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

Anastasiia Evdokimova

Yale University
anastasiia.evdokimova@yale.edu

Latest version available here

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, Unobserved Quality.

JEL CODES. D12, I11, L65.

CONTACT. anastasiia.evdokimova@yale.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 CSIO. 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 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 (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 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 significant consumer impact. Many of these drugs share the same active components, making them 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 similarity in the perceived efficacy dimension based on symptom/brand-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 latent perceived efficacy space identified in the experiment into the model. Extend-

ing Magnolfi et al. (2024), who use crowdsourced product similarity measures to capture products' substitution, 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, in which equivalent efficacy is explicitly reported, increases substitution between products with identical formulations by 30% and between branded and generic products by 19%. The stronger effect within brands suggests consumers were previously less aware 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; 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). 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 disclosure, as examined in prior research (Dranove and Jin, 2010; Vatter, 2021).

Methodologically, this paper contributes to the integration of RCTs with structural models to enhance policy analysis, echoing the call by Todd and Wolpin (2023). It provides a concrete implementation of this integration, motivated by the increasing use of machine learning to reveal unobserved aspects of decision-making behavior (Magnolfi et al., 2024). While their approach recovers aggregate substitution patterns that blend observable and unobservable influences, I show how to directly estimate a specific unobservable parameter, perceived product quality. This enables counterfactual analysis of information provision policies.

The paper is structured as follows: Section II describes the institutional background, data, and experimental design. Section II.C presents reduced-form evidence. Section IV introduces the demand model, with results in Section V. Section VI outlines the supply side. Section VII incorporates experimental findings into the equilibrium models to analyze impacts on decisions, savings, and welfare.

II. Setting and data

II.A. OTC drug market

This paper examines the OTC drug market, where consumers often 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.

II.B. Data

My analysis focuses on pain medications that alleviate migraine and headache symptoms. I focus on this market since headache disorders being among the most prevalent neurological disorders, 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 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 major health-related shocks.[1] Products are defined at the UPC level, which uniquely pins down the brand, active ingredient combination, symptom label, dosage form, and package size. Symptom labels are inferred from the UPC description codes.

The final sample contains 498 products spanning seven major brands—Advil, Tylenol, Bayer Aspirin, Aleve, Excedrin, and Motrin—along with comparable non-brand/private-label equivalents. Table 1 summarizes revenues, unit-weighted mean prices, and unit shares by active ingredient combination, branded status (brand=1), and migraine labeling. Sales are concentrated in standard analgesics: ibuprofen accounts for roughly 47% of total revenue (and 46.8% of units), while acetaminophen-only and naproxen sodium products each contribute about 19% of revenue. Combination products containing acetaminophen/aspirin/caffeine represent about 10% of revenue and are the primary category in which migraine-labeled variants appear. Overall, migraine-labeled products comprise a small portion of the market (about 3.5% of units and 3.7% of revenue). Across categories, branded products command higher unit-weighted prices than non-branded/private-label alternatives (e.g., ibuprofen \$8.06 vs. \$5.37; naproxen sodium \$8.23 vs. \$6.01), while the unit-weighted average price across all observations is \$6.62.

A market is defined by month, retailer, and Nielsen Designated Marketing Area (DMA region(s)), 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 weighted average within DMA region(s). Aggregating to the DMA region(s) level is common (DellaVigna and Gentzkow, 2019). 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 0.06.

To narrow the focus, the analysis is restricted to the 600 largest markets. Twenty-two retailers serve these markets. On average, retailers offer approximately 112 products, with a median of 115. Most retailers list 6 active components, indicating a standardized product composition across firms.

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

[1] I exclude 2018 due to potential disruptions from Affordable Care Act changes and later years due to (post)pandemic trends.

Active ingredient(s)	Brand	Migraine	Revenue	Mean price	Unit share
Acetaminophen					
Acetaminophen	No	No	80,667,365.55	5.08	0.1197
Acetaminophen	Yes	No	84,462,384.85	7.33	0.0868
Acetaminophen combinations					
Acetaminophen/Aspirin/Caffeine	No	Yes	7,698,290.19	5.46	0.0106
Acetaminophen/Aspirin/Caffeine	No	No	13,550,366.32	4.86	0.0210
Acetaminophen/Aspirin/Caffeine	Yes	Yes	22,184,750.22	7.73	0.0216
Acetaminophen/Aspirin/Caffeine	Yes	No	44,967,056.57	7.46	0.0455
Acetaminophen/Caffeine	No	No	2,082,231.95	5.77	0.0027
Acetaminophen/Caffeine	Yes	No	5,147,132.50	7.47	0.0052
Aspirin and combinations					
Aspirin	No	No	15,386,690.62	3.96	0.0293
Aspirin	Yes	No	21,763,053.72	7.56	0.0217
Aspirin/Caffeine	No	No	1,075,942.86	4.60	0.0018
Aspirin/Caffeine	Yes	No	1,417,822.69	11.35	0.0009
Ibuprofen					
Ibuprofen	No	No	172,629,341.93	5.37	0.2422
Ibuprofen	Yes	Yes	2,919,293.30	6.94	0.0032
Ibuprofen	Yes	No	238,060,335.13	8.06	0.2226
Naproxen sodium					
Naproxen sodium	No	No	42,054,654.43	6.01	0.0527
Naproxen sodium	Yes	No	122,744,477.46	8.23	0.1124

Notes: "Brand" indicates branded products (Yes) versus non-branded/private-label equivalents (No). "Migraine" indicates whether the UPC carries a migraine symptom label. Revenue is in dollars; mean price is unit-weighted; unit share is computed over all units in the sample.

TABLE 1: Revenue, Average Price, and Unit Share by Active Ingredients, Brand Status, and Migraine Label (Units Omitted)

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 data from the FDA Label Repository, pharmaceutical companies frequently register biologically identical products under the same brand name as distinct offerings by assigning them different symptom labels. To examine whether this phenomenon constitutes a market failure, I first evaluate pricing strategies.

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.

Under second-degree price discrimination based on the label, either regime (2) or (3) should dominate: a consistent premium for one symptom label would imply that one product is systematically priced above the other. As shown in Figure 1, for 100-caplet packages, the

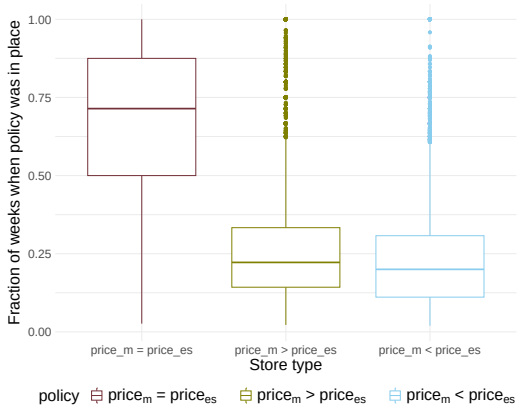


Figure 1: Pricing policy for 100 Caplets: Excedrin Migraine and Excedrin Extra Strength

median store maintains identical pricing for 67.8% 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 more expensive. Pricing variation is even more pronounced for smaller packages (see Online Appendix, Figure 37). These findings suggest no systematic second-degree price discrimination based on observable front-label attributes.

This conclusion is reinforced by the distribution of the week–store–UPC level relative price wedge, $\Delta \equiv (p_m - p_{es})/p_{es}$. Figure 2 shows that the distribution of Δ is sharply concentrated around zero, indicating that most observations correspond to parity pricing or only small deviations from parity. At the same time, the distribution exhibits long two-sided tails: large positive and negative price wedges occur occasionally, consistent with short-run retail pricing dynamics such as temporary promotions and price resets rather than a stable, one-sided pricing rule tied to the symptom label. Importantly, substantial mass on both sides of zero implies that, when price differences arise, they do not exhibit a systematic direction across stores and weeks, further undermining an interpretation of persistent label-based second-degree price discrimination.

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

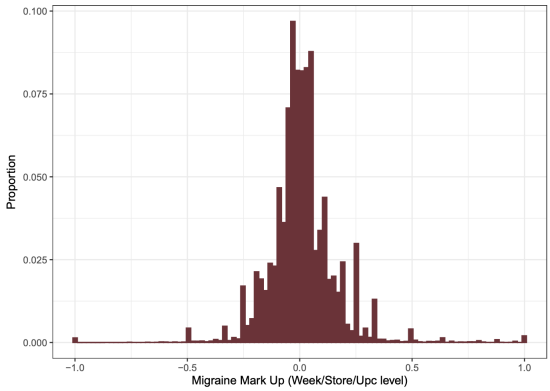


Figure 2: Pricing policy: Migraine Label Mark Up

and then plotted a histogram for all stores/weeks in 2017. As shown in Figure 3, 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.

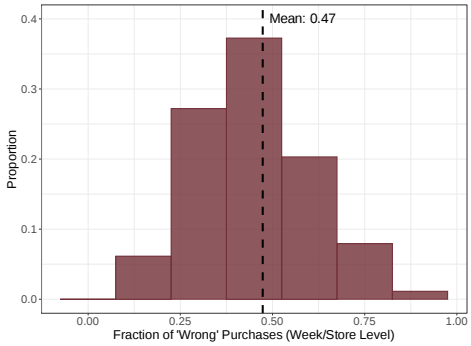


Figure 3: Histogram of Wrong Choices, Week/Store Level

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 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. They are shown the following instruction: *"In this section, imagine you have a headache and are given a choice from one of three painkillers, all sold at the same price. The choice of painkillers will change*

from one question to the next. Your task in each question is to identify the two drugs that are most similar to each other in terms of their effectiveness for headache relief. Effectiveness in this context refers to how well the drug reduces your headache and provides relief."[2] Respondents were then presented with a rotating set of triplets drawn from the full product universe, enabling inference on perceived similarities across all drug pairs:



Figure 4: Example of the Triplet

Each respondent answered 15 questions of this type.[3] 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 I.B. The only tuning parameter in the triplet embedding algorithm is the number of dimensions in the similarity space. I use a two-dimensional space to balance interpretability and complexity. This captures the multidimensional nature of efficacy judgments while avoiding the curse of dimensionality in the demand model (Section IV). Each dimension likely reflects multiple factors, such as past experience or labeling. The resulting graphs offer insights into the relative similarity of the drugs in the latent perceived efficacy space.

One way to illustrate the spaces generated by the relative efficacy triplet embedding algorithm is 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 similarity. 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.

[2] 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.

[3] 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.



