}_{\text{non symptom-specific idiosyncratic shock}}
\end{aligned}
\tag{5}
$$

Under the assumption that variation in preferences, mean utility and logit error is independent of the symptom realization, the utility function follows the standard BLP form (Berry et al., 1995; Nevo, 2000), with a modification. Expected efficacy is not included in the product fixed effect; rather, it is a product characteristic that only appears when interacting with

[12] Defined on the DMA area, month, retailer level

consumer characteristics. This structure reflects the fact that there is no universal valuation of expected efficacy for a product—it only becomes relevant based on the specific symptom the individual is experiencing.

I denote the set of consumers that choose a product j in market t as $\Theta_{jt} = \{i \in \mathcal{I}_t : \mathbb{E}(u_{ijt}) \geq \mathbb{E}(u_{ikt}), \forall k \in \mathcal{J}_t\}$, where $\mathcal{J}_t$ represents the set of products available in market t, and $\mathcal{I}_t$ denotes the set of consumers who shop at least once in market t, normalized to have a mass of one. The market share of product j in market t is given by $s_{jt} = \int_{i \in \Theta_{jt}} di$, and under my assumptions, it is a moment sum of the symptom-specific shares (Berry and Jia, 2010).

IV.C. Identification of expected efficacy

To separately identify the impact of beliefs about efficacy on decision-making and accurately capture substitution patterns, I develop a measure of expected efficacy and integrate it into the choice model. This approach adapts the method by Magnolfi et al. (2024), who substitute observable and unobservable choice parameters in the standard BLP model (Berry et al., 1995) with an aggregate crowdsourced measure of product similarity. While their objective is to estimate an unbiased aggregate measure of substitution between products, my goal is to decompose substitution patterns into components stemming from observable drug characteristics, unobservable beliefs about efficacy, and other unobserved choice parameters.

Given that beliefs about efficacy are inherently unobservable, I designed a survey (Section II) to derive a measure of relative expected efficacy. Using the triplet embedding algorithm (Vankadara et al., 2023), I mapped expected efficacy as coordinates in a two-dimensional space. From the survey, I obtained two-dimensional coordinates that are specific to each product and symptom. The interaction between expected efficacy and the probability of entering the market with particular beliefs inters demand model as a random variable, modeled as a Bernoulli distribution defined by the data. In other words, this approach not only identifies the role of expected efficacy but also quantifies its value, enabling us to observe how heavily consumers rely on perceived efficacy and to measure its impact on demand. This symptom-specific measure, defined at the market (t) and product (j) levels, is given by

$$\sum_{s \in \{m,h\}} \mathbf{Pr}_t(s) \cdot \sum_{d \in \{\text{dim1,dim2}\}} \beta^{\text{eff.},d}_{s,t,j} \cdot \mathbf{E}[\textit{Efficacy}_j \mid s]^d$$

where $\mathbf{Pr}_t(s)$ represents the probability of a consumer seeking relief for symptom s at

time t, and $\beta_{s,t,j}^{\text{eff.},d}$ captures the effect of expected efficacy for symptom s along dimension d.

The primary constraint of this approach is that I have variation in expected efficacy only at the product level. Given that matching drugs to treatments requires specific health-related knowledge, which is not readily available, the extent of this knowledge could lead to varied behavior. A potential solution to this issue is to conduct the triplet embedding component of the survey with a larger and more diverse population group. This would allow for the computation of expected efficacy valuations for different socio-economic groups. Based on the reduced form analysis (Section II.C) and previous research on differences between consumer groups purchasing branded and generic drugs (Bronnenberg et al., 2015; Muthukrishnan et al., 2009; Carrera and Villas-Boas, 2023), these valuations can serve as predictors of biased beliefs. Furthermore, within each socio-demographic cluster, a more sophisticated method could be employed to transform responses to the triplet questions into expected probabilistic efficacy (Diallo and Fürnkranz, 2021), rather than using point estimates. However, due to limitations in the sample size of the experiment, this approach is not feasible in this version of the paper.

IV.D. Estimation

To estimate the model, I utilized a conformant likelihood-based estimator with exogeneity restrictions (CLER) (Grieco et al., 2022). This is a modification of the standard BLP model (Berry et al., 1995) that provides a fast convergence model with valid inference. The quickest convergence is crucial in my case due to utilizing a large data set. Additionally this method accounts for the nearly zero shares of generic products.

IV.D.1. Instruments

The unobservable product-market level effect, denoted as ξ_{jt}, captures a wide range of product characteristics that I do not observe, such as the package color or the product's placement on the shelf. These attributes are likely correlated with prices. To address the correlation between prices and ξ_{jt}, I use instrumental variables. The potential endogeneity of expected efficacy is managed by the fact that expected efficacy is constant across time, therefore time fixed effects control for the endogeneity.

I construct Differentiation IVs as described in Gandhi and Houde (2023). For continuous attributes, such as package size, I calculate the Euclidean distance of each drug from others within the same market (t) along this dimension. For discrete attributes—such as active

components, symptom labels, and brand names—I measure the number of products in the market that share identical characteristics.

IV.D.2. Demographic draws

To account for demographic variations across markets, I generated consumer-specific demographic draws by sampling 10,000 consumers with replacement per unit from the Homescan data, using projection weights provided by NielsenIQ (Kim et al., 2023). Since the data does not specify whether the head of the household or other members made the purchase, to identify the social demographic characteristics of the consumer I randomly selected the first or second member of the household, along with their corresponding age and gender. I also excluded potential buyers younger than 18 or older than 90 years.

V. Demand results

The results of my demand specification indicate that while misbeliefs about expected efficacy significantly influence decision-making, they do not fully account for consumers' preference for more expensive biologically equivalent products. Consistent with prior literature, individuals demonstrate a higher willingness to pay for branded versions over generics. Additionally, my findings show that consumers strongly favor labels corresponding to the symptoms they seek to alleviate when entering the market. The estimated average own-price elasticity in the OTC drug market is -1.22, aligning with previous estimates for OTC analgesics.

V.A. General results and elasticity

Table 11 presents estimated demand parameters. They are aligned with standard economic intuition. For example, individuals with higher income levels show a preference for larger package sizes, as they are less constrained by income, while older consumers tend to be less price-sensitive. Although not all dimensions of the embedding are statistically significant, the likelihood ratio test demonstrates that the model incorporating expected efficacy provides a significantly better fit than the restricted model without expected efficacy.[13] For a randomly chosen month, April, the average own price elasticity across markets and products is -1.22. This number is in line with Liaukonyte (2015) in which median elasticity for OTC analgesics

[13] Some lack of statistical significance can be attributed to high levels of correlation between parameters and the abundance of socio-demographic characteristics.

	Mean Utility	Age	Income	Female	Size HH	Belief Need Migraine Relief	Belief Need Headache Relief	Random Coefficient
Constant	-7.570	0.030	0.042	1.024	-0.195	-0.727	-	0.001
	(14.069)	(0.014)	(0.029)	(0.393)	(0.053)	(0.591)	-	(0.000)
Price	-0.112	0.001	-0.002	-0.013	-0.005	-0.007	0.001	0.000
	(0.040)	(0.000)	(0.001)	(0.017)	(0.003)	(0.018)	(0.072)	(0.000)
Brand	0.895	0.022	0.091	0.064	-	0.298	-	-
	(10.371)	(0.013)	(0.026)	(0.362)	-	(0.387)	-	-
Migraine Label	2.442	-0.031	-0.010	0.033	-	0.495	-	-
	(19.117)	(0.004)	(0.009)	(0.115)	-	(0.379)	-	-
Number Pills	-0.009	-	0.001	0.001	-	0.001	-	0.025
	(10.371)	-	(0.017)	(0.000)	-	(0.002)	-	(0.108)
Embedding Dim 1 (Headache Relief)	-	-	-	-	-	-	6.338	-
	-	-	-	-	-	-	(5.929)	-
Embedding Dim 2 (Headache Relief)	-	-	-	-	-	-	2.224	-
	-	-	-	-	-	-	(3.187)	-
Embedding Dim 1 (Migraine Relief)	-	-	-	-	-	7.881	-	-
	-	-	-	-	-	(6.661)	-	-
Embedding Dim 2 (Migraine Relief)	-	-	-	-	-	-0.359	-	-
	-	-	-	-	-	(1.682)	-	-
Acetaminophen	0.386	-	-	-	-	-	-	-
	(2.299)	-	-	-	-	-	-	-
Acetaminophen + Aspirin + Caffeine	0.227	-	-	-	-	-	-	-
	(11.336)	-	-	-	-	-	-	-
Acetaminophen + Caffeine	1.988	-	-	-	-	-	-	-
	(58.508)	-	-	-	-	-	-	-
Aspirin	-0.312	-	-	-	-	-	-	-
	(3.246)	-	-	-	-	-	-	-
Ibuprofen	-0.329	-	-	-	-	-	-	-
	(13.824)	-	-	-	-	-	-	-

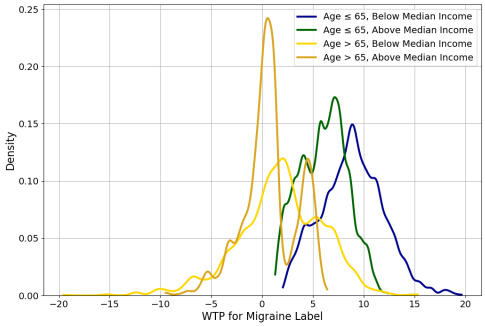
TABLE 5: Regression Results with Mean Utility and Deviations

ranges from -0.8 to -1.3. There is notable variation in own price elasticity across different package sizes. Demand for small packages tends to be inelastic, whereas demand for larger packages averages -2.01. This elasticity is slightly lower than that of general consumption goods, such as ready-to-eat cereals, which have an elasticity of -2.29 (Döpper et al., 2024). Several factors could explain the difference between small and large packages. First, price sensitivity along with other preferences are not the only factors of choice, for instance, some consumers may opt for the cheapest drug available based on their preferences or may not need a large package if they do not consume the drug frequently enough to use it before its expiration date. Second, small packages might be perceived as immediate remedies and, therefore consumed no matter what the price. More granular own price elasticity of distribution is discussed in online Appendix Figure 27.

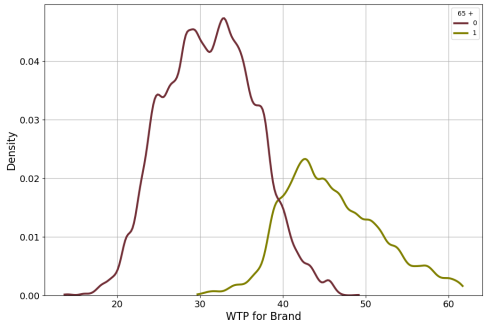
V.B. Brand and label preferences

The willingness to pay (WTP) for the migraine label averages $5.50, which represents 58% of the average drug price. Individuals who are younger and have income below the median have the highest wtp for the migraine label (see Figure 9a). This WTP reflects preferences that extend beyond the belief that products with the migraine label are more effective than

those with the headache label in alleviating migraine symptoms. On average, the WTP for branded products is $36.40 (Figure 9b). Consumers over the age of 65 demonstrate a greater willingness to pay for branded products compared to younger consumers. Generally, individuals exhibit lower price elasticity when purchasing branded products relative to generic versions, branded products being less elastic (see Online Appendix Table 16).



(a) Willingness to Pay for the Migraine Label



(b) Willingness to Pay for the Brand Label

Figure 9: Distribution of Willingness to Pay for Different Labels, Individual Level

V.C. Role of expected efficacy

Comparing total expected efficacy with front-label preferences reveals key insights into consumer decision-making.

Active Components	Brand	Migraine Label	$\beta_m^{eff.} * \mathbf{E}[Efficacy\|m.]$	$\beta_h^{eff.} * \mathbf{E}[Efficacy\|h.]$
Acetaminophen	0	0	-0.81	-1.02
	1	0	-0.86	-1.13
Acetaminophen + Aspirin + Caffeine	0	0	-0.79	-1.06
	1	0	-1.08	-1.36
		1	-0.92	-0.87
Acetaminophen + Aspirin	0	0	-0.89	-1.05
	1	0	-1.10	-1.13
Aspirin	0	0	-1.05	-1.29
	1	0	-0.95	-1.03
Ibuprofen	0	0	12.31	15.41
	1	0	-0.54	-0.78
		1	-0.65	-0.80
Naproxen Sodium	0	0	-0.19	-0.64
	1	0	-0.58	-0.83

TABLE 6: Expected Efficacy by Active Components, Brand Label, and Migraine Label

As shown in Table 6, consumers seeking migraine or headache relief assign higher utility to biologically equivalent drugs containing acetaminophen+aspirin+caffeine when labeled for migraines. Their greater willingness to pay for the migraine label suggests a strong preference for the migraine-branded version over its identical headache-labeled counterpart. This valuation gap is especially pronounced among migraine sufferers, highlighting the difficulty in recognizing biological equivalence across labels—a factor that firms can exploit. The findings also indicate that expected efficacy dampens brand preference. Unlike biological equivalence across labels within the same brand, branded drugs tend to have lower expected efficacy than generics, aligning with the observed substitution patterns in Section . While some consumers recognize generics as equally effective, brand loyalty, inertia, or healthcare provider influence still drive choices. This suggests that policies promoting generics should address behavioral barriers beyond efficacy awareness.