(a) Relieve Migraine (b) Relieve Headache

Figure 5: Distance from Excedrin Products to Other Drugs for Migraine and Headache Relief

The network graphs in Figure 5 highlight key patterns that both reinforce the earlier findings on mistaken choices and offer additional insight into how substitution patterns shift across informational treatments. Notably, respondents consistently struggle to identify biologically equivalent products—those with identical active ingredients and efficacy—as their nearest substitutes. The results also underscore the powerful contextual role of symptom labels in shaping consumer decisions. Individuals tend to cluster products based on similar symptom labels before considering distinctions such as brand versus generic or equivalence within a brand. The positioning of drugs in the perceived efficacy space suggests that current OTC labeling practices generate biased beliefs about drug effectiveness. Crucially, this bias appears to differ between consumers seeking relief for "migraine" versus those seeking relief for "headache." In the Online Appendix I.C, I provide an alternative visualization of the triplet embeddings by calculating the relative distance from each product to all others; these results are consistent with the patterns observed in the network analysis.

Notably, triplet embedding offers key advantages over second-choice data and conjoint analysis. Unlike the second-choice approach, which only identifies relative similarity patterns (e.g., "A is most closest to B"), triplet embedding also provides quantitative similarity,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). Online Appendix E, 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.

The first treatment involves modifying the current strategy used by generic pharmaceutical companies to encourage consumers to choose generics over branded versions. Typically, generics emphasize that they contain the same active ingredients as branded products (Figure 6a). 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 6b). 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 7a, this approach helps bridge potential gaps in consumer knowledge by highlighting biochemical similarity, even if differences in branding or labeling initially obscure it. While one could argue that, in a real-world implementation of such labeling policies, firms might engage marketing consultants to design labels that are less transparent or more strategically ambiguous, the purpose of this experiment is precisely to identify the degree of transparency at which consumer choices begin to align with actual medical efficacy. One way to interpret the results, therefore, is as an upper bound on the potential shift in consumer response.

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 in response to 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 7b). 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.



(a) Generic Actual Market



(b) "Same Active Components" Treatment

Figure 6: 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 7: Second and Third Treatments

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.[4] For each information treatment design, I compute new values of the relative expected efficacy.[5]

The validity of the experiment as a representation of real-world choices, rather than a purely "lab" setting, raises several standard concerns. First, one might argue that because participants are not actually purchasing the drug, they may not fully reveal their true valuation of the product. While this is a legitimate concern, the fact that participants' relative judgment patterns are consistent with intuitive expectations and align with observable trends in actual purchase data helps to mitigate this issue. Second, in the experimental

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

[5]I discuss the computation of this value in more detail in the Online Appendix IV.C.1. New expected efficacy presented in Online Appendix E

setting, individuals are presented with the full set of drug options, whereas in real-world environments, consumer attention may be limited. This is a valid limitation. However, the results can be interpreted as a lower bound on potential bias. In practice, limited attention may lead to even greater distortions in how consumers perceive relative efficacy. While this particular threat is specific to the structure of the experiment, it reflects a broader challenge inherent in most demand estimation frameworks.

Despite these concerns, the experimental approach remains the most suitable method for addressing my research question. Disentangling preferences from beliefs using only choice data is extremely difficult and likely infeasible without imposing strong assumptions about belief formation. The experimental design provides a more direct means of identifying belief-driven variation in choices without relying on restrictive or unverifiable functional form assumptions.

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 rather than general headache is computed based on the survey responses.[6] To account for potential socio-demographic differences in migraine prevalence, I follow the medical literature and categorize respondents into four groups based on median income and age (above and below 65). The results of this classification are presented in Table 2. 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.

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

TABLE 2: Probability to Enter the Market with the Belief That You Need Migraine Relief

[6]Detailed computation is discussed in the Online Appendix I.A.

III.B. Descriptive evidence

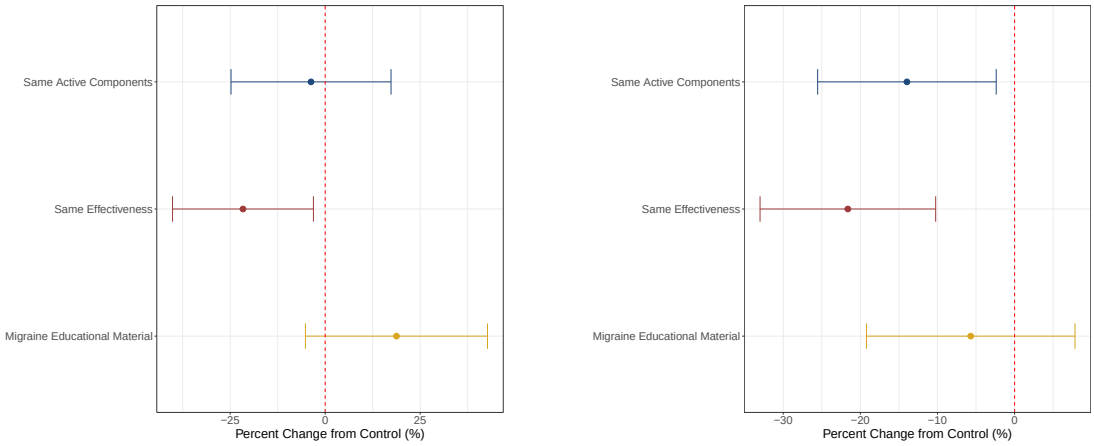
III.B.1. Wrong choices in perceived efficacy dimension

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 8 presents the average percent change in the fraction of "wrong" 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 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 de-



(a) Brand/Generic

(b) Different label biologically equivalent

Figure 8: Change in the Fraction of "Wrong" Choices Across Different Treatments Compared to Control Group

mand 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 more similar in perceived efficacy.[7] 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 (39) indicates that individuals with higher ex-ante information acquisition costs were the most responsive to the information treatments.

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

[7]This hypothesis is explored further in the section on changes in similarity within the latent perceived efficacy space (III.B.3).

following regression model:

$$\begin{aligned}\mathbf{Pr}(\text{wrong choice"})_i = &\beta_0 + \beta_1 \cdot \text{female}_i + \beta_2 \cdot \text{age}_i + \beta_3 \cdot \text{age}_i^2 + \beta_4 \cdot \log(\text{income})_i + \\ &\beta_5 \cdot \text{high school education}_i + \sum_k \gamma_k \cdot X_i^k + \epsilon_i \end{aligned} \tag{1}$$

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	-0.018	-0.024	-0.021
	(0.008)	(0.008)	(0.008)
Age^2	0.0002	0.0002	0.0002
	(0.0000)	(0.0000)	(0.0000)
High School or Higher	-0.097	-0.131	-0.072
	(0.056)	(0.065)	(0.053)
Observations	224	210	224

Mean: 42% wrong choices.
Specifications include group, region, industry, and purchase frequency fixed effects.
Standard errors are heteroskedasticity-robust.

TABLE 3: Impact of Information Acquisition Costs on Likelihood of Wrong Choices

Table 3 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 similarity in efficacy dimension

While the previous analysis demonstrated that information treatments reduce the likelihood of "wrong" choices in triplet comparisons, the triplet embedding approach allows for a

broader structural interpretation. Specifically, it reveals how substitution patterns across drugs evolve under different information treatments, providing insight into consumers' underlying perceptions of medical similarity. If consumers correctly internalize pharmacological equivalence, then products with the same active ingredients, such as branded and generic versions, or variants with different labels under the same brand, should appear closer together in the perceived efficacy space. Conversely, products with different active components should remain more distinct.

To assess these similarity patterns, I use the previously constructed networks based on the triplet-based efficacy embeddings. Changes in network structure across different information structures (Figure 9) provide a visual and analytic lens for understanding how beliefs about similarity in the latent perceived efficacy space shift in response to new information.

Information treatments shift these patterns in distinct ways. Under the "Same Active Ingredients" treatment, products sharing the same chemical composition—such as branded and generic pairs—become more closely connected, reflecting stronger perceived medical similarity. For instance, Excedrin Extra Strength no longer clusters with Tylenol but instead links to its generic equivalent, both labeled for migraine and headache relief. The "Same Effectiveness" treatment, while less effective in establishing brand-generic connections, significantly increases perceived efficacy similarity among biologically identical products that differ only in labeling. This suggests that consumers become more attuned to equivalence within brands when presented with simplified efficacy comparisons. By contrast, the "Migraine Education" treatment produces more diffuse effects. While it modestly improves some medically consistent linkages, its primary impact is a general increase in overall similarity, effectively flattening the perceived differentiation among products. This aligns with earlier findings indicating that the educational material had limited success in correcting specific misconceptions.

IV. Demand on the OTC drug market

Consumers are typically assumed to base drug choices 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, used for symptom relief, perceived quality reflects the consumer's belief about a drug's ability to alleviate a symptom, an inherently unobservable measure.

The previous sections motivated that misconceptions about efficacy might influence



(a) Control Group

(b) Same Active Components

(c) Same Efficacy

(d) Migraine Education Material

Figure 9: Triplet Embedding Network for Headache Relief

choices. Building on this, I use survey data to isolate the role of expected efficacy in decision-making. Under specific assumptions, standard mixed-logit methods can be applied using iterated expectations. I then outline the estimation strategy and instruments used to address potential endogeneity.

Identifying beliefs about efficacy is essential for modeling how consumers respond to information treatments. Without this step, efficacy cannot be distinguished from product-specific fixed effects or preferences for observable attributes like labeling and active ingredients, features consumers often use as proxies for effectiveness.[8]

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 (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 with a set of socio-demographic characteristics ($\mathbf{x}$) 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):

$$\mathcal{C}_j : \big(l_j, ac_j, g_j, p_{j,t}, w_j, q_{i,j,s}, \boldsymbol{v}_i\big),$$

where: (1) $p_{j,t}$ is price; (2) $q_{i,j,s} = \mathbf{E}\big[\textit{Efficacy}_{i,j} \mid s_i\big]$ is the symptom (s_i)–specific expected efficacy; (3) $\boldsymbol{v}_i$ are factors unobservable to the researcher (previous experience, recommendations, etc.); (4) $\mathbf{z}$: (1) l_j is the symptom label; (2) ac_j are the active components; (3) g_j indicates whether the drug is branded or generic; (4) w_j is package size;

Incorporating these parameters in the standard mixed-logit demand model (Berry et al., 1995), in a market (t) the utility of consumer i who is in symptom state $s \in \mathcal{S}$ for product j

[8]Consumers use the front label to infer which products are more effective for symptom relief.

has the following form:

$$u_{i,s,j,t} = \underbrace{\hat{\delta}_{j,t}}_{\text{mean utility}} - \alpha_{i,s}\, p_{j,t} + \sum_{k,r} p_{j,k}\, z_{i,r}\, \gamma^{\text{obs}}_{k,r} + \beta^{\text{eff}}_{s}\, \mathbf{E}\big[\textit{Efficacy}_{i,j} \mid s\big] + \sum_{k,r} x_{j,k}\, z_{i,r}\, \beta^{\text{obs}}_{k,r} + \sum_{k} x_{j,k}\, v_{i,k}\, \beta^{\text{unobs}}_{k} + \varepsilon_{i,s,j,t}, \tag{2}$$

where $\varepsilon_{i,s,j,t}$ is an i.i.d. Type-I Extreme Value shock (logit). The mean utility is

$$\hat{\delta}_{j,t} = \sum_{k} x_{j,k}\, \bar{\beta}_{k} + \tilde{\delta}_{j,\text{DMA}} + \delta_{\text{DMA},t} + \xi_{j,t}, \tag{3}$$

with product and market (DMA×month×retailer) fixed effects and an unobserved product-market component $\xi_{j,t}$. Random coefficients (e.g., $\alpha_{i,s}$) allow flexible substitution patterns.