VI. Supply

In this section, I introduce my supply model. To better capture the patterns that I observe in the data I modify the standard Nash-Bertrand model. Based on the modified Nash-Bertrand I recover marginal cost on the product - market level. Per pill, marginal costs are positively correlated with the larger number of active components and the number of grams per active component.

VI.A. Set up

There are R retailers, each offering a subset of products, $R^{r,dma}$, which varies across months and designated market areas that form a market (t). Retailers sell both branded and private-label headache and migraine medications, where manufacturers are identified by the brand name on the package. I assume that retailers are price takers (Villas-Boas, 2007; Döpper et al., 2024). For packages of 80+ units, manufacturers set prices following pure Nash-Bertrand equilibrium in the horizontal competition across markets (T). For smaller package sizes, manufacturers use a simplification rule, pricing them as a fixed fraction of the larger size's price (Cohen et al., 2021; Laitinen, 2009). Online Appendix IX.B.1 provides a detailed discussion of these assumptions. In summary, the pricing patterns of smaller packages relative to larger ones cannot be fully explained by the standard Nash-Bertrand model, whether through supply or demand shock responses. For instance, under a pure demand shock, the prices of smaller packages do not adjust in response to changes in the

pricing of larger packages. To simplify notation, I assume a single manufacturer and a single time period in the analysis.

The manufacturer's profit function across all markets (T) is given by:

$$\mathbf{\Pi}^f = \sum_{t \in T^f} \sum_{j \in \mu^{f,t}} (p_j^{t,f} - mc_j^{t,f}) * s_j^{t,f}(\mathbf{p}) \tag{6}$$

Solving the manufacturer's profit maximization and aggregating equation 6 across all products supplied to a market (t), I recover marginal costs from the following condition:

$$(\mathbf{p}^{t,f} - \mathbf{mc}^{t,f}) = -(\mathbf{\Omega})^{-1} \cdot \mathbf{S}^{\mathbf{t,f}}(\mathbf{p}) \tag{7}$$

Where Ω is the own- and cross-price elasticity matrix of substitution. The problem is solved simultaneously and independently for all manufacturers and markets.

This setup relies on two key assumptions: (1) Retailers are price takers, meaning there is no double marginalization; (2) Manufacturers follow Nash-Bertrand pricing for packages of 80+ units, while smaller packages are priced using the simplification rule.

The assumption that pharmaceutical companies act as price setters is justified by the high market concentration in this sector. Additionally, given data limitations, this model is more appropriate than a linear model with double marginalization, as I cannot separately identify retailer costs from manufacturer prices in that framework. Without this separation, I would have to impose the strong assumption that manufacturers do not adjust their prices under counterfactual scenarios—an issue my model avoids. By focusing on information treatments that shift substitution patterns across products from the same manufacturer, this approach offers a more realistic representation of pricing dynamics. The second assumption stems from the misalignment between observed prices and the properties expected under Nash-Bertrand pricing. Online Appendix IX.B.1 presents simulations of Nash-Bertrand responses and tests their validity against the data.

This model has several limitations. First, aggregating demand at the monthly level prevents capturing promotional activity. Since I use average monthly prices, both promotional and non-promotional weeks are combined, making it difficult to identify which products experience deeper or more frequent promotions. This likely underestimates the impact of substitution pattern changes on consumer welfare (Sinitsyn, 2016). Second, some retailers in the dataset are large chains, where a bargaining model might be a more natural fit. However, implementing such a model would require additional assumptions, such as simultaneous bargaining and separate negotiations for each product, which may not be more realistic

than the current approach. Finally, the model assumes that in the counterfactual scenarios, retailers do not adjust their assortment. In reality, if consumers switch to the cheapest biologically equivalent product, some retailers might discontinue certain products to avoid profit losses.

VI.B. Estimation

I recover marginal costs based on Equation 7, following the approach of (Morrow and Skerlos, 2011). The distribution of the recovered marginal costs is shown in Figure 10.

To account for idiosyncratic cost shocks across counterfactuals, where there are no meaningful input costs, I decompose the recovered marginal costs into per-pill costs of the active components, brand fixed effects, and market fixed effects. The brand fixed effects can be interpreted as costs related to packaging, advertising, shipment, etc., while the market fixed effects reflect the costs associated with shelf placement. Table 17 from the Online Appendix presents the results of this decomposition. Products with two active components have lower per-pill costs than those with three components. The higher per-pill costs of acetaminophen compared to acetaminophen, aspirin, and caffeine can be attributed to the dosage, with acetaminophen at 500 mg and the combination drug at 250 mg.

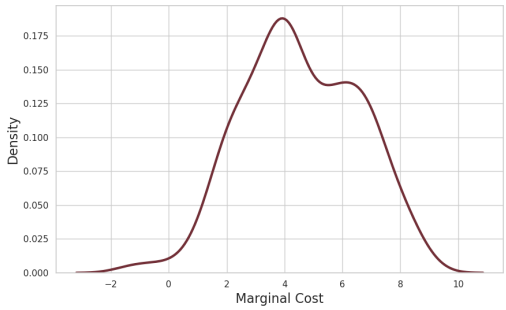


Figure 10: Marginal Costs Distribution, Product Market Level

VII. Equilibrium effects of counterfactual information provision

The primary goal of information treatments is to reduce the information gap between consumers and pharmaceutical companies. Ideally, their effectiveness in mitigating information distortion should be evaluated against a scenario in which consumers possess full information. However, a key challenge arises due to the absence of data on how consumers

value drug efficacy under such conditions. [14] A reasonable assumption is that, with full information, consumers' expected efficacy would align with clinical trial results, which assess efficacy based solely on active ingredients, excluding the influence of labeling or packaging. However, the medical literature lacks a detailed analysis of the relative efficacy among active components, making it infeasible to incorporate a "rational" valuation of efficacy into the demand model (Moore and Derry, 2012).

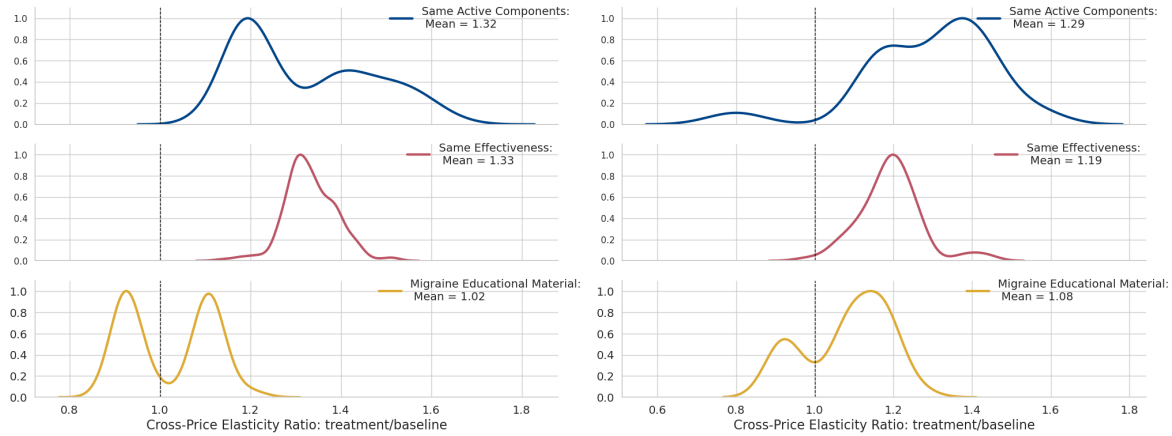
Given this limitation, I assume that biologically equivalent drugs should be closer substitutes than products with different active components, and study whether treatments increase substitution across biologically equivalent drugs. By incorporating new perceived efficacy values into demand I find that demand-side responses alone increase the cross-price elasticity of substitution between biologically equivalent products, which is in line with the reduced form evidence (V.C) that shows that information treatments increase substitutability between biologically equivalent products in the perceived efficacy dimension. When combined with the supply-side response, all counterfactual scenarios yield welfare gains, measured by consumer savings and changes in surplus.

VII.A. Impact of Information Provision Without Price Change

VII.A.1. Change in Substitution Patterns

If consumers treated biologically equivalent products as identical—consistent with FDA standards—they would function as perfect substitutes from a demand perspective. In practice, however, consumer preferences for product attributes prevent full substitutability. A treatment is considered successful if it increases the substitutability between biologically equivalent products. To evaluate this, I computed the ratio of each treatment's cross-price elasticity to the baseline at the individual product level. Specifically, I analyzed the elasticity between biologically equivalent products within the same brand but under different labels (Figure 11a) and the elasticity between branded products and their generic counterparts (Figure 11b), relative to their respective baseline values.

[14] There is no experimental design that can simulate a full-information scenario.



(a) Excedrin Products

(b) Branded to Generic

Figure 11: Cross price elasticity analysis for Excedrin and branded-to-generic products

All treatments result in statistically different average treatment effects on the cross-price elasticity of substitution between biologically equivalent products (see Online Appendix, Table 13a and Table 13b). For biologically identical products within the same brand, treatments featuring front-label statements on equivalent effectiveness (red) increase cross-price elasticity by 33%, while those highlighting shared active components (blue) raise it by 32%. However, these effects vary across consumers, with some experiencing little to no impact—or even a negative effect. The informational brochure (yellow) shows no significant average effect. All treatments increase cross-price substitution from branded to generic products, with front-label statements on shared active components yielding the highest average increase in cross-price elasticity, though some consumers respond negatively. In contrast, front-label statements on equivalent effectiveness are the most effective overall, raising cross-price elasticity by 19%.

These results underscore the importance of information type in shaping consumer responses. Direct references to efficacy provide the clearest and most accessible information, generating a strong and consistent positive effect for a broad consumer segment. In contrast, information about active components is more complex, requiring consumers to connect it to efficacy, benefiting some while confusing others due to self-selection effects. The brochure treatment, which demands the highest cognitive effort, proves the least effective, as it requires consumers to actively engage before making decisions. The relatively smaller impact of these treatments on brand-to-generic substitution—compared to their effect within the same brand—suggests that many consumers were already aware of brand-generic equivalence but

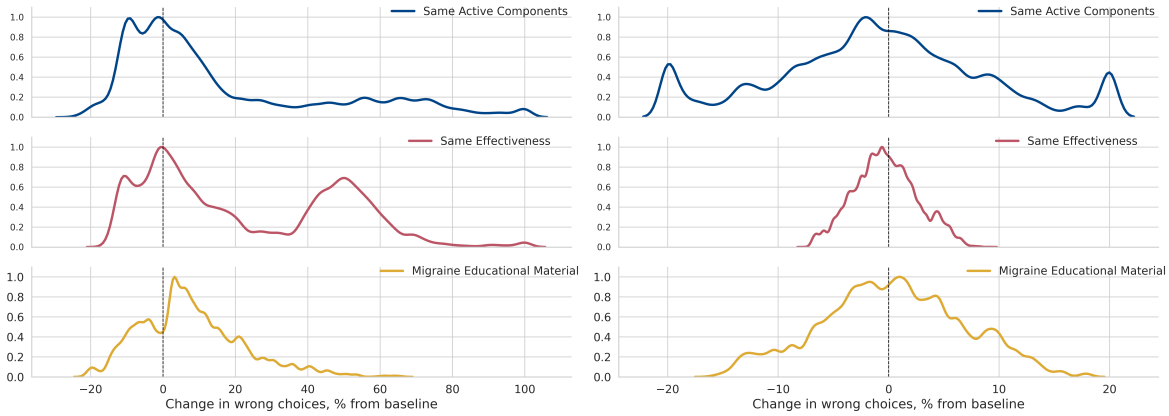
were less familiar with biological equivalence within a brand, prompting a stronger response in that case. This interpretation aligns with the reduced-form evidence.

VII.A.2. Choices and "Wrong" choices

A change in substitution patterns does not necessarily translate into more informed decisions that lead to savings, which is the ultimate goal of my information treatments. Without separating choices on correct and wrong I observe that on average an absolute change in purchase probabilities across counterfactuals is 16.9% (Online appendix Figure 26). A key question is whether these choice shifts translate into fewer wrong choices, defined as purchasing a costlier biologically identical product when a cheaper alternative within the same brand and size is available. The information treatments aim to reduce this information gap, which, in theory, should lead to a decline in wrong choices. However, actual decisions depend not only on perceived efficacy but also on preferences for product attributes. As a result, greater awareness of biological equivalence does not always directly reduce wrong choices, making it essential to assess how new beliefs interact with consumer preferences.

To quantify this, I calculate both the absolute and conditional changes in wrong choices. The absolute change measures the overall shift in wrong choices, but it does not account for differences in product consideration. A consumer may appear to make more wrong choices simply because they are purchasing more products overall, rather than due to a change in decision-making quality. Additionally, an increase in absolute wrong choices could result from broader market engagement, where consumers who previously opted out of the market now make purchases, some of which may be classified as wrong. To address these limitations, I compute the conditional change, which accounts for the probability that a consumer considers a biologically identical product pair before making a choice. This measure isolates the impact of information treatments on decision-making by ensuring that observed changes in wrong choices are not driven by shifts in overall market participation or product exposure.

Figure 12a presents the absolute change, capturing both intensive and extensive margins—an increase in wrong choices can arise from either a general increase in product consumption (e.g., more consumers purchasing Excedrin) or a shift toward more expensive option within the Excedrin product line. After conditioning on product consideration (Figure 12b), the density plots remain highly heterogeneous but skew negative, indicating an average decrease in the probability of making a wrong choice when accounting for consumer choice sets.



(a) Absolute Change (b) Relative Change

Figure 12: Comparison of absolute and relative changes in wrong choices

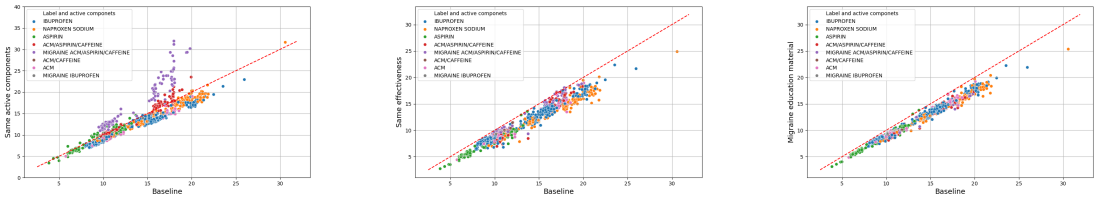
To further examine the heterogeneity in wrong choices, I analyze how increased awareness of biological equivalence interacts with consumer preferences for product attributes. Specifically, I decompose the probability of making a wrong choice based on whether the more expensive product carries a migraine label and the consumer's willingness to pay for it. The findings, detailed in Online Appendix IX.E.1, reveal that consumers with a higher willingness to pay for migraine-labeled products are more likely to make wrong choices when a cheaper biologically identical alternative exists. This effect is particularly strong for information treatments that enhance substitution between biologically identical products, suggesting that labeling influences consumer choices beyond efficacy beliefs.

VII.B. General equilibrium response

VII.B.1. Supply side response

Since markups are a function of elasticities, changes in substitution patterns (Online Appendix IX.F) trigger supply-side reactions, though the nature of these reactions is ambiguous. On one hand, a better understanding of biological equivalence makes products more interchangeable, thereby increasing competition between molecules and leading to lower prices. On the other hand, consumers who remain "loyal" to the more expensive biologically equivalent product after receiving information about equivalence have a higher willingness to pay for the product's attributes. This allows companies to exploit loyalty to the label and widen the gap between this product and its biologically equivalent analog (Grabowski and

Vernon, 1992) and focus on this attribute as the main dimension for product differentiation.



(a) Same Active Components (b) Same Effectiveness (c) Educational Material

Figure 13: Comparison of Baseline Prices with Counterfactual Simulations

To determine which of these forces dominates in my data, I apply the model from the supply section (Equation 7), incorporating costs recovered from the baseline optimization (Section VI) and demand for each information treatment. In the treatment where individuals receive information about the same active components (Figure 13a), there is a heterogeneous response to the change in substitution elasticities, with downward price pressure generally prevailing. Notably, prices for Excedrin Migraine increase, while those for Excedrin Extra Strength decrease, leading to a 17-fold increase in the average price difference between these products. This suggests that the second mechanism—more pronounced label loyalty—is dominant. Consumers who continue to buy Excedrin Migraine are likely those with a high willingness to pay specifically for the migraine label within the Excedrin product line. For the "Same Effectiveness" label treatment (Figure 13b) and the educational material treatment (Figure 13c), downward pressure on prices is the dominant force. However, in the "Same Effectiveness" treatment, the average price difference between biologically equivalent products still increases by 3 times compared to the baseline, highlighting the existence of "loyal exploitation" in strategic pricing.

While increased substitution between products with different molecules or brands generally drives prices down, greater cross-price substitution between biologically identical products under the same brand can lead to strategic pricing adjustments. Figure 14 illustrates shifts in pricing policies for biologically equivalent products with different labels. In the baseline, there is an equal likelihood (50%) that a store prices Excedrin Migraine higher than Excedrin Extra Strength, or vice versa. However, when information treatments enhance substitutability between biologically identical products, efficacy becomes a less relevant differentiating factor since the products are identical in that regard. Instead, consumer decisions are increasingly driven by willingness to pay for the label itself. As a result, treatments that modify labels and increase substitution within the same brand lead to a

shift in pricing strategy: the probability of Excedrin Migraine being priced higher moves toward 1. Consumers consistently face higher prices for migraine-labeled products within the same biologically equivalent pair. In contrast, the treatment providing educational materials—without altering substitution patterns—has minimal impact on firms' pricing strategies for biologically equivalent products.

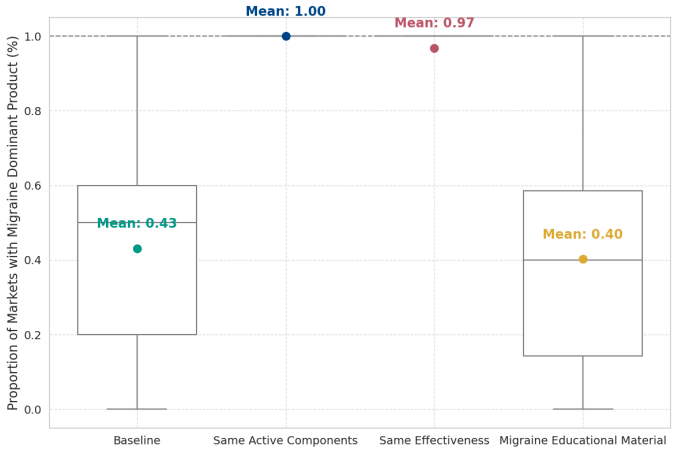


Figure 14: Box Plot: Probability that Product with Migraine Label Has Higher Price in Biologically Identical Pair with Different Labels

VII.B.2. Welfare

The standard approach to evaluating welfare changes in the counterfactual analysis under ex-ante biased information follows (Allcott, 2013), incorporating a "rational" efficacy valuation as the expected efficacy while holding other utility parameters constant. However, this method is infeasible in my case due to the lack of definition or numerical estimates for rational efficacy valuation. To address this, I conduct a two-part welfare analysis. First, I assume that labeling has no intrinsic value and that only consumption matters — meaning that as long as the pill provides relief, the label itself does not contribute to consumer welfare. Under this assumption, the only factor that increases consumer surplus is a reduction in expenditure. Second, I consider a scenario where ex-post value creation from labeling exists. This reflects the possibility that some labels enhance perceived symptom relief, even when, from a medical perspective, they should not. This approach uses a standard formula of change in consumer surplus conditional on entering the market but fixes the perceived efficacy valuation on the counterfactual level:

$$\Delta \text{CW}^m(x) = \text{CW}^m(x) - \text{CW}^m(\text{baseline (x)}) \tag{8}$$

Where:

- $x \in \{\text{Same Active Components; Same Effectiveness; Migraine Education Material}\}$
- $\text{CW}^m(\text{baseline}) = \sum_j \left\{ \int_{\Theta_{jm}^{(\text{baseline})}} \frac{1}{\alpha_i} \left(\delta_{ijm} - \alpha_i p_{jm}^{\text{baseline}} + \beta_{s_i}^{\text{efficacy}} \cdot \mathbf{E}\left[\textit{Efficacy}_j \mid s_i \right]^{(x)} \right) di \right\}$
- $\text{CW}^m(x) = \sum_j \left\{ \int_{\Theta_{jm}^{(x)}} \frac{1}{\alpha_i} \left(\delta_{ijm} - \alpha_i p_{jm}^{x} + \beta_{s_i}^{\text{efficacy}} \cdot \mathbf{E}\left[\textit{Efficacy}_j \mid s_i \right]^{(x)} \right) di \right\}$

Change in the choice patterns along with the downward pressure on prices, in equilibrium, results in savings for all consumers in all markets Figure 15. Analogous to the analysis of the impact of the information treatments without price response, I compute the savings amount conditional on market entry. This allows me to separate the extensive margin of the information treatment—where individuals enter the market due to new information and make new choices—from the intensive margin, where individuals save due to changes in prices and existing choices.

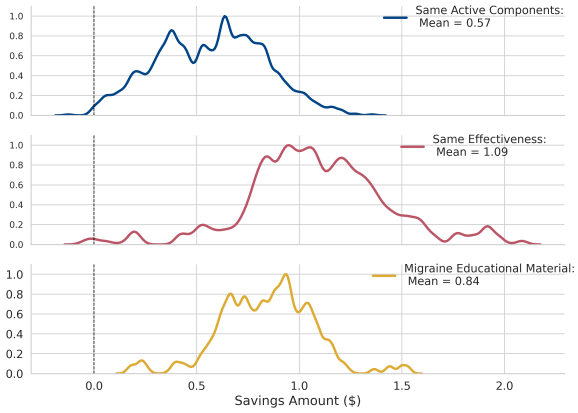


Figure 15: Individual Level Savings, New Equilibrium

The results from equation 8 are presented in Figure 16. The welfare results align with other findings from the information treatments: all treatments result in higher per capita welfare levels compared to the baseline, with the provision of labels about the same effectiveness yielding the highest welfare benefits. Assuming that "ex-post" efficacy levels reflect the information treatments, the welfare increase is $1.20, driven by consumer savings and the fact that consumers were under-consuming products with higher "ex-post" efficacy.

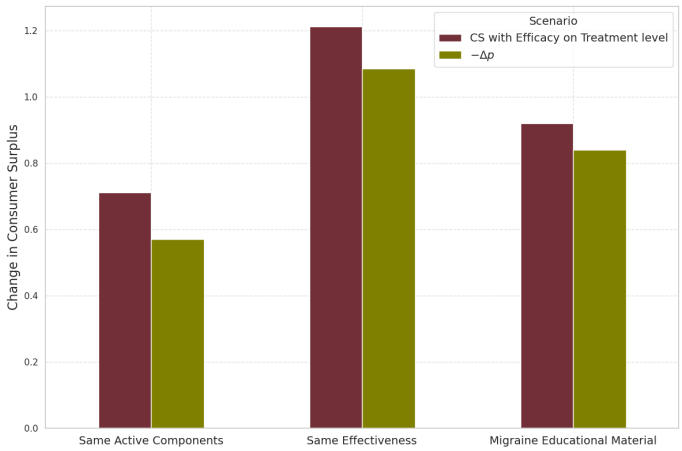


Figure 16: Chang Controlling for Expected Efficacy Levels

VIII. Discussion and conclusion

This paper examines market failure in the OTC drug market, focusing on consumer misperceptions of efficacy. Consumers fail to recognize biologically equivalent products—whether within the same brand or across brand-generic pairs—as equally effective, instead relying on symptom labels to guide their choices. Firms exploit this artificial differentiation to extract additional markups. To assess whether improved information provision can correct these biases and enhance consumer welfare, I propose three counterfactual scenarios that increase transparency about relative efficacy.

To measure market response, I adopt a novel methodology integrating randomized controlled trial (RCT) results into a structural model that separately identifies preferences based on observable drug characteristics, inherently unobservable beliefs about efficacy, and other unobserved choice parameters. To capture firms' pricing responses, I adapt the Nash-Bertrand model, assuming firms first set prices for specific market segments and then apply them market-wide.

All information treatments lead to new market equilibria where consumers experience cost savings and improved welfare. Increased transparency in efficacy not only facilitates informed decision-making but also enhances competition between biologically equivalent drugs and products with different active ingredients, leading to an overall price reduction. The most effective intervention, the "Same Effectiveness" treatment, highlights biological equivalence based on consumer experience. These findings align with (Carrera and Villas-Boas, 2023), who demonstrated that effectiveness similarity combined with consumer

experience is more persuasive than active ingredient information in guiding generic adoption.

However, while increased transparency benefits consumers, it also introduces potential risks. Firms can exploit consumer preferences for symptom-specific labels, increasing price discrimination. My counterfactuals show that firms adjust prices based on consumers' willingness to pay for labels such as "migraine relief." Since young, lower-income consumers demonstrate the highest willingness to pay for such labels, this structural shift could further disadvantage vulnerable populations. To counteract these effects, additional interventions—such as behind-the-counter markets, pharmacist consultations, or price regulations—may be necessary.

The methodological approach of decomposing substitution patterns into preferences for observable attributes and ex-ante unobservable but RCT-identified efficacy beliefs can be extended to various settings where quality is both a key decision factor and initially unobservable. Moreover, this framework enables testing consumer responses to changes in quality valuation due to information treatments.

Several key directions for future research emerge. First, methodological advances should shift from aggregate product-level measures of perceived quality to more heterogeneous consumer valuations. Second, while this model focuses on static information acquisition, real-world consumers may learn over time, affecting pricing strategies and market structure. Third, firm promotional activities could play a strategic role in preventing consumers from learning about biological equivalence, necessitating a dynamic model of firm promotion and consumer learning. Finally, a more precise estimation of welfare requires further examination of how perceived efficacy aligns with ex-post medical outcomes. The broader debate remains: should pharmaceutical market welfare be measured by objective medical effectiveness or by consumers' subjective experience?

References

AHMED, F. (2015): "Headache Disorders: Differentiating and Managing the Common Subtypes," *Journal of Headache & Pain*, 16, 4.

ALLCOTT, H. (2013): "The welfare effects of misperceived product costs: Data and calibrations from the automobile market," *American Economic Journal: Economic Policy*, 5.

ATAL, J. P., J. I. CUESTA, AND M. SÆTHRE (2022): "Quality Regulation and Competi-

tion: Evidence from Pharmaceutical Markets," *Review of Economic Studies*, revise and resubmit.