Given (2), the *type-specific* logit choice probabilities are

$$s^{(s)}_{i,j,t} = \frac{\exp\big(u_{i,s,j,t}\big)}{1 + \sum_{k \in \mathcal{J}_t} \exp\big(u_{i,s,k,t}\big)} \tag{4}$$

IV.B. Aggregation

Assume that for each individual we *observe* a symptom type $b_i \in \{m, h\}$ (migraine, headache) that follows a Bernoulli distribution recovered from the survey:

$$b_i \sim \text{Bernoulli}(\rho_i), \qquad \rho_i \equiv \Pr(s_i = m).$$

With type-specific probabilities $s^{(m)}_{i,j,t}$ and $s^{(h)}_{i,j,t}$ computed from (2), aggregate market shares are the mixture of type-specific probabilities as in Berry and Jia (2010):

$$s^{\text{model}}_{j,t} = \Gamma_{m,t}\, \bar{s}^{(m)}_{j,t} + \Gamma_{h,t}\, \bar{s}^{(h)}_{j,t}, \tag{5}$$

where $\Gamma_{r,t} \equiv \mathbb{E}_i[\mathbf{1}\{b_i = r\}]$ is the market share of type r and $\bar{s}^{(r)}_{j,t} \equiv \mathbb{E}_i[s^{(r)}_{i,j,t} \mid b_i = r]$. This aggregates probabilities, not utilities, and nests the Berry–Jia discrete-types formulation.

I denote the set of consumers that choose a product j in market t as

$$\Theta_{j,t} = \{i \in \mathcal{I}_t : \mathbb{E}(u_{i,j,t}) \geq \mathbb{E}(u_{i,k,t})\,,\ \forall k \in \mathcal{J}_t\}\,,$$

where $\mathcal{J}_t$ is the set of products in market t, and $\mathcal{I}_t$ is the (unit-mass) set of consumers who

shop at least once in market t. The market share of product j is then $s_{j,t} = \int_{i \in \Theta_{j,t}} di$, which equals the mixture in (5).

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 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 enters the demand model through the mixture in (5). 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 summarized by

$$\sum_{s \in \{m,h\}} \Gamma_{s,t} \sum_{d \in \{\text{dim1},\text{dim2}\}} \beta_s^{\text{eff},d} \left(\mathbf{E}[\textit{Efficacy}_j \mid s] \right)^d,$$

where $\Gamma_{s,t}$ is the market-level Bernoulli mean for symptom s at time t, and $\beta_s^{\text{eff},d}$ 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—knowledge that is not uniformly distributed across consumers—differences in this knowledge likely contribute to heterogeneity in behavior. While the current approach improves upon models that ignore beliefs altogether, allowing for heterogeneity in beliefs across consumer groups could further enhance predictive power and policy relevance. A potential solution is to conduct the triplet embedding component of the survey with a larger

and more diverse population group, allowing computation of expected efficacy valuations for different socio-economic groups. Based on the reduced-form analysis (Section II.C) and prior work on differences between consumers 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. Within each socio-demographic cluster, a more sophisticated method could transform responses to triplet questions into expected probabilistic efficacy (Diallo and Fürnkranz, 2021), rather than point estimates; sample size constraints prevent this in the current version.

IV.C.1. Comparability of experiment and control group embeddings

Incorporating counterfactual efficacy perceptions into the demand system requires that the spatial embeddings derived from both the control and treatment groups be directly comparable. This alignment is crucial to ensure that the structural relationships in the demand model, those between the efficacy dimension and other determinants of choice, like price, remain consistent.

Without specific adjustments, each of the four independently generated triplet embeddings, corresponding to the baseline and three counterfactual conditions, produces a distinct coordinate space. While each embedding captures the internal similarity relationships among products within its scenario, it is invariant to absolute position, orientation, and scale. As a result, their spatial configurations may differ arbitrarily across scenarios, posing a challenge for embedding them consistently into a demand model. Since the application of triplet embeddings in counterfactual structural estimation is novel, I develop and implement a methodology to align these representations to a common spatial frame. This alignment ensures that observed differences in consumer choices reflect genuine shifts in perceived efficacy, rather than artifacts of geometric transformation in the embedding space.[9]

To ensure comparability across scenarios, I begin by pooling all relative comparison responses from the baseline and counterfactual conditions. Each product is treated as scenario-specific, for example, Advil in the baseline is considered distinct from Advil under any counterfactual. This setup allows the triplet embedding algorithm to independently

[9]This approach assumes that information treatments shift perceived efficacy in a consistent direction, without fundamentally altering its structural relationship with other decision parameters. That is, average efficacy levels remain stable, and the core structure of the decision space is preserved. Alternatively, compatibility could be enhanced by combining triplet embeddings with conjoint analysis, enabling estimation of how treatments jointly affect efficacy perceptions and their interaction with attributes such as price or brand. This methodology relaxes the need for strong assumptions about comparability across embeddings or the specific tools used to achieve it.

recover the similarity structure within each scenario. Although the algorithm preserves relative distances, i.e., the internal scaling of each cluster, it remains indifferent to absolute positioning and orientation. Consequently, the four embeddings yield structurally consistent but spatially misaligned clusters.

To correct for this misalignment, I apply two alignment transformations to each counterfactual embedding. First, I translate each embedding so that its centroid coincides with that of the baseline cluster. This step ensures that the average level of perceived efficacy remains fixed across scenarios, preserving the interpretation of average substitution to the outside option and the role of other choice parameters under informational change. Second, I rotate each counterfactual embedding such that the vector from the centroid to a designated outlier matches the corresponding vector in the baseline. This rotation preserves the orientation of the embedding and maintains the economic interpretation of horizontal differentiation. By anchoring the alignment on an extreme point, I ensure that while substitution patterns between products may shift, the product furthest from the outside option in the baseline remains the furthest in each counterfactual. This preserves its relative position in the efficacy space and, by extension, its interpretive role — whether it reflects the highest or lowest likelihood of substitution to the outside option, or whether efficacy plays the most or least dominant role in determining choice compared to other parameters such as price or brand.

IV.D. Estimation

I estimate the model using a BLP-style share inversion embedded in a conformant likelihood-based estimator with exogeneity restrictions (CLER) (Grieco et al., 2022). Let θ collect all taste parameters (including β_s^{eff} and the random-coefficient parameters). For a given θ, define model shares via (5). The inner loop recovers δ_t by matching observed shares $s_{j,t}^{\text{obs}}$:

$$\delta_{j,t}^{(m+1)} = \delta_{j,t}^{(m)} + \log s_{j,t}^{\text{obs}} - \log s_{j,t}^{\text{model}}\big(\delta_t^{(m)}, \theta\big),$$

where $s_{j,t}^{\text{model}} = \Gamma_{m,t}\, \bar{s}_{j,t}^{(m)} + \Gamma_{h,t}\, \bar{s}_{j,t}^{(h)}$, and each $\bar{s}_{j,t}^{(r)}$ is computed from the type-specific utilities in (2). After convergence,

$$\widehat{\xi}_{j,t}(\theta) = \widehat{\delta}_{j,t}(\theta) - x_j'\bar{\beta} + \alpha\, p_{j,t},$$

and I estimate θ by IV-GMM using standard BLP moments with instruments $Z_{j,t}$.

IV.D.1. Migraine/Headache beliefs (micro moments assignment)

For the micro moments, I assign each individual a binary type based on the survey-recovered Bernoulli distribution that depends on observables (age and income). Specifically, for each i,

$$b_i \in \{m, h\}, \qquad b_i \sim \text{Bernoulli}\big(\rho_i(\text{age}_i, \text{income}_i)\big),$$

and I use the corresponding type-specific probability in the micro block,

$$\tilde{s}_{i,j,t} = s_{i,j,t}^{(b_i)}.$$

Micro moments are then formed in the usual way, e.g. $\mathbb{E}[(d_{i,j,t} - \tilde{s}_{i,j,t})\, h(z_i)] = 0$ for instruments or conditioning functions $h(\cdot)$ built from observables z_i. The macro type mix $\Gamma_{r,t}$ in (5) is the (weighted) sample mean of $\mathbf{1}\{b_i = r\}$ within market t.

IV.D.2. Instruments

The unobservable product-market level effect, $\xi_{j,t}$, captures product characteristics I do not observe (e.g., package color, shelf placement) that are likely correlated with prices. To address the correlation between prices and $\xi_{j,t}$, I use instrumental variables. Expected efficacy is time-invariant, so time fixed effects absorb common unobservables correlated with $q_{i,j,s}$.

I construct Differentiation IVs as in Gandhi and Houde (2023). For continuous attributes (e.g., package size), I calculate Euclidean distances of each drug to others in the same market (t) along that dimension. For discrete attributes—active components, symptom labels, brands—I use counts of products in the market that share the attribute.

IV.D.3. Demographic draws

To account for demographic variation across markets, I generate consumer-specific demographic draws by sampling 10,000 consumers with replacement per unit from the Homescan data, using NielsenIQ projection weights. Because the data do not specify which household member made the purchase, I randomly select the first or second adult in the household and use their age and gender. I exclude 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 -2.16, slightly less than the general product elasticity.

V.A. General results and elasticity

Table 10 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.[10]

	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 4: Regression Results with Mean Utility and Deviations

The average own price elasticity across markets and products is -2.16. This elasticity is

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

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), which can be explained by the nature of urgency in drug consumption. 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 -3.47.

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 29.