BARAHONA, N., C. OTERO, AND S. OTERO (2023): "Equilibrium Effects of Food Labeling Policies," *Econometrica*.

BERNDT, E. R., I. M. COCKBURN, Z. GRILICHES, T. E. KEELER, AND M. N. BAILY (1996): "Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Antidepressant Drugs," *Brookings Papers on Economic Activity. Microeconomics*, 1996.

BERRY, S. AND P. JIA (2010): "Tracing the Woes: An Empirical Analysis of the Airline Industry," *American Economic Journal: Microeconomics*, 2, 1–43.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63.

BLEWETT, L. A., J. A. R. DREW, D. BACKMAN, A. CHEN, G. COOPER, M. SCHOUWEILER, S. RICHARDS, AND M. WESTBERRY (2023): "IPUMS Health Surveys: Medical Expenditure Panel Survey, Version 2.3 [dataset]," .

BRONNENBERG, B. J., J. P. DUBÉ, M. GENTZKOW, AND J. M. SHAPIRO (2015): "Do pharmacists buy bayer? Informed shoppers and the brand premium," *Quarterly Journal of Economics*, 130.

BUGNI, F. A., I. A. CANAY, AND A. M. SHAIKH (2019): "Inference under covariate-adaptive randomization with multiple treatments," *Quantitative Economics*, 10.

CARRERA, M. AND S. VILLAS-BOAS (2023): "Generic Aversion and Observational Learning in the Over-the-Counter Drug Market," *American Economic Journal: Applied Economics*, 15.

CATLIN, J. R. AND E. P. BRASS (2018): "The Effectiveness of Nonprescription Drug Labels in the United States: Insights from Recent Research and Opportunities for the Future," *Pharmacy (Basel)*, 6, 119.

COHEN, M., G. PERAKIS, AND R. PINDYCK (2021): "A Simple Rule for Pricing with Limited Knowledge of Demand," *Management Science*, cited by 27.

CRAWFORD, G. S. AND M. SHUM (2005): "Uncertainty and learning in pharmaceutical demand," *Econometrica*, 73.

DELLAVIGNA, S. AND M. GENTZKOW (2019): "Uniform Pricing in U.S. Retail Chains," *The Quarterly Journal of Economics*, 134, 2011–2084.

DIALLO, A. AND J. FÜRNKRANZ (2021): "Elliptical Ordinal Embedding," in *Proceedings of the IFIP Working Conference on Computer Science and Mathematics*, Springer.

DRANOVE, D. AND G. Z. JIN (2010): "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 48, 935–963.

DRANOVE, D., D. KESSLER, M. MCCLELLAN, AND M. SATTERTHWAITE (2003): "Is More Information Better? The Effects of Report Cards on Health Care Providers," *Journal of Political Economy*, 111, 555–588.

DÖPPER, H., A. MACKAY, J. STIEBALE, AND N. MILLER (2024): "Rising Markups and the Role of Consumer Preferences," Resubmitted to *Journal of Political Economy*.

ELLISON, G. AND S. F. ELLISON (2011): "Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbents Prior to Patent Expiration," *American Economic Journal: Microeconomics*, 3, 1–36.

ELLISON, G. AND A. WOLITZKY (2012): "A Search Cost Model of Obfuscation," *The RAND Journal of Economics*, 43, 417–441.

GANDHI, A. AND J.-F. HOUDE (2023): "Measuring Substitution Patterns in Differentiated-Products Industries," NBER Working Paper w26375, National Bureau of Economic Research.

GRABOWSKI, H. G. AND J. M. VERNON (1992): "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals After the 1984 Drug Act," *The Journal of Law and Economics*, 35, 331–350.

GREENSTONE, M., P. OYER, AND A. VISSING-JORGENSEN (2006): "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments," *The Quarterly Journal of Economics*, 121, 399–460.

GRIECO, P. L., C. MURRY, J. PINKSE, AND S. SAGL (2022): "Conformant and Efficient Estimation of Discrete Choice Demand Models," Unpublished Manuscript.

HANDEL, B. R. (2013): "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review*, 103, 2643–2682.

HARBEN, A. L., D. A. KASHY, S. ESFAHANIAN, L. LIU, L. BIX, AND M. W. BECKER (2021): "Using change detection to objectively evaluate whether novel over-the-counter drug labels can increase attention to critical health information among older adults," *Cognitive Research: Principles and Implications*, 6, 40, published online 2021 May 26.

HOUDE, S. (2018): "Bunching with the Stars: How Firms Respond to Environmental Certification," *SSRN Electronic Journal*.

JIN, G. Z. AND P. LESLIE (2003): "The Effect of Information on Product Quality: Evidence From Restaurant Hygiene Grade Cards," *The Quarterly Journal of Economics*, 118, 409–451.

KIESEL, K. AND S. VILLAS-BOAS (2013): "Can Information Costs Affect Consumer Choice? Nutritional Labels in a Supermarket Experiment," *International Journal of Industrial Organization*, 31, 153–163.

KIM, M., R. MANSLEY, M. REMER, AND M. WEINBERG (2023): "Price-Fixing Allegations in the Canned Tuna Industry: A Look at the Data," *Antitrust Bulletin*, 68, 154–163.

KINGMA, D. P. AND J. BA (2014): "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*.

KNOBLAUCH, K. AND L. T. MALONEY (2012): "Maximum Likelihood Conjoint Measurement," in *Modeling Psychophysical Data in R*, Springer, vol. 32 of *Use R!*, 229–256, access provided by Northwestern University.

KÜNSTLE, D.-E., U. VON LUXBURG, AND F. A. WICHMANN (2022): "Estimating the perceived dimension of psychophysical stimuli using triplet accuracy and hypothesis testing," *Journal of Vision*, 22, 5.

LAITINEN, E. K. (2009): "Complexities of Pricing Decisions: Evidence from Finnish Firms," *Journal of Business Economics*.

LIAUKONYTE, J. (2015): "Is Comparative Advertising an Active Ingredient in the Market for Pain Relief?" *Working Paper, Cornell University*.

LUCE, R. D. AND J. W. TUKEY (1964): "Simultaneous conjoint measurement: A new type of fundamental measurement," *Journal of Mathematical Psychology*, 1, 1–27.

MAGNOLFI, L., J. MCCLURE, AND A. SORENSEN (2024): "Triplet Embeddings for Demand Estimation," *AEJ: Microeconomics*, extended Abstract in EC 22: "Embeddings and Distance-based Demand for Differentiated Products".

MAĆKOWIAK, B., F. MATĚJKA, AND M. WIEDERHOLT (2023): "Rational Inattention: A Review," *Journal of Economic Literature*, 61, 226–273.

MOORE, R. A. AND C. DERRY (2012): "Efficacy of OTC analgesics," *International Journal of Clinical Practice*.

MORROW, W. R. AND S. J. SKERLOS (2011): "Fixed-Point Approaches to Computing Bertrand-Nash Equilibrium Prices Under Mixed-Logit Demand," *Operations Research*, 59, 328–345.

MUTHUKRISHNAN, A. V., L. WATHIEU, AND A. J. XU (2009): "Ambiguity aversion and the preference for established brands," *Management Science*, 55, 1933–1941.

NEVO, A. (2000): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9, 513–548.

ROE, B. E., M. F. TEISL, AND C. R. DEANS (2014): "The Economics of Voluntary versus Mandatory Labels," *Annual Review of Resource Economics*, 6, 407–427.

SINITSYN, M. (2016): "Managing price promotions within a product line," *Marketing Science*, 35, 304–318.

STEWART, W. F., R. B. LIPTON, D. D. CELENTANO, AND M. L. REED (1992): "Prevalence of migraine headache in the United States. Relation to age, income, race, and other sociodemographic factors," *JAMA*, 267, 64–69.

TERADA, Y. AND U. VON LUXBURG (2014): "Local Ordinal Embedding," *Journal of Machine Learning Research*, 15, 746–754.

TODD, P. E. AND K. I. WOLPIN (2023): "The Best of Both Worlds: Combining Randomized Controlled Trials with Structural Modeling," *Journal of Economic Literature*, 61, 41–85.

VANKADARA, L. C., M. LOHAUS, S. HAGHIRI, F. U. WAHAB, AND U. VON LUXBURG (2023): "Insights into Ordinal Embedding Algorithms: A Systematic Evaluation," *Journal of Machine Learning Research*.

VATTER, B. (2021): "Quality Disclosure and Regulation: Scoring Design in Medicare Advantage," Working Paper.

VILLAS-BOAS, S. B. (2007): "Vertical relationships between manufacturers and retailers: Inference with limited data," *Review of Economic Studies*, 74, 651–676.

WORLD HEALTH ORGANIZATION (2024): "Headache Disorders: How Common Are Headaches?" *WHO report*, accessed: 2024-09-18.

WOŹNIAK-HOLECKA, J., M. GRAJEK, K. SIWOZAD, K. MAZGAJ, AND E. CZECH (2012): "Consumer behavior in OTC medicines market," *Przeglad Epidemiologiczny*, 66, 189–95, free article.

ZHU, C., R. A. LOPEZ, AND X. LIU (2015): "Information Cost and Consumer Choices of Healthy Foods," *American Journal of Agricultural Economics*, 98, 41–53.

IX. Appendix

IX.A. Additional information RCT

IX.A.1. Probability of entering the market with the belief to have headache/migraine

Retailer/DMA/month-level probability of entering the market with the belief that you need to search for the drug that elevates migraine symptoms is computed based on the social-demographic characteristics of respondents and the following questions:

- Do you believe that migraines and general headaches are the same condition?
  Yes/No/Maybe
- When you experience headache symptoms, do you have a:
  Migraine / General headache/ Both/ Not sure
- How frequently have you purchased a drug to relieve migraine symptoms?
  Weekly/ Monthly/ Annually/ Never

- How frequently have you purchased a drug to relieve headache symptoms? Weekly/Monthly/Annually/Never

The first two questions divide respondents into two groups: 1) those who can differentiate between migraine and headache symptoms—these individuals answered "No" to the first question and did not indicate that they experience both migraine and headache symptoms at the same time or are unsure which one they are experiencing; and 2) those who cannot differentiate between the two. For individuals in the first group, I can confidently state that their responses about the frequency of purchasing medication for headache or migraine relief are consistent and would not change if asked again. However, I do not have this level of certainty for the second group. Therefore, I randomly assign them to the group that enters the market believing they need to alleviate either migraine or headache symptoms.

If a respondent can differentiate between migraine and headache symptoms and answered that when they experience a headache symptom, they are having a migraine, I assume that when they buy a drug to relieve migraine symptoms or headache symptoms, in both cases they seek for drug to alleviate migraine symptoms.

To compute the probability of entering the market every month to alleviate migraine or headache symptoms, I first calculate the number of drugs purchased for each symptom at the monthly level and then derive the probability. The number of drugs purchased within a month to relieve migraines, and similarly for headaches, is based on the last two questions. For the first group—those who can differentiate between migraine and headache symptoms—the quantity is calculated as follows:

$$group1_m = w_m * 4 + m_m + a_m/12 \tag{9}$$

Where:

- w_m - number of respondents that purchase drugs to relieve migraine symptoms weekly
- m_m -number of respondents that purchase drugs to relieve migraine symptoms monthly
- a_m - number of respondents that purchase drugs to relieve migraine symptoms annually

For the second group, I sum the answers for the two last questions on the individual level using the above formula 9. Then, I randomly assign this individual to either the group

that purchases drugs to relieve migraine symptoms or to the group that purchases drugs to relieve headache symptoms.

The final probability of entering the market with the belief that you need a drug that alleviates migraine symptoms is:

$$\mathbf{Pr}(\textit{enter market to buy a drug to relief migraine symptom}) = \frac{group1_m + group2_m}{group1_m + group2_m + group1_h + group2_h} \tag{10}$$

Where:

- $group1_m$ - number of drugs purchased by group 1 within a month to relieve migraine
- $group2_m$ - number of drugs purchased by group 2 within a month to relieve migraine

To account for the potential social-demographic differences in the prevalence of migraines across populations, I've divided respondents into those above and below the median income and above and below age 65. The computation results are presented in Table 3. The probabilities I got resemble the medical literature that claims that low-income populations and younger people have higher chances of suffering from migraine pain and/or self-diagnosing themselves with migraine (Stewart et al., 1992).

Using these probabilities, I estimate the probability of entering the market for each individual in the socio-demographic draws' data with the belief that an individual needs a drug to relieve migraine symptoms or headaches.

IX.A.2. Triplet embedding theory

The goal of my triplet embedding is to map responses about relative effectiveness into a multi-dimensional space, where the distance between points represents relative effectiveness. This multi-dimensionality does not impose assumptions about the number or nature of factors that define relative expected efficacy; instead, it captures the inherent complexity of judgments about relative efficacy.

There are various embedding methods available today, each suited to different needs, such as high-dimensional or noisy data. In the economic paper that discusses the use of embeddings in demand models (Magnolfi et al., 2024), t-STE is applied. However, based on recommendations from (Vankadara et al., 2023), I chose Soft Ordinal Embedding (SOE). The key difference between t-STE and SOE is that, given enough triplets, SOE is

more efficient at reconstructing local neighborhood distances. Since the distances between products in the expected efficacy space are a crucial parameter in my project, this advantage led to my choice.

Any triplet embedding method can be described by the following algorithm: it takes a function $\mathcal{O}$ as input and provides, as output, a multi-dimensional location $(x_i, y_i, z_i, \ldots, k_i)$ for each drug i.

The function $\mathcal{O}$, known as the Oracle function, transforms the actual data—responses to relative comparison questions—into an indicator function. Specifically, in my setup, each question involves three drugs: $(\langle i, j, k \rangle)$. For these three drugs, the answer can be represented as:

$$\mathcal{O}(\langle i, j, k \rangle) = \begin{cases} +1 & \text{if } \delta(i,j) < \delta(i,k) \\ 0 & \text{if } \delta(i,j) > \delta(i,k) \end{cases} \tag{11}$$

Here, the function $\mathcal{O}$ yields a positive value if drugs i and j are more similar in terms of effectiveness compared to drugs i and k, and zero otherwise. Since respondents are asked to identify the two drugs that are most similar in effectiveness, each question yields two outcomes from the Oracle function. This approach effectively reduces the total number of answers required by half.

The most natural objective function under this setup, for any subset of available triplets resulting in positive oracle $\mathcal{A}$, has the form:

$$\text{Err}_{\text{hard}}(X \mid \mathcal{A}) := \sum_{(i,j,k) \in \mathcal{A}} \mathbb{I}\left[\delta_{ij}(X) \geq \delta_{ik}(X)\right] \tag{12}$$

However, the drawback of this function is that it is binary and does not provide information on the degree of "wrongness" of the fitted distances.

(Terada and von Luxburg, 2014) provide an alternative optimization function based on the hinge triplet margin loss to satisfy the set of input triplet comparisons:

$$\text{Err}_{\text{soft}}(X \mid \mathcal{A}) := \sum_{(i,j,k) \in \mathcal{A}} \max\left\{0, 1 + \sqrt{\delta_{i,j}} - \sqrt{\delta_{i,k}}\right\} \tag{13}$$

Where Adam ((Kingma and Ba, 2014)), an algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments, is used for optimization.

IX.A.3. Comparability of experiment and control group embeddings

The novelty of my research, compared to both economics papers that use this technique (Magnolfi et al., 2024) and machine learning studies (Vankadara et al., 2023), lies in my ability to recover similarity measures not only in the actual scenario but also in counterfactual ones. I then use a structural model based on the data and baseline measures of expected efficacy to compute the new equilibrium. This approach requires that the coordinates of the drugs in both the baseline and across all counterfactuals share a consistent definition.

To ensure this consistency, I first pooled all responses from the relative comparison questions in both the baseline and counterfactual scenarios. I assume that each treatment changes the context of the products, meaning that, for instance, Advil in the baseline and Advil in any counterfactual are considered different drugs. As a result, there is no overlap of the same product across treatments.

The outcome of the triplet embedding results in four distinct clusters, each representing the drugs' similarity relationships under the baseline and the three different counterfactual scenarios. However, triplet embedding is inherently indifferent to the scale, orientation, and rotation of the points in the embedding space. This means that the spatial relationships within each cluster are preserved, but the clusters themselves can be shifted, rotated, or scaled relative to one another. For example, the cluster from a counterfactual scenario may have a different centroid (central point) or be oriented differently than the baseline cluster, even though the internal structure (i.e., the relative distances between points) remains the same.

To ensure that I can meaningfully compare the counterfactual clusters to the baseline cluster and integrate them into the structural demand model, I needed to apply transformations to each counterfactual cluster. These transformations ensure that the clusters are aligned and can be interpreted in the same coordinate system as the baseline. Specifically, I performed two key adjustments: (1) Aligning Centroids: First, I aligned the centroids (the central points) of each counterfactual cluster with the centroid of the baseline cluster. This ensures that all clusters are centered at the same point, making it easier to compare the spatial relationships of drugs across scenarios; (2) Adjusting Outliers: Second, I adjusted the position of the outliers within each counterfactual cluster to match those in the baseline cluster. By aligning these outliers, I ensured that the relative scale and spread of points in each counterfactual scenario closely mirror those in the baseline. This step corrects for differences in the size or dispersion of clusters, preserving the original relationships between products while ensuring that they are all represented on the same scale.

These transformations ensure that the similarity measures across both the baseline and counterfactual scenarios are consistent and can be accurately incorporated into the structural demand model. By doing so, I can compute new equilibria that account for the changing context of the drugs across different treatments.

IX.B. Supply side discussion

IX.B.1. Nash-Bertrand properties

The goal is to demonstrate that the current pricing of the same product across different package sizes cannot be explained by standard Nash-Bertrand competition. To test this, I examine the joint price response of products to demand and supply shocks, assessing whether the observed price adjustments align with the expected response of one product conditional on the response of another.

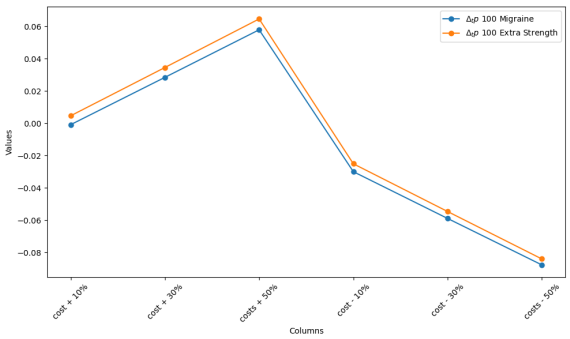
Claim 1. *Under a demand shock driven by label preferences, the prices of all products with the affected label adjust consistently across all package sizes.*

Claim 2. *Under a supply shock affecting biologically equivalent products, prices of these products within the same size and firm exhibit a highly correlated response.*

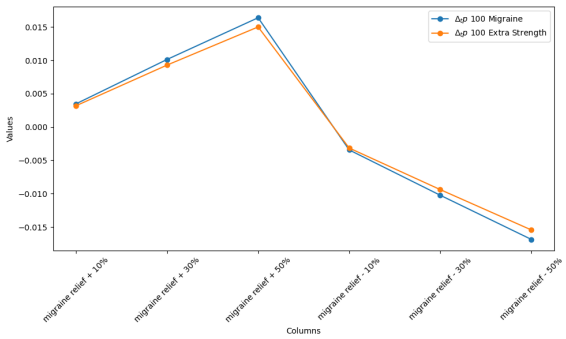
To motivate these propositions, I simulate both shocks within a simplified version of the market. I hold market composition and individual price sensitivity constant, allowing for a clean comparison of price dynamics under different competitive conditions. This approach helps isolate the role of label-driven demand effects and biological equivalence in pricing behavior, further highlighting deviations from standard Nash-Bertrand pricing predictions.

Assume that products in the market are characterized by 3 characteristics: brand: {brand; generic}, symptom label: {migraine, headache}, and size of the package: {24; 100}. Consumers prefer branded products, and they prefer migraine labels in case they enter the market with the belief that they need migraine relief and headaches in another case. In the simulation, I assume that brand preferences are 0.5, migraine preferences for those who need migraine relief are 1.5, and size preferences are -0.01.

For simplicity, all the products in the market have the same active components and are biologically equivalent. Two firms produce products with brand labels and the other with generic labels.



(a) Response to shock in costs



(b) Response to shock in demand

Figure 17: Responses to Shocks in Costs and Demand

IX.B.2. Test of Nash-Bertrand assumptions

Assuming there was no cost shock since the same products in different sizes are substitutes, the change in the price of one size due to the demand shock should affect the price of the product with the other size. To check this pattern in my data, I run the regression [15]:

$$\Delta_t p_{j,24} = \alpha_{t,j} + \Delta_t p_{j,100} + \gamma_m + \delta_b + \omega_a + \epsilon_{j,t}$$

Where:

$$\Delta_t p = \frac{p_t - p_{t-1}}{p_{t-1}}$$

The price is a non-promotional price for this month.

Results of this regression are presented in table 7. Based on these results, on average, there is no response to the price change of smaller sizes to the change in the bigger sizes. Now, assume that I observed these results because there was a cost shock that affected high prices but not small prices. Based on the Nash-Bertrand, products that face the same cost shock, biologically equivalent products of the same brand and size, should have nearly the same price response. Results of the same regression as for small and big sizes but for the biologically equivalent products of the same brand and size are presented in Table 8. They reject the fact that price movement for size 100 was solely due to the cost shock.

Additionally, across all markets for most products, the difference between the same products of sizes 24 and 100 lies between 40 and 60 percent (Figure 31). Based on the demand estimation for the small packages, demand is inelastic or has lower elasticity than big-size products (Figure 27). Therefore, if the whole panel of products would face a

[15] The two most popular products for any brand are the 24 or 100-pill package. **??**

Nash-Bertrand price setting, prices for the small packages should be higher than those for the big packages. This is an infeasible market equilibrium. Finally, I capture 61% of revenue by focusing on the 80+ size market (Figure 33).

Dependent Variable:	$\Delta_t p_k$			
Model:	(1) $k = 20$	(2) $k = 24$	(3) $k = 80$	(4) $k = 200$
Variables:				
$\Delta_t p_{100}$	0.1117	-0.0053	0.5328***	-0.0532**
	(0.0783)	(0.1202)	(0.1210)	(0.0200)
Fixed-effects:				
retailer code	Yes	Yes	Yes	Yes
DMA code	Yes	Yes	Yes	Yes
brand	Yes	Yes	Yes	Yes
active components + symptom label	Yes	Yes	Yes	Yes
Fit statistics:				
Observations	2,736	3,853	2,391	2,903
R^2	0.02102	0.01664	0.15374	0.01044
Within R^2	0.01586	2.51×10^{-6}	0.14238	0.00352

Clustered (retailer code) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

TABLE 7: Regression Results for $\Delta_t p_k$ with Different Values of k

IX.C. Triplet embedding additional results

Figures 18a and 18b present the results of the two-dimensional triplet embedding for headache and migraine relief, with zoomed versions shown in Figures 18c and 18d. The triplet embedding algorithm allows flexibility in selecting the number of dimensions in the relative efficacy space; here, I chose two dimensions. Using more than one dimension captures the multidimensionality of relative efficacy judgments, reflecting the complexity of these assessments. However, to incorporate this dimension into the demand model (Section IV) while avoiding the curse of dimensionality, I limit the number of dimensions to two, assuming that two primary groups of factors guide individuals in comparing drugs based on expected efficacy. Each dimension may reflect multiple underlying influences, such as past experience and product labeling. Interpreting these graphs involves two main aspects: (1) a qualitative measure, indicating the relative distances between products, and (2) a sequential measure, identifying the first, second, third, etc., substitutes for each product.

The results showing the distances for Excedrin Migraine and Extra Strength under the

Dependent Variable:	$\Delta_t p_{100,\ excedrin\ migraine}$
Model:	(1)
Variables	
Constant	-0.0005
	(0.0016)
$\Delta_t p_{100,\ excedrin\ extra\ strength}$	0.8071***
	(0.0235)
Fit statistics	
Observations	616
R^2	0.65792
Adjusted R^2	0.65736

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

TABLE 8: Regression Results for $\Delta_t p_k$ with k determine biological equivalence inside brand and size

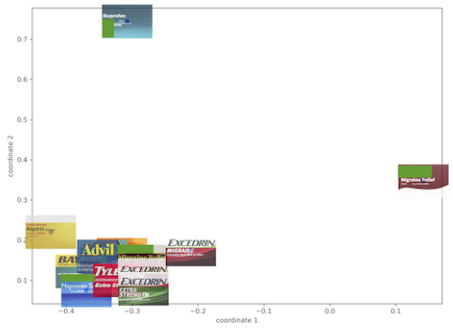
"Same Active Components" treatment are presented for migraine relief in figures: 19a and 19b, and for headache relief 19c and 19d.

The "Same Effectiveness" treatment results are presented for migraine relief in figures: 20a and 20b, and for headache relief 20c and 20d.

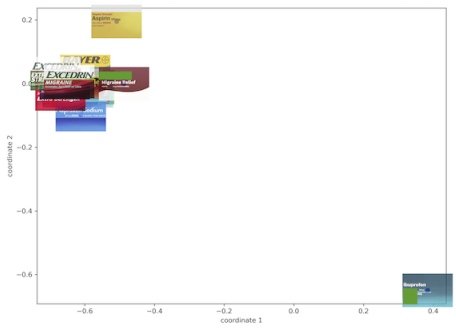
Finally, for "Migraine Educational Material" treatment results are presented for migraine relief in figures: 21a and 21b, and for headache relief 21c and 21d.