V.B. Brand and label preferences

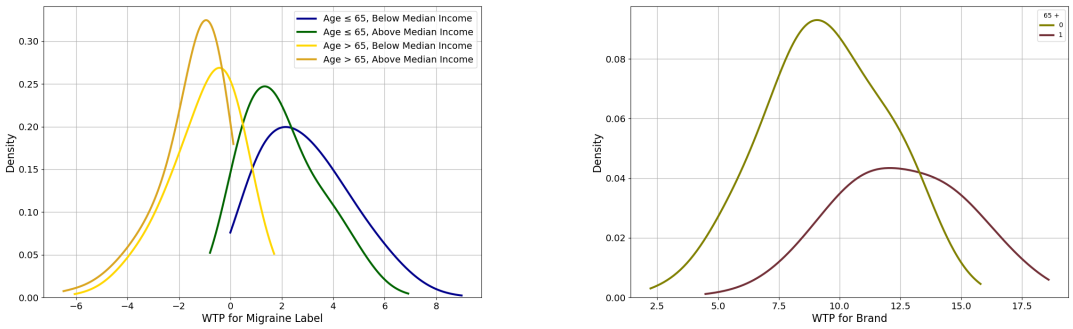
The willingness to pay (WTP) for the migraine label averages $1.2, which represents 58% of the average drug price. Individuals who are younger and have income below the median have the highest wtp for the migraine label (see Figure 10a). This WTP reflects preferences that extend beyond the belief that 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 $10.5 (Figure 10b). 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 12).

V.C. Role of expected efficacy

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

As shown in Table 5, 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 much less pronounced among migraine sufferers, highlighting the diffi-



(a) Willingness to Pay for the Migraine Label (b) Willingness to Pay for the Brand Label

Figure 10: Distribution of Willingness to Pay for Different Labels, Individual Level

culty 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 relative similarity in the latent space of perceived efficacy patterns in Section III.B.3. 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, $R^{r,DMAregion(s)}$, 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

Active Components	Brand	Migraine Label	$\beta_m^{eff} \times \mathbb{E}[\textit{Efficacy} \mid m.]$	$\beta_h^{eff} \times \mathbb{E}[\textit{Effica}$
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	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	1	-0.65	-0.80
Naproxen Sodium	0	0	-0.19	-0.64
	1	0	-0.58	-0.83

TABLE 5: Expected Efficacy by Active Components, Brand Label, and Migraine Label

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 II.A 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 II.A 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 11. Approximately 1.2% of the marginal cost estimates are negative and were excluded from

the analysis.[11] To account for idiosyncratic cost shocks across counterfactuals, where there

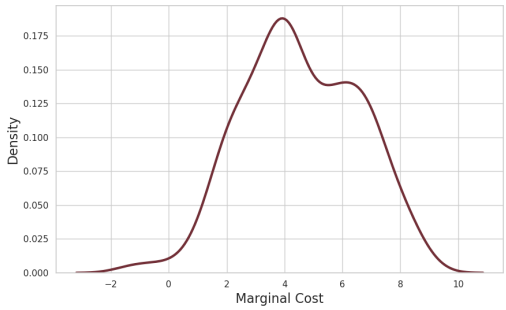


Figure 11: Marginal Costs Distribution, Product Market Level

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 13 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.

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. [12] 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).

[11] These negative values correspond to generic products, which likely have near-zero marginal costs and may be subject to strategic pricing behavior not captured in the current supply specification.

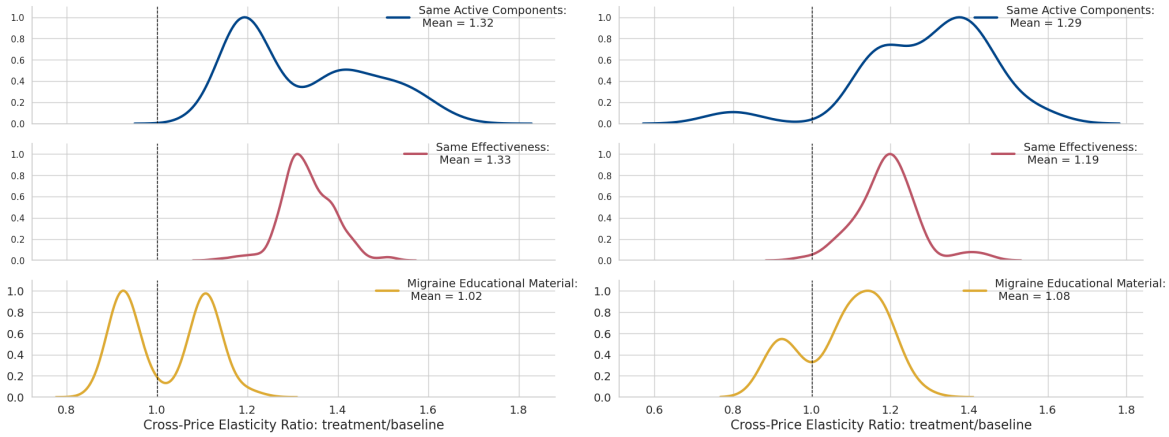
[12] There is no experimental design that can simulate a full-information scenario.

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 (III.B.3) 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 12a) and the elasticity between branded products and their generic counterparts (Figure 12b), relative to their respective baseline values.



(a) Excedrin Products

(b) Branded to Generic

Figure 12: 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 11a and Table 11b). 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.

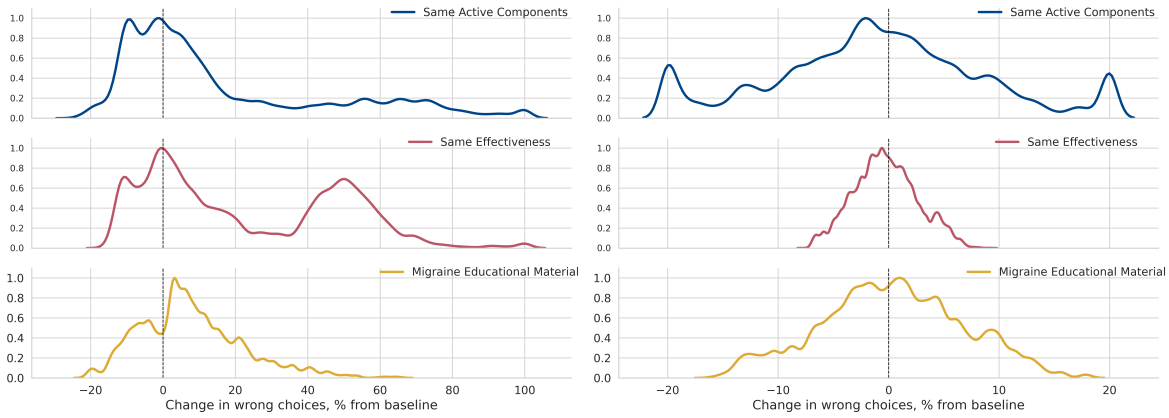
VII.A.2. Choices and "Wrong" choices

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 28). 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 13a 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 13b), 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 13: Comparison of absolute and relative changes in wrong choices

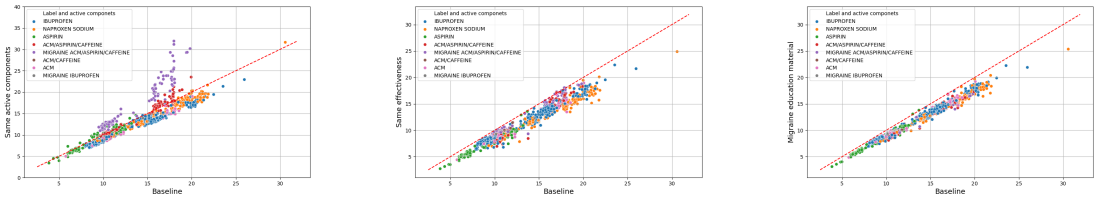
To further examine the heterogeneity in wrong choices, I analyze how increased aware-

ness 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 IV.A, 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. Equilibrium response

VII.B.1. Supply side response

Since markups are a function of elasticities, changes in substitution patterns (Online Appendix E) 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 14: 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 14a), 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 14b) and the educational material treatment (Figure 14c), 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 15 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 the 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.

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

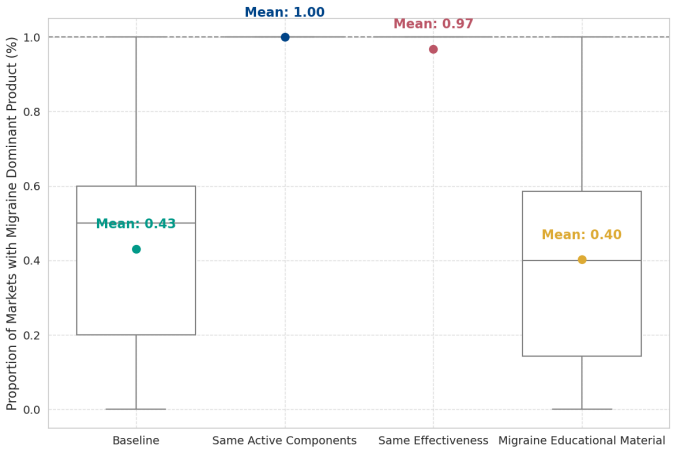


Figure 15: Box Plot: Probability that Product with Migraine Label Has Higher Price in Biologically Identical Pair with Different Labels

(Formula 8). 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, but fixes the perceived efficacy valuation on the counterfactual level. This design isolates welfare gains from labeling transparency that arise purely from improved allocation of choices, not from changes in how consumers value efficacy. In both scenarios, 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, I compute the savings amount conditional on market entry.

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

Where:

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

Change in the choice patterns along with the downward pressure on prices, in equilibrium, results in savings for all consumers in all markets (Figure 16).

The results from equation 8 are presented in Figure 17. The welfare results align with other findings from the information treatments: all treatments result in higher per

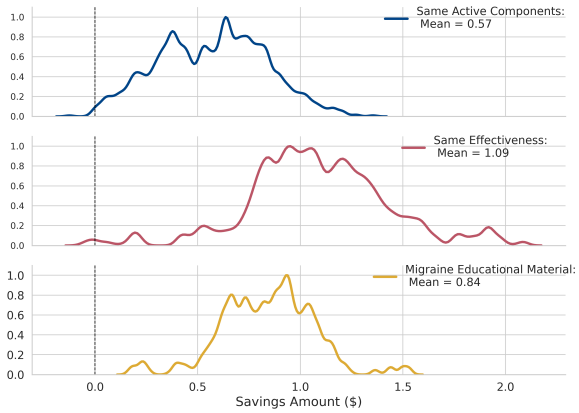


Figure 16: Individual Level Savings, New Equilibrium

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 on average $1.20 per transaction, driven by consumer savings and the fact that consumers were under-consuming products with higher "ex-post" efficacy.

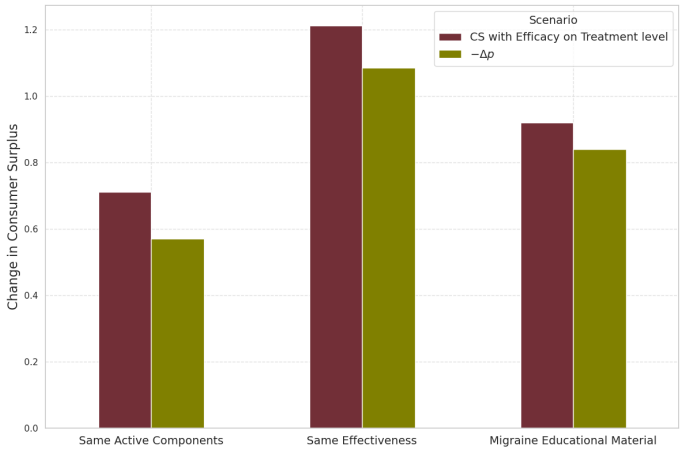


Figure 17: Chang Controlling for Expected Efficacy Levels

There are two key takeaways from this exercise. First, the current over-the-counter (OTC) drug market relies on labels that often convey complex or confusing information. As a result, consumers may struggle to accurately interpret product differences. The analysis shows that increasing label transparency—by making information easier to understand—enhances consumer responsiveness. In other words, the clearer and more accessible the labeling, the

more effectively consumers can evaluate product efficacy and make informed choices. Second, enhancing information transparency improves consumer surplus through two channels: a direct effect by improving the match between consumers and products, and an indirect effect by reshaping market pricing dynamics. Directly, it reduces misperceived differences between biologically identical products sold under the same brand, allowing consumers to make more informed choices. Indirectly, transparency reduces artificial differentiation by narrowing perceived efficacy gaps, thereby weakening unnecessary segmentation in the market. This reshapes supply-side pricing behavior: as products within the same brand become more substitutable, the patterns of substitution across brands also shift.

The main limitation of these welfare estimates is that ex-post perceived efficacy may differ from the ex-ante perceived efficacy (which I observe and estimate). While consumer choices are made based on their ex-ante utility expectations, actual surplus is ultimately determined by ex-post evaluations. In such a scenario, the expenditure savings estimated above remain valid, as they reflect observed changes in prices and choices. However, any surplus gains that include utility derived from perceived efficacy may be misestimated, since they rely on the assumption that ex-ante and ex-post valuations align.

VIII. 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.

References

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

ATAL, JUAN PABLO, JOSÉ IGNACIO CUESTA, AND MORTEN SÆTHRE (2022): "Quality Regulation and Competition: Evidence from Pharmaceutical Markets," *Review of Economic Studies*, revise and resubmit.

BARAHONA, NANO, CRISTÓBAL OTERO, AND SEBASTIÁN OTERO (2023): "Equilibrium Effects of Food Labeling Policies," *Econometrica*.

BERRY, STEVEN AND PANLE JIA (2010): "Tracing the Woes: An Empirical Analysis of the Airline Industry," *American Economic Journal: Microeconomics*, 2 (3), 1–43.

BERRY, STEVEN, JAMES LEVINSOHN, AND ARIEL PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63.

BLEWETT, LYNN A., JULIA A. RIVERA DREW, DANIEL BACKMAN, ANNIE CHEN, GRACE COOPER, MEGAN SCHOUWEILER, STEPHANIE RICHARDS, AND MICHAEL WESTBERRY (2023): "IPUMS Health Surveys: Medical Expenditure Panel Survey, Version 2.3 [dataset]," .

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

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

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

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

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

DELLAVIGNA, STEFANO AND MATTHEW GENTZKOW (2019): "Uniform Pricing in U.S. Retail Chains," *The Quarterly Journal of Economics*, 134 (4), 2011–2084.

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

DRANOVE, DAVID AND GINGER ZHE JIN (2010): "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 48 (4), 935–963.

DÖPPER, HENRIK, ALEXANDER MACKAY, JOEL STIEBALE, AND NATHAN MILLER (2024): "Rising Markups and the Role of Consumer Preferences," Resubmitted to *Journal of Political Economy*.

ELLISON, GLENN AND ALEXANDER WOLITZKY (2012): "A Search Cost Model of Obfuscation," *The RAND Journal of Economics*, 43 (3), 417–441.

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

GRABOWSKI, HENRY G. AND JOHN M. VERNON (1992): "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals After the 1984 Drug Act," *The Journal of Law and Economics*, 35 (2), 331–350.

GRIECO, PAUL LE, CHARLES MURRY, JORIS PINKSE, AND STEPHAN SAGL (2022): "Conformant and Efficient Estimation of Discrete Choice Demand Models," Unpublished Manuscript.

HARBEN, ALYSSA L., DEBORAH A. KASHY, SHIVA ESFAHANIAN, LANQING LIU, LAURA BIX, AND MARK 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.

KIESEL, KRISTIN AND SOFIA VILLAS-BOAS (2013): "Can Information Costs Affect Consumer Choice? Nutritional Labels in a Supermarket Experiment," *International*

Journal of Industrial Organization, 31, 153–163.

KÜNSTLE, DAVID-ELIAS, ULRIKE VON LUXBURG, AND FELIX A. WICHMANN (2022): "Estimating the perceived dimension of psychophysical stimuli using triplet accuracy and hypothesis testing," *Journal of Vision*, 22 (13), 5.

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

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

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

MAĆKOWIAK, BARTOSZ, FILIP MATĚJKA, AND MIRKO WIEDERHOLT (2023): "Rational Inattention: A Review," *Journal of Economic Literature*, 61 (1), 226–273.

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

MORROW, W. ROSS AND STEVEN J. SKERLOS (2011): "Fixed-Point Approaches to Computing Bertrand-Nash Equilibrium Prices Under Mixed-Logit Demand," *Operations Research*, 59 (2), 328–345.

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

SINITSYN, MAXIM (2016): "Managing price promotions within a product line," *Marketing Science*, 35 (2), 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 (1), 64–69.

TODD, PETRA E. AND KENNETH I. WOLPIN (2023): "The Best of Both Worlds: Combining Randomized Controlled Trials with Structural Modeling," *Journal of Economic Literature*, 61 (1), 41–85.

VANKADARA, LEENA C., MICHAEL LOHAUS, SIAVASH HAGHIRI, FAIZ UL WAHAB, AND ULRIKE VON LUXBURG (2023): "Insights into Ordinal Embedding Algorithms: A Systematic Evaluation," *Journal of Machine Learning Research*.

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

VILLAS-BOAS, SOFIA BERTO (2007): "Vertical relationships between manufacturers and retailers: Inference with limited data," *Review of Economic Studies*, 74 (2), 651–676.

WOŹNIAK-HOLECKA, JOANNA, MATEUSZ GRAJEK, KAROLINA SIWOZAD, KAMILA MAZGAJ, AND ELZBIETA CZECH (2012): "Consumer behavior in OTC medicines market," *Przeglad Epidemiologiczny*, 66 (1), 189–95, free article.

ZHU, CHEN, RIGOBERTO A. LOPEZ, AND XIAOOU LIU (2015): "Information Cost and Consumer Choices of Healthy Foods," *American Journal of Agricultural Economics*, 98,

41–53.

A. Additional information RCT

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

Retailer/DMA region(s)/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 who purchase drugs to relieve migraine symptoms weekly
- m_m -number of respondents who purchase drugs to relieve migraine symptoms monthly
- a_m - number of respondents who 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:

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

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 2. 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.

I.B. 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.

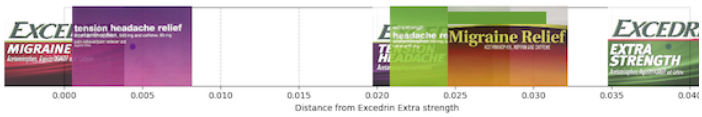
An alternative optimization function based on the hinge triplet margin loss to satisfy the set of input triplet comparisons:

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

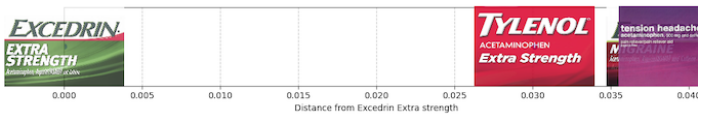
Where Adam, an algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments, is used for optimization.

I.C. Triplet embedding additional results

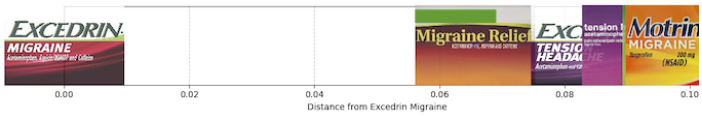
One way to illustrate triplet embedding results is by computing the relative distance between a given product and all others. Figure 18a shows the distances from Excedrin Migraine to its closest substitutes, while Figure 18b presents the corresponding distances for Excedrin Extra Strength, based on responses from individuals selecting products for migraine relief.



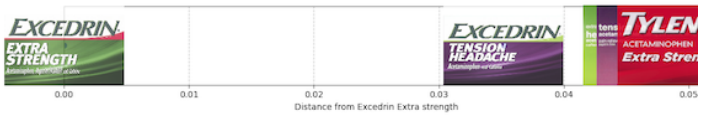
(a) Excedrin Migraine; Relieve Migraine



(b) Excedrin Extra Strength; Relieve Migraine



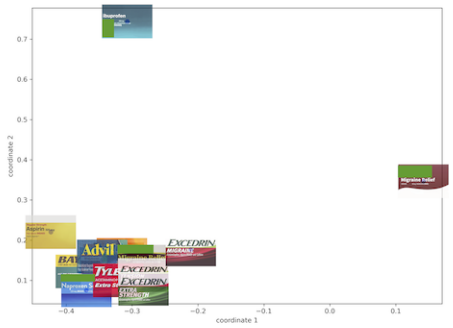
(c) Excedrin Migraine; Relieve Headache



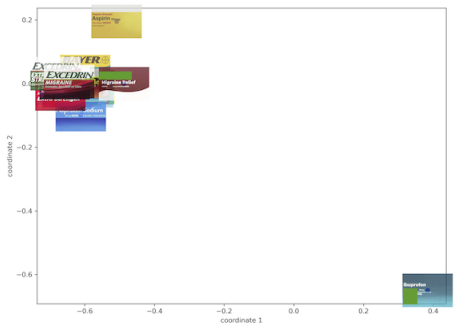
(d) Excedrin Extra Strength; Relieve Headache

Figure 18: Distance from Excedrin Products to Other Drugs for Migraine and Headache Relief

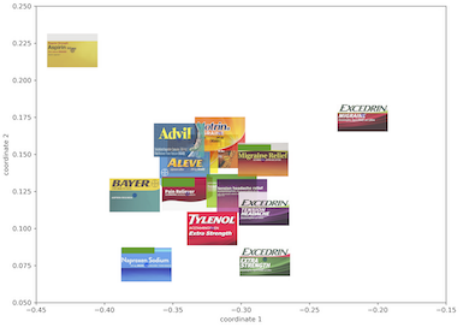
Figures 19a and 19b present the results of the two-dimensional triplet embedding for headache and migraine relief, with zoomed versions shown in Figures 19c and 19d.