IX.D. Savings without price response

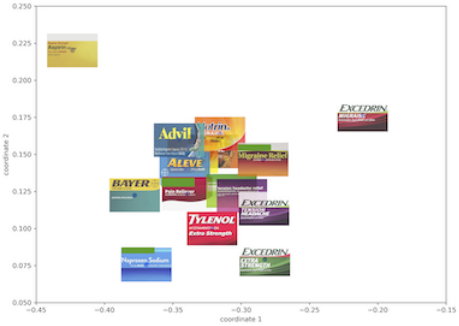
Changes in consumer choices lead to savings of two cents on average per individual in the treatments that involve front-label modifications (Figure 22), while no average savings are observed for the treatment that provides educational material to consumers. Negative savings do not necessarily translate to a welfare loss. For example, consider a consumer who, in the baseline scenario, prefers branded products and is unaware that Excedrin Migraine and Excedrin Extra Strength have the same efficacy. This consumer may have substituted Excedrin Migraine with its generic version, which is cheaper than any brand. After the treatment, if the consumer learns about the biological equivalence between these products, they might switch from Excedrin Migraine to Excedrin Extra Strength, which is more expensive than the generic version. The small average savings reflect the fact that each



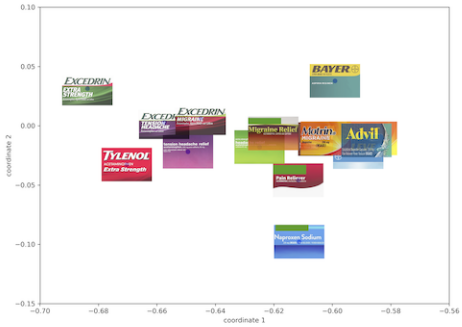
(a) Headache



(b) Migraine



(c) Headache, Zoomed



(d) Migraine, Zoomed

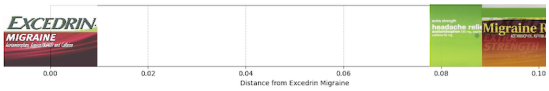
Figure 18: Relative and Zoomed Effectiveness Spaces for Drugs to Relieve Headache and Migraine

consumer's choice set includes products, not all of which have a cheaper biologically equivalent alternative. Furthermore, as demonstrated in the following section, transparency in information does not necessarily lead consumers to choose the cheaper product. Where the explanation is a significant role that preferences for specific product attributes play in decision-making.

IX.E. "Wrong" choices

IX.E.1. Heterogeneity, no price response

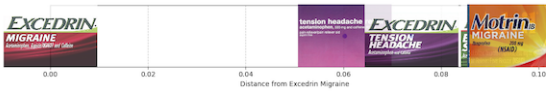
To further investigate the source of this heterogeneity, I decompose the individual and product-pair-level probability of making wrong choices (conditional on choosing a biologically identical pair of the same size) based on two parameters: (1) a binary variable indicating whether the symptom label of the more expensive product is "migraine," and (2)



(a) Excedrin Migraine; Relieve Migraine



(b) Excedrin Extra Strength; Relieve Migraine



(c) Excedrin Migraine; Relieve Headache



(d) Excedrin Extra Strength; Relieve Headache

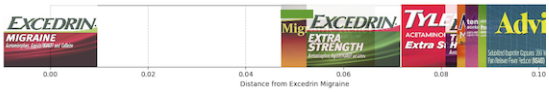
Figure 19: "Same Active Components" Treatment: Distances from Excedrin to Other Drugs for Migraine and Headache Relief



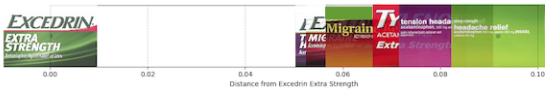
(a) Excedrin Migraine; Relieve Migraine



(b) Excedrin Extra Strength; Relieve Migraine



(c) Excedrin Migraine; Relieve Headache



(d) Excedrin Extra Strength; Relieve Headache

Figure 20: "Same Effectiveness" Treatment: Distances from Excedrin to Other Drugs for Migraine and Headache Relief

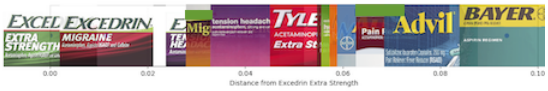
the consumer's willingness to pay for the migraine label. This decomposition allows me to explore the relationship between increased awareness of biological equivalence—revealed by changes in cross-price substitution—and consumer preferences for product attributes. Since I focus on the same brand, active components, and size, the only varying attribute is the symptom label. One potential source of heterogeneity is whether consumers enter the market seeking relief for migraine or headaches. However, since this value is highly correlated with the willingness to pay for the migraine label (Figure 9a), I do not include it in the regression.

The regression formula is as follows:

$$\begin{aligned}\mathbf{Pr}_{i,j}(&\textit{purchase product } j \mid \textit{biologically identical product } k \textit{ of same size is available}, p_k < p_j) = \alpha_{i,j} \\ &+ \beta_j \cdot \mathcal{I}(\textit{symptom label} = \textit{migraine}) + \gamma_i \cdot (\textit{wtp migraine label}) \\ &+ \gamma_{i,j} \cdot (\textit{wtp migraine label}) \cdot \mathcal{I}(\textit{symptom label} = \textit{migraine}) + \epsilon_{i,j}\end{aligned} \tag{14}$$



(a) Excedrin Migraine; Relieve Migraine



(b) 'Excedrin Extra Strength; Relieve Migraine



(c) Excedrin Migraine; Relieve Headache



(d) Excedrin Extra Strength; Relieve Headache

Figure 21: "Migraine Educational Material" Treatment: Distances from Excedrin Products to Other Drugs for Migraine and Headache Relief

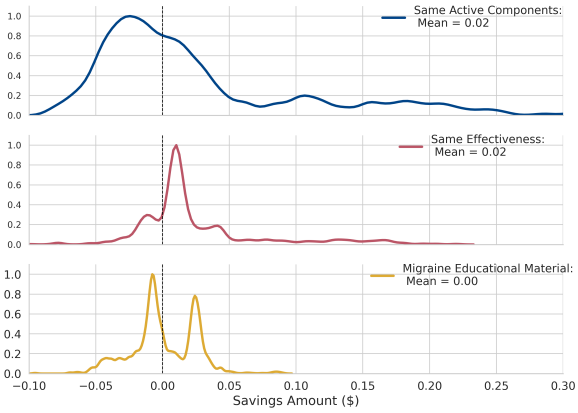


Figure 22: Result of expected efficacy change: Savings

The results of this regression are presented in Table 9. The main takeaway is that for treatments that increase substitution between biologically identical products (Figure 11a), there is an increase in wrong choices for products with symptom labels for which consumers have a high willingness to pay. For example, if a consumer has a negative willingness to pay for the migraine label, the relative change in the probability of making a wrong choice for a product with the headache symptom label would be $6.2 + (-2.2) \times$ (negative wtp for migraine label) > 0. Conversely, if the consumer is buying a product with the headache label but has a positive willingness to pay for the migraine label, the probability of making a wrong choice will be lower: $6.2 + (-2.2) \times$ (positive wtp for migraine label).

Similarly, for migraine-labeled products, if a consumer has a negative willingness to pay for the migraine label, the probability of making a wrong choice will decrease more for those with a lower willingness to pay, as shown in $6.2 - 13.6 - 2.1 \times$ (negative wtp) $<$

$6.2 - 13.6 + 2.2 \times$ (positive wtp).

Notably, in the case of headache products, only individuals with a positive willingness to pay for the migraine label experience a potential reduction in wrong choices. However, for migraine-labeled products, even those with a lower willingness to pay for the migraine label show a reduction in wrong choices. This aligns with reduced-form evidence suggesting that the migraine group is more responsive to the information treatments in changing the relative effectiveness space.

Furthermore, the change in the sign of the interaction between the migraine label and the willingness to pay for the migraine label, coupled with the fact that this treatment does not lead to an increase in cross-price elasticity between biologically equivalent products (Figure 11a), suggests that the reduction in wrong choices for the treatment providing consumers with brochures has a different underlying mechanism—one that is less reliant on increased awareness of biological equivalence and more related to preferences for product attributes.

	Same Active Components	Same Effectiveness	Migraine Educational Material
Intercept	6.2144***	3.3859***	7.4548***
	(0.0196)	(0.0054)	(0.0038)
Migraine	-13.5932***	-8.1730***	-19.7342***
	(0.0276)	(0.0075)	(0.0053)
WTP for Migraine Label	-2.1691***	-0.5558***	0.5249***
	(0.0027)	(0.0008)	(0.0005)
Migraine · WTP for Migraine Label	4.3102***	1.2483***	-0.2274***
	(0.0039)	(0.0011)	(0.0007)
R-squared	0.3456	0.3120	0.9329
N	3,240,000	3,240,000	3,240,000

TABLE 9: Decomposition of Wrong Choice Probability

In this analysis, I focus solely on biologically identical product pairs, as I do not aim to take a stance on the consumption of generic versus same-brand but differently labeled products.

IX.E.2. Heterogeneity in equilibrium

The response of firms producing biologically identical products suggests a hypothesis: treatments that convey new information through front-label modifications lead to second-degree price discrimination. This results in certain consumer types missing out on potential savings from price decreases driven by changes in the pricing scheme.

To test this hypothesis I first compute how the probability of making a "wrong" choice, conditional on picking a product inside a biologically equivalent pare of the products of

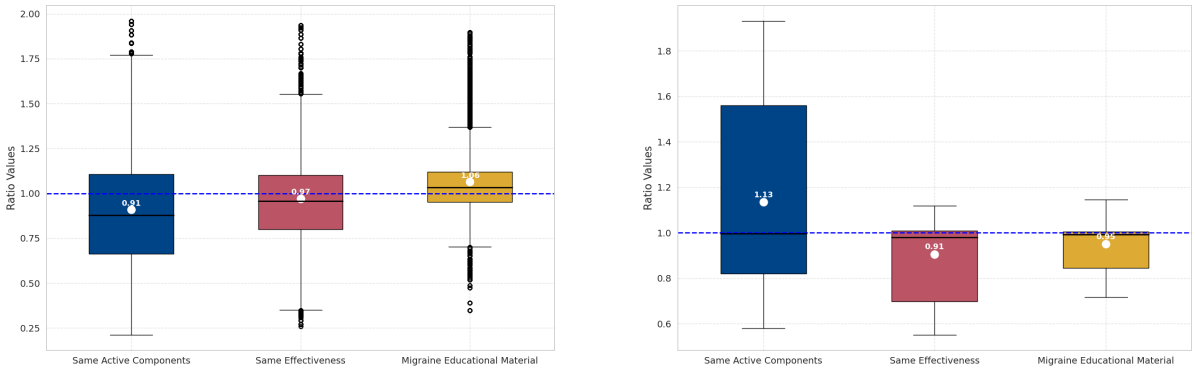
the same size what is the probability to choose a more expensive one, changes in each counterfactual compared to the baseline:

$$\frac{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(x)}, p_x\right)}{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(b)}, p_b\right)} \tag{15}$$

Where:

- x $\in$ { "Same Active Components", "Same Effectiveness" , "Migraine Educational Material"}
- b - baseline equilibrium

Figures 23a and 23b illustrate the ratio of "wrong" choices in each treatment compared to "wrong" choices in the baseline, measured at the individual level for biologically equivalent products within the same brand and for brand/generic pairs. First, inside same brand comparison in the "Same Effectiveness" treatment while on average consumers decrease the probability to make "wrong" choices this number seems quite small conditional on a 30% average increase in the substitution between biologically equivalent products inside same brand (11a), analogously, for "Same Active Components" Treatment there is an increase in the number of the "wrong" choices in the brand/generic comparison even though this treatment leads to the highest increase in the substitution from branded version to generic (11b).

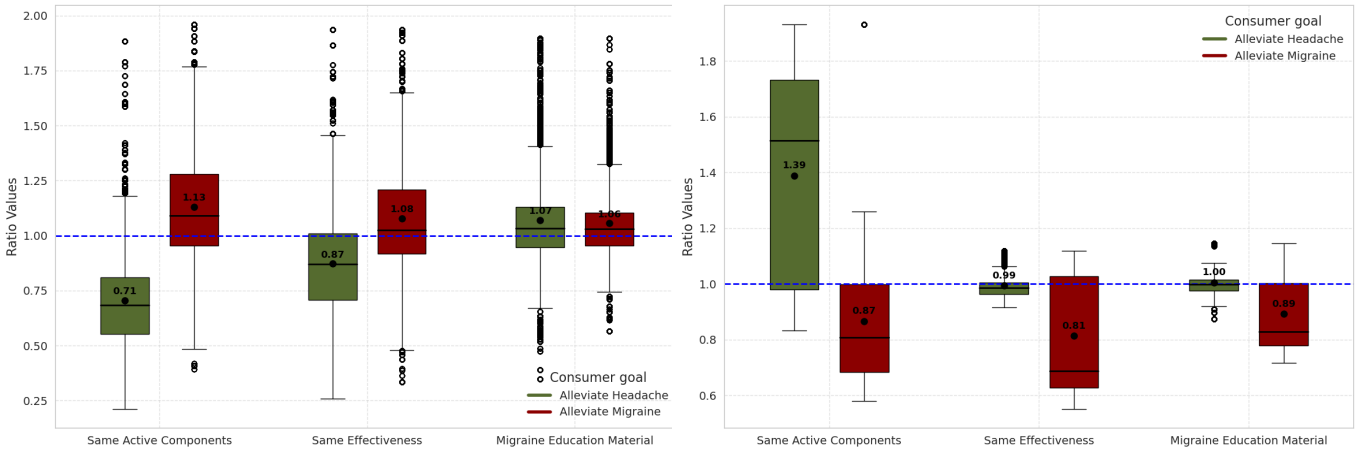


(a) Between Biologically Identical Products Inside Same Brand

(b) Between Products With Same Active Components Different Brand

Figure 23: Box Plot with Mean as a Point: Comparison of Changes in the Probability to Make "Wrong" Choices at the Individual Level

To evaluate these ambiguities, I first decompose "wrong" choices by distinguishing whether they occur among consumers who enter the market seeking migraine relief or headache relief. Figure 24a breaks down "wrong" choices within the biologically equivalent products of the same brand, while Figure 24b does so for the biologically equivalent group of brand and generic products. The results reveal that individuals aiming to alleviate migraines are the primary drivers of the high probability of "wrong" choices in the counterfactual scenarios for the inside same brand comparison. While individuals who search for the Headache relief drive wrong choices in the brand/generic comparison. Since in the general equilibrium model, this effect could stem from either price changes, strategic response to the altered information structure (see Figure 14), or changes in efficacy levels I formulate two propositions that test whether I observe a new version of second-degree price discrimination or just change in choices due to information transparency.