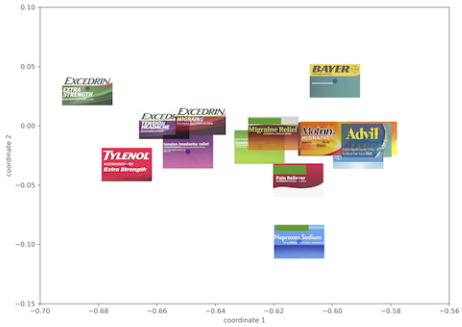
(a) Headache



(b) Migraine



(c) Headache, Zoomed



(d) Migraine, Zoomed

Figure 19: Relative and Zoomed Effectiveness Spaces for Drugs to Relieve Headache and Migraine

The results showing the distances for Excedrin Migraine and Extra Strength under the "Same Active Components" treatment are presented for migraine relief in figures: 25a and 25b, and for headache relief 25c and 25d.

The "Same Effectiveness" treatment results are presented for migraine relief in figures: 26a and 26b, and for headache relief 26c and 26d.

Finally, for "Migraine Educational Material" treatment results are presented for migraine relief in figures: 27a and 27b, and for headache relief 27c and 27d.

B. Supply side discussion

II.A. 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.

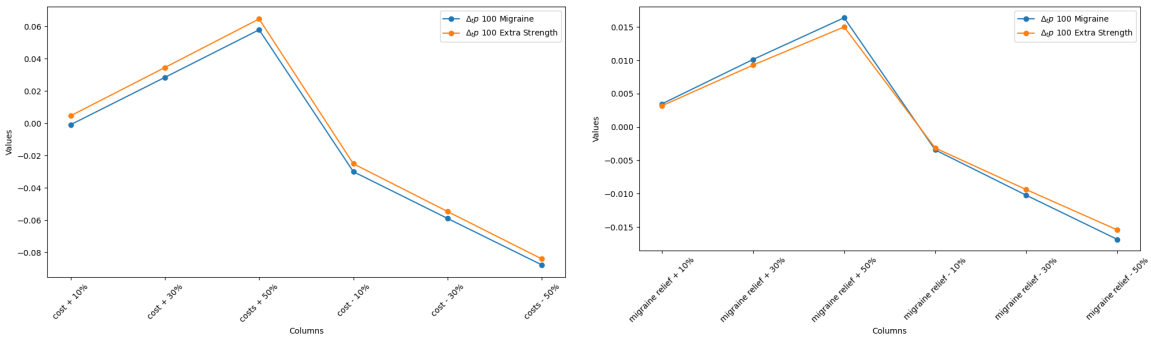
For simplicity, all the products in the market have the same active components and are biologically equivalent. Two firms produce products with brand labels and the other with generic labels. Based on the simulation results (Figure 20), both single label demand shock and production cost shock, that affects both of the products, result in the synchronous and nearly identical price adjustment for products with both labels supporting the claims discussed above.

II.B. Test of Nash-Bertrand assumptions

Assume 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 [13]:

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

[13]The two most popular products for any brand are the 24 or 100-pill package. 34



(a) Response to shock in costs

(b) Response to shock in demand

Figure 20: Responses to Shocks in Costs and Demand

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 6. 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 7. 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 33). Based on the demand estimation for the small packages, demand is inelastic or has lower elasticity than big-size products (Figure 29). Therefore, if the whole panel of products would face a 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 35).

C. 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 21), while no average savings

Dependent Variable:	$\Delta_t p_k$			
Model:	(1) $k = 20$	(2) $k = 24$	(3) $k = 80$	(4) $k = 200$
$\Delta_t p_{100}$	0.112	-0.005	0.533	-0.053
	(0.078)	(0.120)	(0.121)	(0.020)
Observations	2,736	3,853	2,391	2,903
R^2	0.021	0.017	0.154	0.010

Clustered on retailer level standard errors in parentheses.
Fixed effects: retailer code, DMA region(s) code, brand, active components + symptom label.

TABLE 6: Regression Results for $\Delta_t p_k$ with Different Values of k

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	0.6579

IID standard-errors in parentheses

TABLE 7: Regression Results for $\Delta_t p_k$ with k determine biological equivalence inside brand and size

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 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.

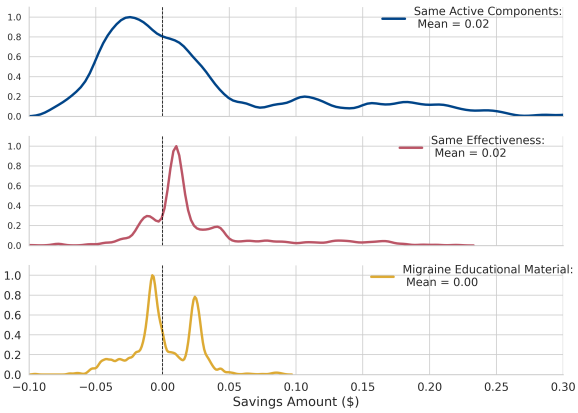


Figure 21: Result of expected efficacy change: Savings

D. "Wrong" choices

IV.A. 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) 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 10a), I do not include it in the regression. The regression formula is as follows:

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

The results of this regression are presented in Table 8. The main takeaway is that for treatments that increase substitution between biologically identical products (Figure 12a), 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 12a), 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	6.214	3.386	7.455
	(0.020)	(0.005)	(0.004)
Migraine	-13.593	-8.173	-19.734
	(0.028)	(0.007)	(0.005)
WTP for Migraine Label	-2.169	-0.556	0.525
	(0.003)	(0.001)	(0.001)
Migraine · WTP for Migraine Label	4.310	1.248	-0.227
	(0.004)	(0.001)	(0.001)
R-squared	0.346	0.312	0.933
N	3,240,000	3,240,000	3,240,000

TABLE 8: 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.

IV.B. 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:

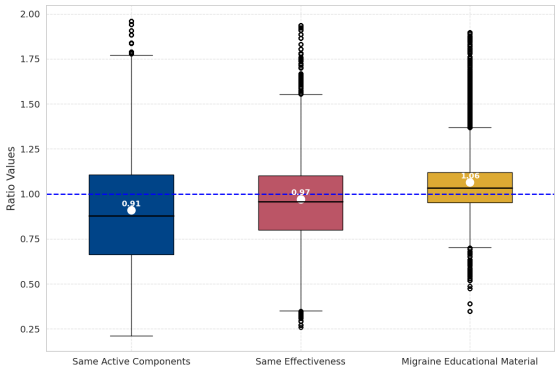
$$\frac{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(x)}, p_x\right)}{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(b)}, p_b\right)} \tag{15}$$

Where:

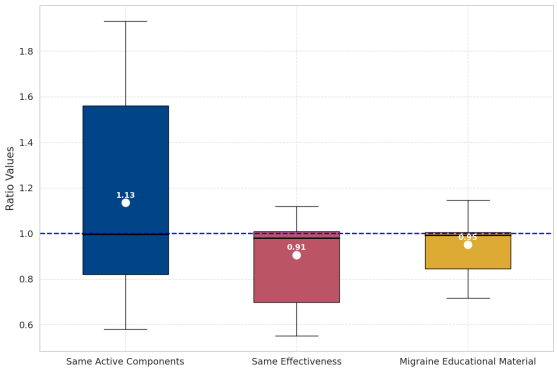
- x $\in$ { "Same Active Components", "Same Effectiveness" , "Migraine Educational Material"}
- b - baseline equilibrium

Figures 22a and 22b 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 (12a), 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 (12b).

To evaluate these ambiguities, I first decompose "wrong" choices by distinguishing whether they occur among consumers who enter the market seeking migraine relief or



(a) Between Biologically Identical Products Inside Same Brand



(b) Between Products With Same Active Components Different Brand

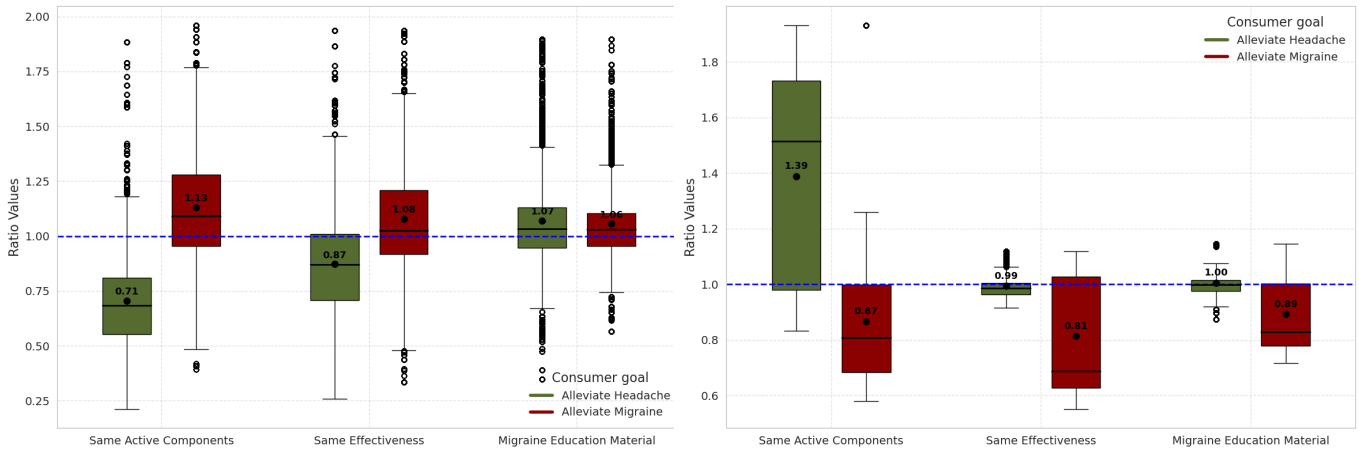
Figure 22: Box Plot with Mean as a Point: Comparison of Changes in the Probability to Make "Wrong" Choices at the Individual Level

headache relief. Figure 23a breaks down "wrong" choices within the biologically equivalent products of the same brand, while Figure 23b 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 15), 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.

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 (10b). 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.*



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

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

Figure 23: Box Plot: Decomposition of Changes in Probability to Make "Wrong" Choices at the Individual Level

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:

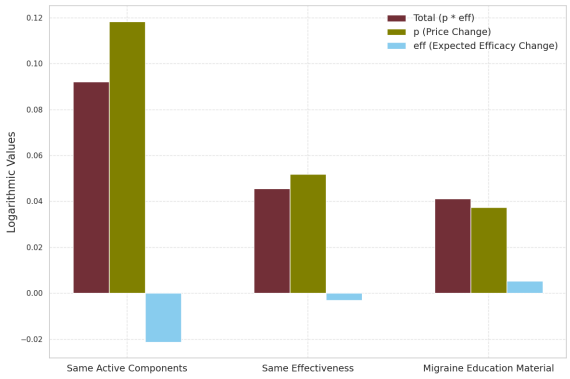
$$\frac{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(x)}, p_x\right)}{\mathbf{Pr}_{i,p}\left(\text{wrong} \mid \mathbf{E}\left[\textit{Efficacy}_j \mid s_i\right]^{(b)}, p_b\right)} \tag{16}$$

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

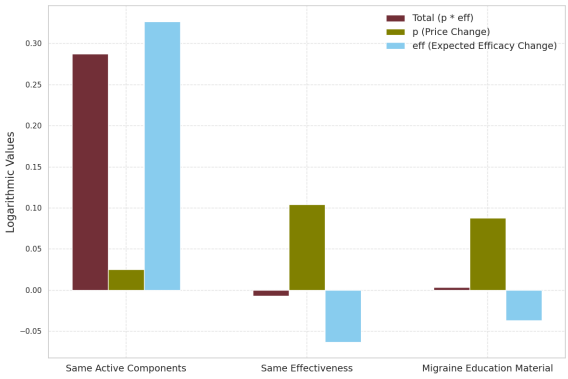
Moreover, the logarithms of both the right- and left-hand sides are taken to facilitate easier interpretation. If change in efficacy is the main driving force, the positive total value of the change in the "wrong" choices will be primarily explained by a dominant positive value in the "Impact of Efficacy" 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 24a and 24b 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.



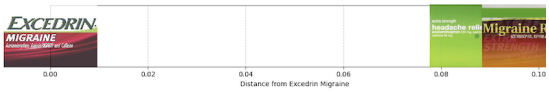
(a) Same Brand; Migraine Relief



(b) Brand vs Generic; Headache Relief

Figure 24: Decomposition of Changes in the Probability of Making "Wrong" Choices by Product and Condition

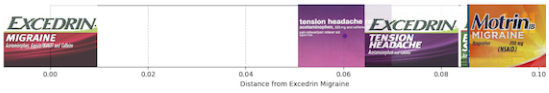
E. Additional Figures and Tables



(a) Excedrin Migraine; Relieve Migraine



(b) Excedrin Extra Strength; Relieve Migraine



(c) Excedrin Migraine; Relieve Headache



(d) Excedrin Extra Strength; Relieve Headache

Figure 25: "Same Active Components" Treatment: Distances from Excedrin to Other Drugs for Migraine and Headache Relief

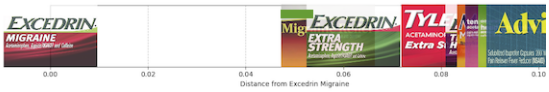
Figure 28 illustrates changes in individual product-level choice patterns, conditional on market entry, by plotting the distribution of absolute changes in purchase probabilities.



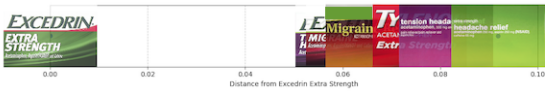
(a) Excedrin Migraine; Relieve Migraine



(b) Excedrin Extra Strength; Relieve Migraine



(c) Excedrin Migraine; Relieve Headache

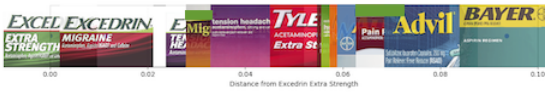


(d) Excedrin Extra Strength; Relieve Headache

Figure 26: "Same Effectiveness" 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 27: "Migraine Educational Material" Treatment: Distances from Excedrin Products to Other Drugs for Migraine and Headache Relief

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 C, I present the impact of these choice changes on total spending.

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

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

. Results of this equation are the following:

Since I assume that firms set up prices based on a market segment I recompute my

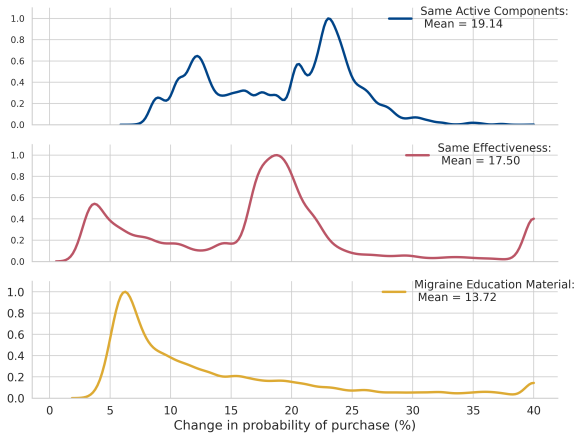


Figure 28: Effect of Expected Efficacy Change on Purchase Probability

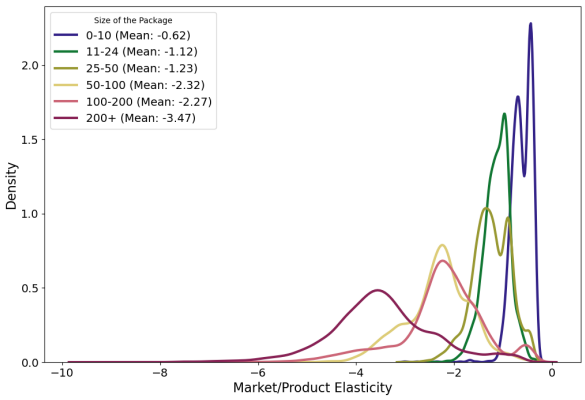


Figure 29: Distribution of Product Own Price Elasticity, Individual Level

demand model for this segment: Table 10.

Change in the cross-price elasticity across treatments is statistically different:

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 the 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 30 and 31. The major reason for the treatment of migraine education material is an increase in the substitution between products with different active components 32.

To support the hypothesis that the misbeliefs about efficacy in this market may stem, at

Dependent Variable:	$\mathbf{Pr}$("wrong choice")
chosen drugs have same label	-0.083
	(0.011)
chosen drugs have same label · tr: same active components	-0.028
	(0.011)
chosen drugs have same label · tr: same effectiveness	-0.124
	(0.011)
chosen drugs have same label · tr: education material	0.026
	(0.019)
chosen drugs have same label · group: migraine relief	0.093
	(0.013)
tr: same active components · group: migraine relief	-0.011
	(0.011)
tr: same effectiveness · group: migraine relief	-0.104
	(0.020)
tr: education material · group: migraine relief	0.090
	(0.017)
chosen drugs have same label · tr: same active components · group: migraine relief	-0.050
	(0.013)
chosen drugs have same label · tr: same effectiveness · group: migraine relief	0.055
	(0.011)
chosen drugs have same label · tr: education material · group: migraine relief	-0.208
	(0.023)
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

TABLE 9: Decomposition of the Probability of Making Wrong Choices

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 39, 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 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.

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 10: Demand Model used in the Supply Response

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 11: T-test comparison of treatment effects

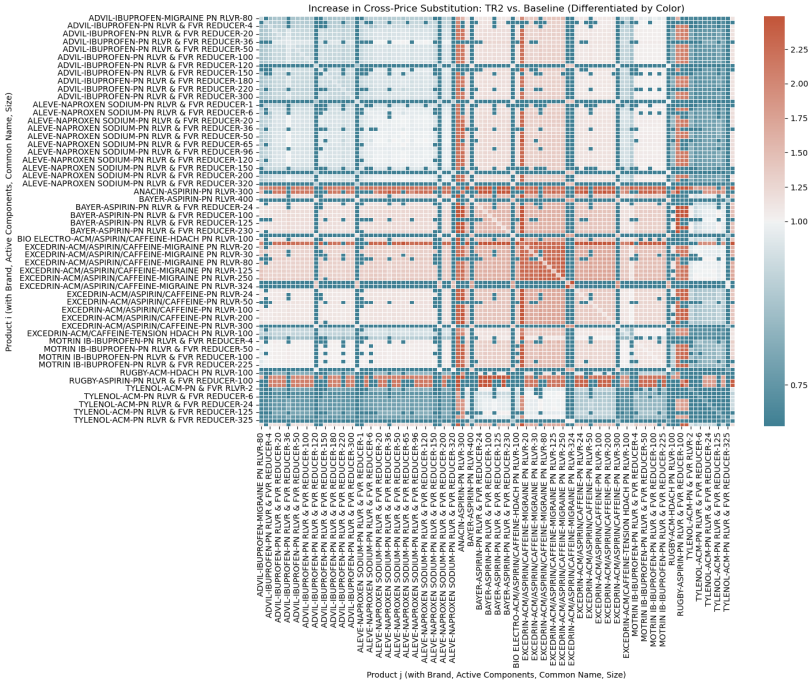


Figure 30: Ratio of the cross-price elasticities: “Same Active Components” to baseline

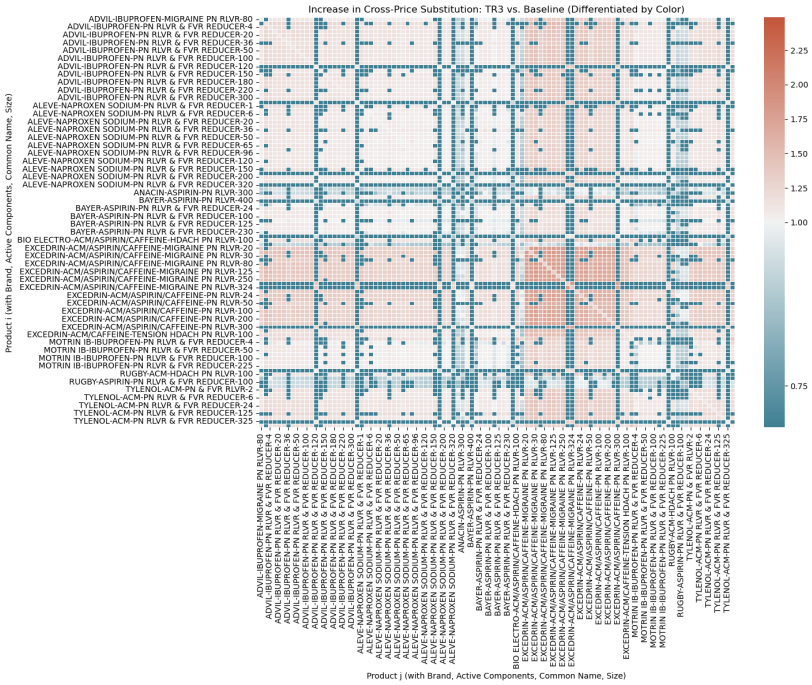


Figure 31: Ratio of the cross-price elasticities: Same Effectiveness to baseline



Figure 32: Ratio of the cross-price elasticities: “Migraine Educational Material” to baseline