(a) Between Biologically Identical Products Inside Same Brand, Individual Level

(b) Between Products With Same Active Components Different Brand, Individual Level

Figure 24: Box Plot: Decomposition of Changes in Probability to Make "Wrong" Choices at the Individual Level

Claim 3. *In the "Same Active Components" treatment, people who search for headache relief switch from a biologically equivalent product inside the brand / generic pair to a biologically equivalent pair inside the same brand.*

The intuition behind this hypothesis is that the "Same Active Components" treatment educates consumers about the presence of a biologically equivalent product within the same brand. For consumers seeking relief from headaches, brand preferences tend to outweigh

label considerations (9b). As a result, this product becomes a more suitable substitute than the generic version because it is biologically equivalent and shares the same brand.

Claim 4. *In the "Same Effectiveness" treatment, people who search for migraine relief switch from biologically equivalent products within the same brand to the generic version.*

The intuition of this proposition is reversed to the previous one. For consumers who enter the market to alleviate migraine, label preferences play a crucial role in decision-making.

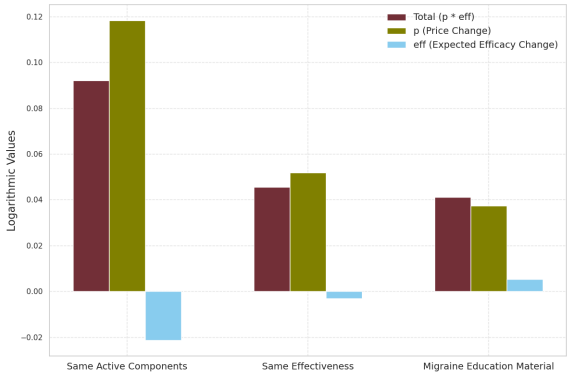
To test these hypotheses, I further decompose the change in the "wrong" choice into the change in the "wrong" choice due to the price change, keeping expected efficiency at the same level of the new treatment, and due to the change in the expected efficiency, keeping the price at the baseline level:

$$\frac{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(x)}, p_x\right)}{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(b)}, p_b\right)} \tag{16}$$

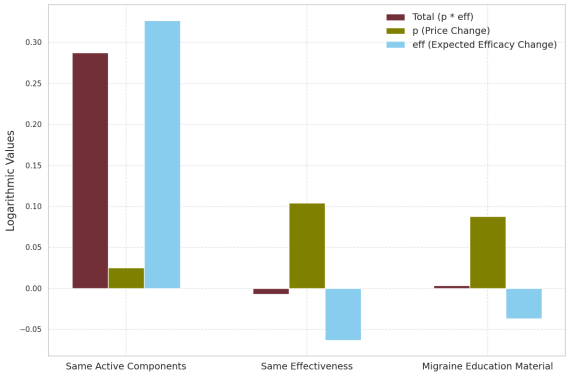
$$= \underbrace{\frac{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(x)}, p_b\right)}{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(b)}, p_b\right)}}_{\text{Impact of Efficacy}} \cdot \underbrace{\frac{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(x)}, p_x\right)}{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(x)}, p_b\right)}}_{\text{Impact of Price}} \tag{17}$$

Moreover, the logarithms of both the right- and left-hand sides are taken to facilitate easier interpretation. If change in efficacy is the main driving force, the positive total value of the change in the "wrong" choices will be primarily explained by a dominant positive value in the "Impact of Efficacy" component. However, if the "Impact of Price" component is dominant and positive, strategic price adjustment and second-degree price discrimination are based on either preference for branded products or strong label preferences among individuals seeking relief for migraines.

Results from the Figures 25a and 25b support Proposition 3 and reject Proposition 4. This suggests that the new pricing schemes in the "Same Active Components" and "Same Effectiveness" treatments lead to second-degree price discrimination against individuals with strong preferences for the migraine label.



(a) Same Brand; Migraine Relief



(b) Brand vs Generic; Headache Relief

Figure 25: Decomposition of Changes in the Probability of Making "Wrong" Choices by Product and Condition

IX.F. Additional Figures and Tables

Figure 26 illustrates changes in individual product-level choice patterns, conditional on market entry, by plotting the distribution of absolute changes in purchase probabilities. Front-label statements highlighting shared active components led to an average choice shift of 16.69%, while statements indicating equivalent effectiveness resulted in a 10.58% shift. The educational brochure had the smallest effect, with an average shift of 5.38%. These changes closely resemble substitution pattern responses, reinforcing the idea that information provision influences consumer choices in a way that aligns with cross-price elasticity changes. In Online Appendix IX.D, I present the impact of these choice changes on total spending.

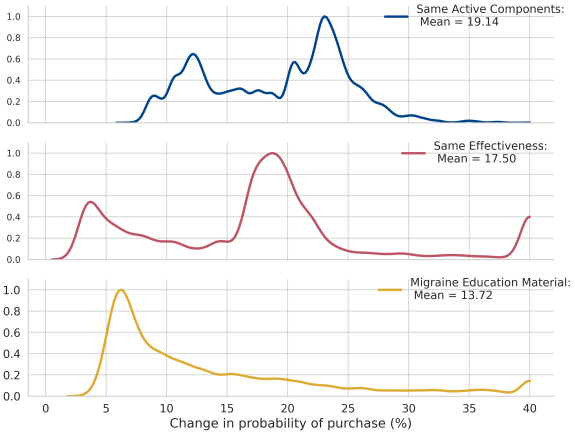


Figure 26: Effect of Expected Efficacy Change on Purchase Probability

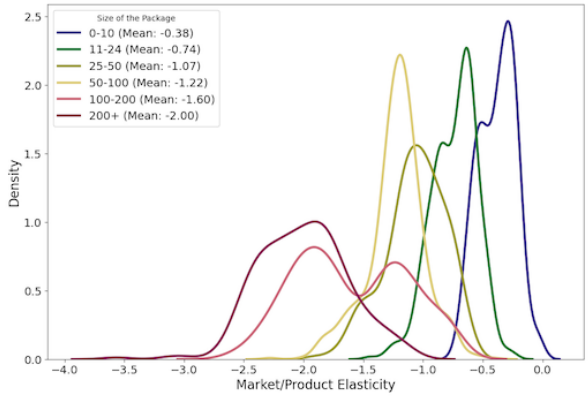


Figure 27: Distribution of Product Own Price Elasticity, Individual Level

Since I assume that firms set up prices based on a market segment I recompute my demand model for this segment: Table 11.

I generalize results of the section by running the following regression::

$$\mathbf{Pr}(\text{“wrong choice”})_i = \beta_0 + \beta_1 \cdot \text{chosen drugs have same label} \cdot \text{treatment type} \cdot \text{group type} + \sum_k \gamma_k \cdot X_i^k + \epsilon_i \tag{18}$$

.

While all treatments lead to a downward pressure on prices due to an increase in the cross-price substitution both inside biologically equivalent products and across products with different active components, the source of the downward pressure depends on the treatment. The same active components and same effectiveness treatment lead to downward prices due to an increase in the substitution between biologically equivalent products and substitution of these biologically equivalent products with other 28 and 29. The major reason for the treatment of migraine education material is an increase in the substitution between products with different active components 30.

To support the hypothesis that the misbeliefs about efficacy in this market may stem, at least in part, from a lack of information, I test whether individuals with higher acquisition costs—those without a college education—benefit more from information treatments that provide low-cost information, such as modifications to product labels. The results, shown in Figure 37, suggest a decrease in the fraction of incorrect choices among individuals without a college education in treatments that modify the label. Although the analysis is underpowered

TABLE 10

	Dependent variable:	
	mean	
	(1)	(2)
sample_labelSame Active Components	−0.034	−0.035
	(0.027)	(0.027)
sample_labelSame Effectiveness	−0.053**	−0.053**
	(0.027)	(0.027)
sample_labelMigraine Educational Material	−0.014	−0.015
	(0.027)	(0.027)
educ		−0.010
		(0.025)
age		−0.005
		(0.004)
age_sq		0.00004
		(0.00004)
Constant	0.244***	0.359***
	(0.019)	(0.085)
Observations	878	878
R^2	0.005	0.007
Adjusted R^2	0.002	0.001
Residual Std. Error	0.280 (df = 874)	0.280 (df = 871)
F Statistic	1.491 (df = 3; 874)	1.079 (df = 6; 871)
Note:		*p<0.1; **p<0.05; ***p<0.01

Variable Name	Coefficient
Household Income * Price	0.0019
Household Size * Price	-0.0121
Age * Price	0.0025
Household Income * size1 amount	0.0003
Household Income * migraine	-0.0256
Age * Migraine Label	-0.0303
Household Income * const	0.0947
Household Size * const	-0.2433
Age	0.0253
Female	0.7392
Female * Migraine Label	0.0280
Female * Number Pills	-0.0002
Female * Price	0.0109
Belief Need Migraine Relief * avg price	0.0136
Belief Need Migraine Relief * size1 amount	0.0008
Belief Need Migraine Relief * migraine	1.5210
Belief Need Migraine Relief * brand	1.3787
Belief Need Headache Relief * Dim1 Expected Efficacy	17.3461
Belief Need Headache Relief * Dim2 Expected Efficacy	12.0633
Belief Need Migraine Relief * Dim1 Expected Efficacy	21.1804
Belief Need Migraine Relief * Dim2 Expected Efficacy	0.4614
Belief Need Migraine Relief	-2.5905
rc on const	0.0000
rc on price	0.1987
rc on size cluster	0.0000
Const	-0.1463
Price	-0.4206
Migraine Label	2.2121
Number Pills	-0.0265
Brand	8.5571

TABLE 11: Demand Model used in the Supply Response

due to sample size limitations and heterogeneous responses, the evidence points to a larger reduction in incorrect choices for this group when exposed to label modifications.

Dependent Variable:	$\mathbf{Pr}$("wrong choice")
chosen drugs have same label	-0.0826*
	(0.0114)
chosen drugs have same label · tr: same active components	-0.0275
	(0.0112)
chosen drugs have same label · tr: same effectiveness	-0.1241*
	(0.0113)
chosen drugs have same label · tr: education material	0.0259
	(0.0185)
chosen drugs have same label · group: migraine relief	0.0930*
	(0.0126)
tr: same active components · group: migraine relief	-0.0111
	(0.0108)
tr: same effectiveness · group: migraine relief	-0.1041
	(0.0199)
tr: education material · group: migraine relief	0.0896
	(0.0173)
chosen drugs have same label · tr: same active components · group: migraine relief	-0.0501
	(0.0129)
chosen drugs have same label · tr: same effectiveness · group: migraine relief	0.0554
	(0.0105)
chosen drugs have same label · tr: education material · group: migraine relief	-0.2079*
	(0.0227)
Fixed-effects	
Treatment type	Yes
Group assignment	Yes
Region	Yes
Social-demographic characteristics	Yes
Frequency Migraine	Yes
Frequency Headache	Yes
Observations	1,927
R^2	0.06251
Within R^2	0.02187

Clustered on treatment level standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

TABLE 12: Decomposition of the Probability of Making Wrong Choices

Treatment 1	Treatment 2	t-statistic
Same Active Components	Same Effectiveness	-4.0952
Same Active Components	Migraine Education Material	30.7741
Same Effectiveness	Migraine Education Material	56.0180

(a) Same brand different labels

Treatment 1	Treatment 2	t-statistic
Same Active Components	Same Effectiveness	4.7522
Same Active Components	Migraine Education Material	9.3487
Same Effectiveness	Migraine Education Material	7.3213

(b) Brand to Generic

TABLE 13: T-test comparison of treatment effects

Brand	Total Quantity	Market Share
Advil	11,125,423	40.91%
Aleve	5,390,743	19.84%
Tylenol	4,160,926	15.32%
Excedrin	3,535,312	13.01%
Motrin IB	1,473,888	5.43%
Bayer	964,318	3.55%
Generic	517,131	1.90%

TABLE 14: Brand Descriptions and Shares



Figure 28: Ratio of the cross-price elasticities: “Same Active Components” to baseline