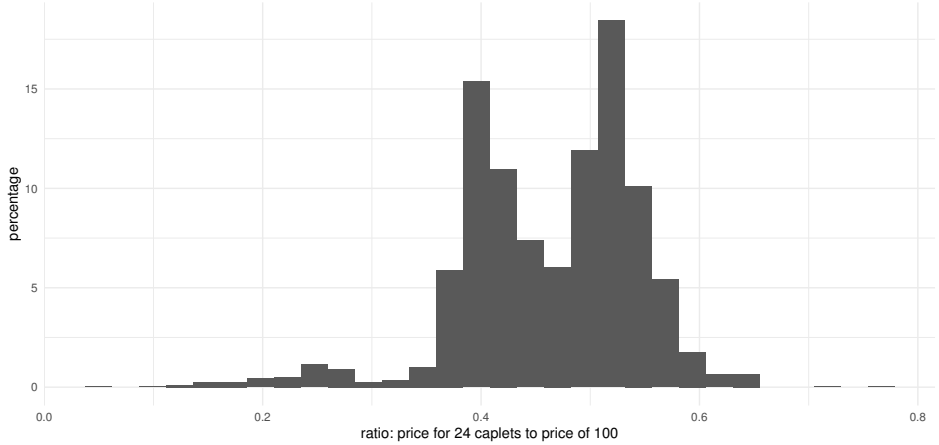


Figure 33: Distribution on the product market level: price for 24 pills to price for 100 pills

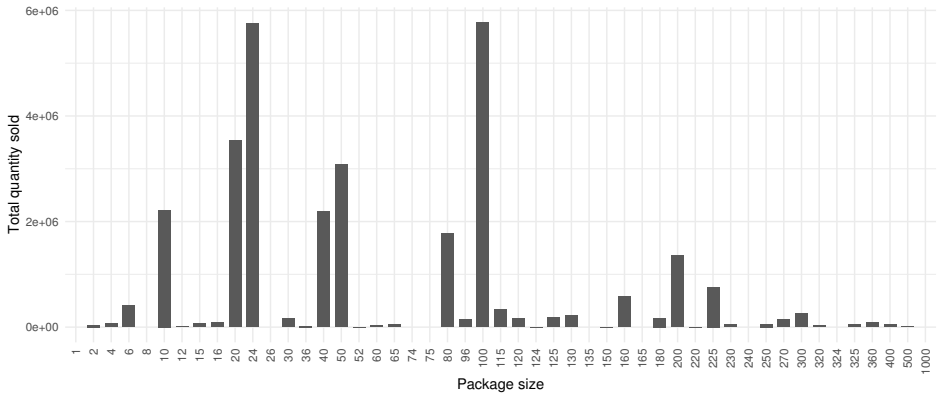


Figure 34: Distribution of total quantities sold by size of the package

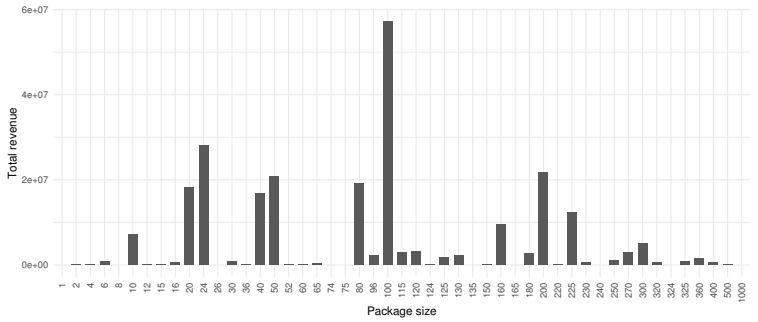


Figure 35: Distribution of total revenue sold by size of the package

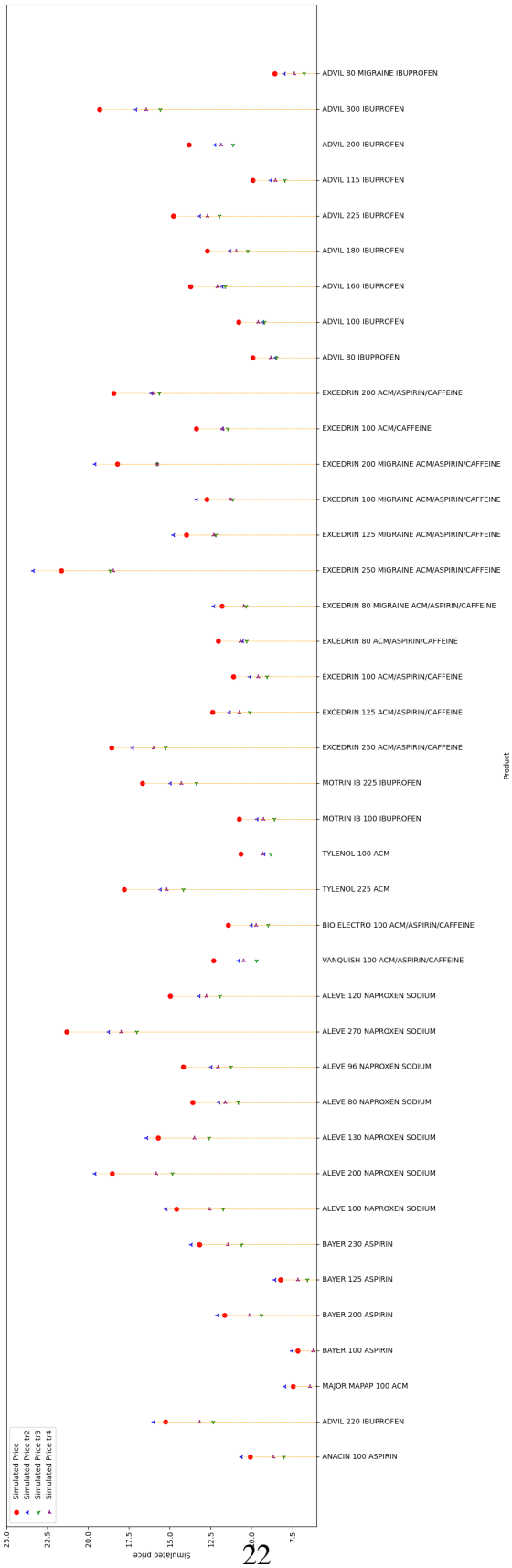


Figure 36: Example of price change, randomly selected market

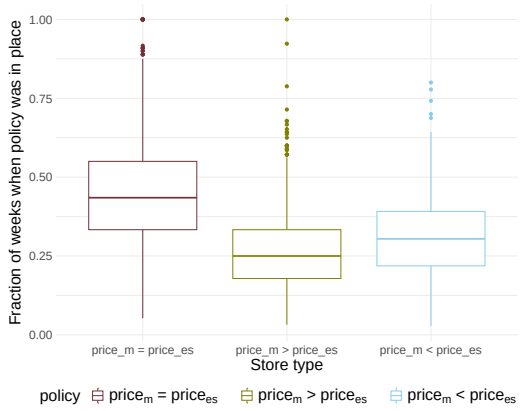
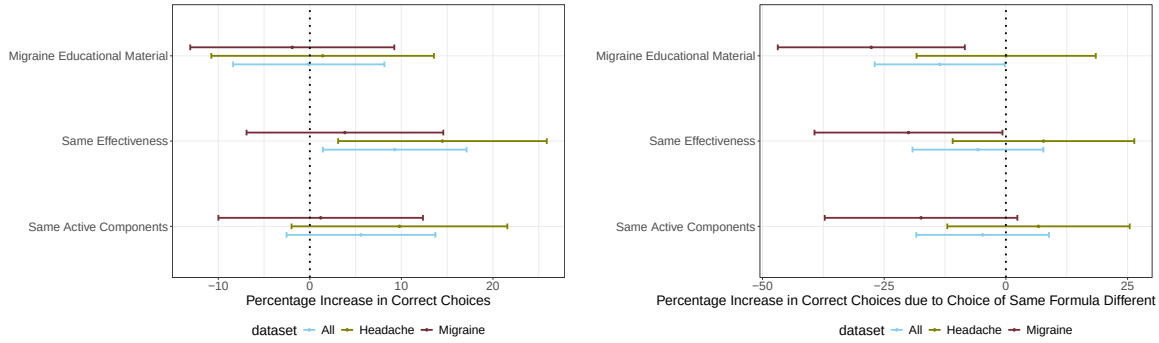


Figure 37: Pricing policy for 50 Caplets: Excedrin Migraine and Excedrin Extra Strength



(a) Choice of Generic Product

(b) Choice of Different Label Product

Figure 38: Comparison of Changes in Correct Choices Driven by Generic and Different Label Product Choices Across Treatments

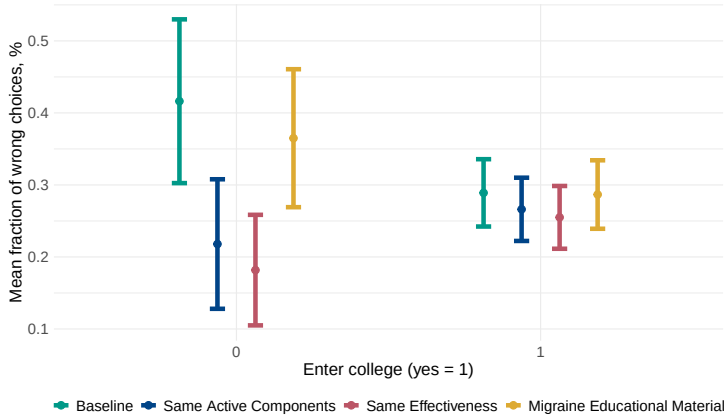


Figure 39: Impact of Label Modifications on Wrong Choices by Education Level

	Coefficient
Intercept	-0.560
	(0.040)
Acetaminophen+Aspirin+Caffeine	-0.134
	(0.025)
Acetaminophen+Caffeine	-0.041
	(0.038)
Aspirin	0.264
	(0.026)
Ibuprofen	-0.114
	(0.022)
Naproxen Sodium	-0.325
	(0.023)
Migraine Label	-0.107
	(0.022)
Brand	0.058
	(0.035)
Number of Pills	-0.006
	(0.000)
R-squared: 0.771, Observations: 2155	
Mean of Dependent Variable: -1.2183	

TABLE 12: 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 13: Regression Results: Costs Fitting