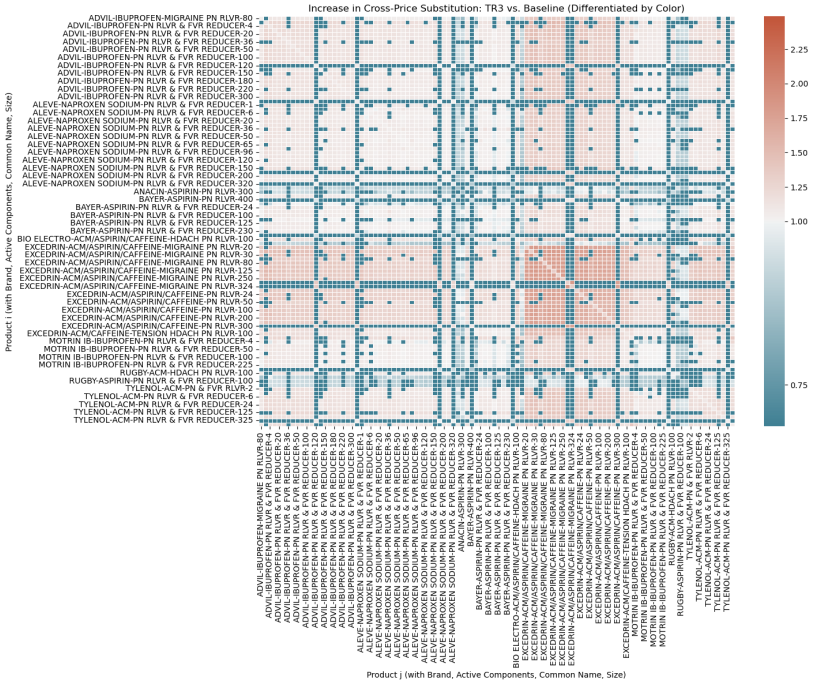


Figure 29: Ratio of the cross-price elasticities: Same Effectiveness to baseline



Figure 30: Ratio of the cross-price elasticities: “Migraine Educational Material” to baseline

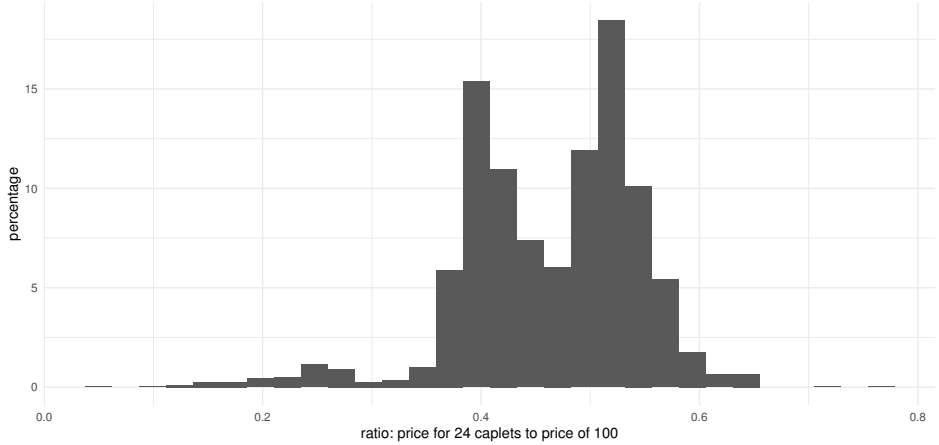


Figure 31: Distribution on the product market level: price for 24 pills to price for 100 pills

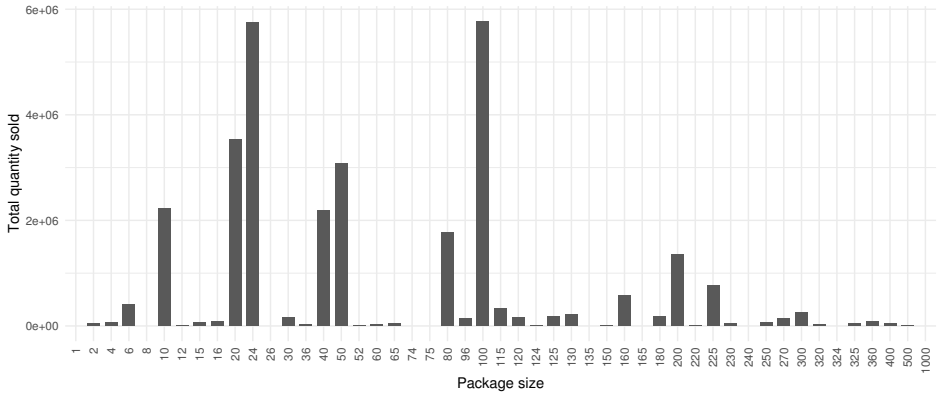


Figure 32: Distribution of total quantities sold by size of the package

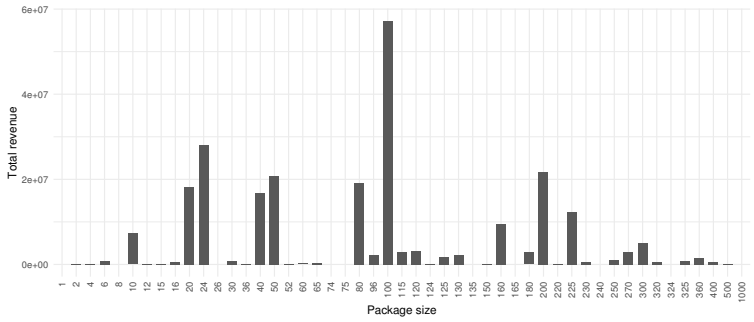


Figure 33: Distribution of total revenue sold by size of the package

Retailer	Products	Active Components	Markets	Variation in products
1	98	6	19	0.04
2	70	6	1	0.03
3	64	6	4	0.04
4	64	6	6	0.03
5	62	6	6	0.02
6	60	6	1	0.02
7	60	6	2	0.02
8	59	6	1	0.03
9	57	6	1	0.02
10	57	6	1	0.04
11	53	6	2	0.06
12	53	6	1	0.01
13	53	6	1	0.01
14	49	6	1	0.02
15	28	6	1	0.05
16	28	5	1	0.03
17	10	5	1	0.00

TABLE 15: Retailer Information

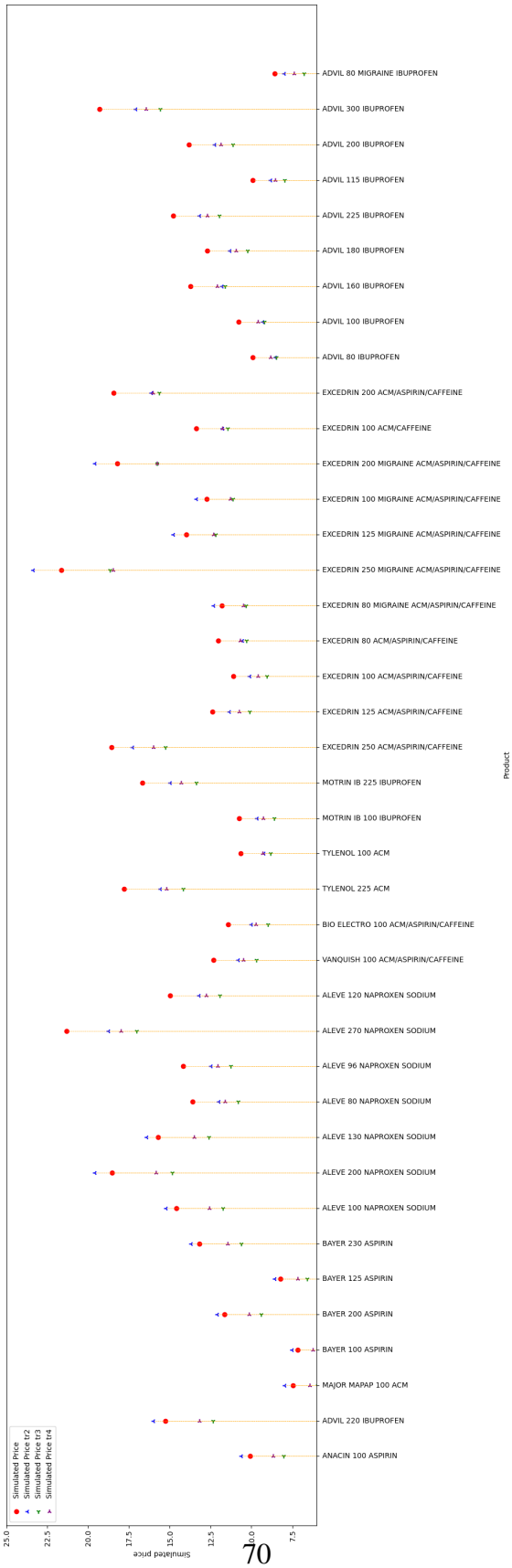


Figure 34: Example of price change, randomly selected market

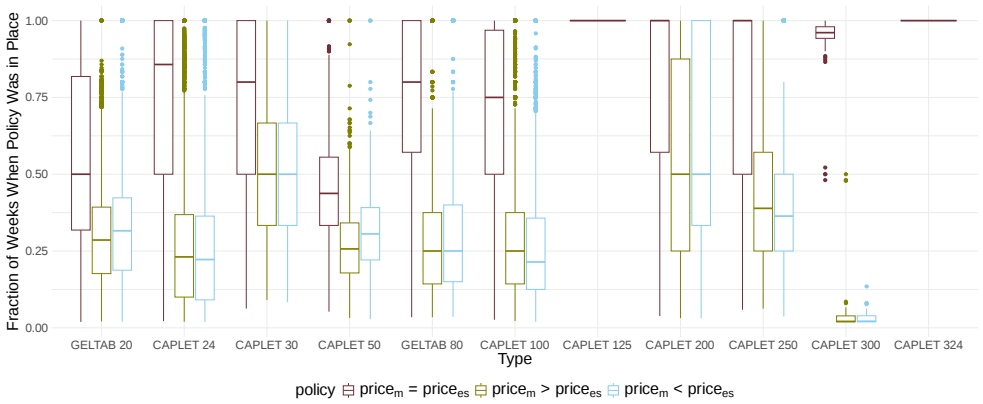
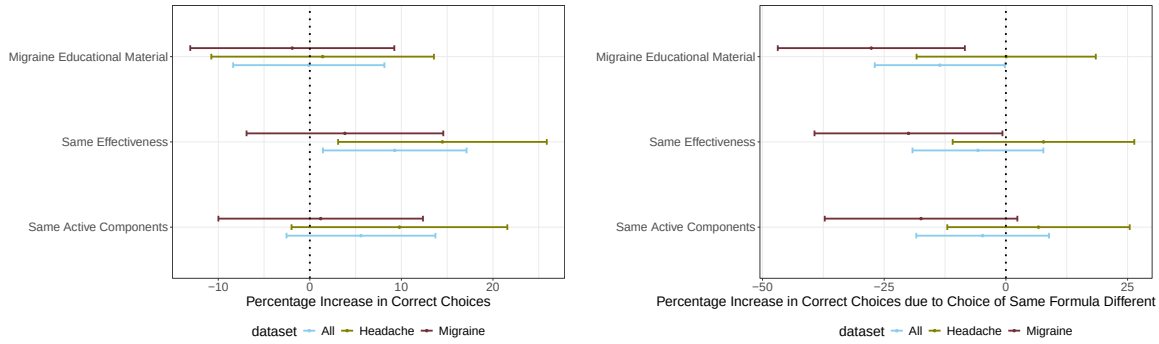


Figure 35: Pricing policy: Excedrin Migraine and Excedrin Extra Strength



(a) Choice of Generic Product

(b) Choice of Different Label Product

Figure 36: Comparison of Changes in Correct Choices Driven by Generic and Different Label Product Choices Across Treatments

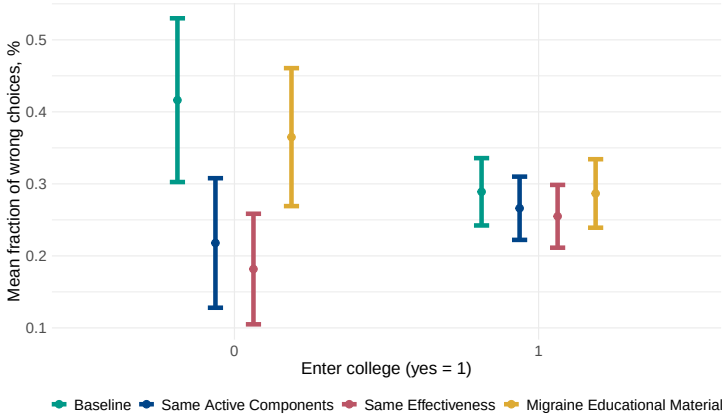


Figure 37: Impact of Label Modifications on Wrong Choices by Education Level

	Coefficient
Intercept	-0.5594
	(0.040)
Acetaminophen+Aspirin+Caffeine	-0.1338
	(0.025)
Acetaminophen+Caffeine	-0.0413
	(0.038)
Aspirin	0.2637
	(0.026)
Ibuprofen	-0.1143
	(0.022)
Naproxen Sodium	-0.3252
	(0.023)
Migraine Label	-0.1067
	(0.022)
Brand	0.0583
	(0.035)
Number of Pills	-0.0060
	(7.35e-05)
R-squared: 0.771, Observations: 2155	
Mean of Dependent Variable: -1.2183	

TABLE 16: Correlation Between Own-Price Elasticity and Product Characteristics

	Coefficient
Intercept	-0.5156
	(0.949)
Acetaminophen · Number of Pills	0.0291
	(0.001)
Acetaminophen+Aspirin+Caffeine · Number of Pills	0.0247
	(0.001)
Acetaminophen+Caffeine · Number of Pills	0.0238
	(0.002)
Aspirin · Number of Pills	0.0239
	(0.001)
Ibuprofen · Number of Pills	0.0216
	(0.001)
Naproxen Sodium · Number of Pills	0.0163
	(0.001)
Brand FE	Yes
Market FE	Yes

TABLE 17: Regression Results: Costs Fitting