NBER WORKING PAPER SERIES

MARKET EFFECTS OF ADVERSE REGULATORY EVENTS: EVIDENCE FROM DRUG RELABELING

Matthew J. Higgins
Xin Yan
Chirantan Chatterjee

Working Paper 24957
http://www.nber.org/papers/w24957

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2018

We thank Maria Arbatskya, Lee Branstetter, David Howard, Sara Markowitz and seminar participants at Emory University for helpful comments and suggestions. We thank IMS Health We thank Maria Arbatskya, Lee Branstetter, David Howard, Sara Markowitz and seminar participants at Emory University for helpful comments and suggestions. We thank IMS Health Incorporated for their generous support and access to their data. The statements, findings, conclusions, views, and opinions contained and expressed herein are not necessarily those of IMS Health Incorporated or any of its affiliated or subsidiary entities. The statements, findings, conclusions, views, and opinions contained and expressed in this article are based in part on data obtained under license from the following IMS Health Incorporated or affiliate information service(s): IMS Midas™. The usual disclaimers apply and any remaining errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Market Effects of Adverse Regulatory Events: Evidence from Drug Relabeling
Matthew J. Higgins, Xin Yan, and Chirantan Chatterjee
NBER Working Paper No. 24957
August 2018
JEL No. I18,L51,L65

ABSTRACT

The FDA maintains post-approval safety surveillance programs to monitor the safety of drugs. As adverse events are reported, the FDA may choose to intervene and change the safety labeling associated with a drug. We provide causal evidence of the impact that these regulatory interventions have on aggregate demand for pharmaceuticals. We find that aggregate demand declines by 16.9 percent within two years of a relabeling event. After accounting for substitution patterns by physicians along with competitor actions, aggregate demand declines by 5.1 percent. Critically, this decline represents consumers that leave the market. The overall effect appears to be driven by 'high-intensity' markets or those with significant relabeling activity. Results control for the level of advertising and are robust to variation across types of relabeling, market sizes, levels of competition and degrees of cross-molecular substitution.

Matthew J. Higgins
Scheller College of Business
Georgia Institute of Technology
800 West Peachtree Street
Atlanta, GA 30308
and NBER
matt.higgins@scheller.gatech.edu

Chirantan Chatterjee
Indian Institute of Management Ahmedabad
Room 15F, Wing 15, Heritage Campus
India
chirantan@gmail.com

Xin Yan
Emory University
Department of Economics
1602 Fishburne Drive
Atlanta, GA 30307
xin.yan@emory.edu

1.0 Introduction

The drug development process is long and expensive with a low probability of receiving FDA approval (Wong *et al*., 2018).[1] As part of the approval process drug candidates undergo clinical trials designed to test their safety and efficacy. These studies culminate with large-scale, randomized clinical trials to test a drug candidate's effectiveness. After successful completion of these trials drug candidates are submitted to the FDA for approval. The FDA maintains the FDA Adverse Events Reporting System (FAERS) database to support its post-approval safety surveillance program. FAERS collects complaints and adverse events related to drugs and depending on the situation, the FDA will act upon this data and move to change the safety label associated with a drug. While the regulatory process underlying these safety label changes or relabeling is well defined, the overall impacts of those changes on demand is under investigated.

While prior studies have focused on various types of relabeling (*e.g*., Macher and Wade, 2016; Qureshi *et al.,* 2011; Dorsey *et al*., 2010), most have limited their analyses to a single or limited number of therapeutic markets (*e.g*., Olfson *et al.*, 2008; Jacoby *et al.*, 2005). These studies are important because we learn about the intricacies and nuances of specific markets but we are unable to draw overall conclusions about the impact of relabeling across markets. Using a dataset of drugs sold across all markets in the U.S. and U.K. we provide causal evidence relating to the impacts of FDA drug relabeling on aggregate consumer demand.[2] We find that, on average, aggregate demand declines by 16.9 percent within two years of a relabeling event. Our data allows us to capture intra- and inter-market substitution patterns as well as competitive responses. Critically, after accounting for these factors we still find that demand declined by 5.1 percent, an estimate that plausibly represents consumers that prematurely leave the market.[3]

Our results control for the level of advertising and are robust to variation in the types of relabeling activity, market sizes, level of market competition and degree of cross molecular substitution (CMS). Interestingly, in markets with low levels of relabeling activity we find declines in aggregate focal drug demand completely absorbed by intra-market substitution or physicians switching patients to other drugs within that same market. In contrast, in markets with high levels of relabeling activity, after

[1] Wong *et al* (2018) place the probability of a drug candidate reaching FDA approval at 13.8 percent.
[2] For ease of exposition we use the term aggregate consumer demand interchangeably with demand. To be precise we are referring to aggregate consumer demand. Our data is at the standard unit level and not at the individual prescription level. Standard units are determined by IMS Health and are intended to equate pills, tablets and liquids.
[3] We engage in back of the envelope calculations to translate these changes to represent numbers of prescriptions and consumers. As a conservative lower bound, we can assume that the entire decline, after accounting for substitution patterns, represents consumers that are chronic patients. This allows us to transform the decline in aggregate demand into monthly prescriptions and as a result an estimate of chronic patients. Clearly, all conditions are not chronic so as the number of acute prescriptions increases in the sample so will the number of consumers that plausibly leave the market.

accounting for plausible substitution patterns, we find consumers leave the market. Finally, we follow prior literature (Branstetter *et al*, 2016 and 2014) and split our sample across varying degrees of CMS. Cross molecule substitution in our analysis measures the ability of patients to switch easily from one molecule to another within the same disease market induced either by the insurer or by the prescribing physician. In low CMS markets, such as those relating to the nervous system, we find that declines in aggregate focal drug demand are again absorbed by intra-market substitution. However, in high CMS markets, such as anti-infectives, we find that consumers leave the market.

Our findings have implications for firms. Our results for focal firms suggest that their current efforts to counteract the impacts from these negative product shocks, on average, appear to be failing. Importantly, the magnitude of our results for relatively minor safety relabeling suggests that physicians may be proactively shifting consumers (or those consumers are requesting to be switched) to other drugs. This implies that while detailing (*i.e.*, direct advertising to physicians) may be effective at influencing initial physician prescription behavior (*e.g.*, Datta and Dave, 2017), this influence appears to break down when confronted with negative safety information. Unfortunately, while we can detect the shift in behavior we can only conjecture on the underlying motivations driving physician behavior. Explanations range from physicians practicing "defensive medicine" and switching consumers due to potential liability concerns to physicians being concerned that less serious safety concerns will eventually unmask more serious concerns or simply being induced by competitor firm detailing efforts.

Finally, there are plausible welfare implications from our findings. If consumers that leave the market should be treated, then this shift to the non-treated population could be a detriment to welfare. Moreover, if consumers remain treated but are switched to drugs that are less effective, this will again be a detriment to welfare. On the other hand, it is widely believed that some drugs are overprescribed (Lembke *et al*., 2018; Sacarny *et al*., 2016; Forgacs, 2008; Price *et al*., 1986). If it is these consumers that exit the market then the impact on welfare will be dampened. Additionally, it is plausible that the benefits of having avoided an adverse effect outweigh the negative impacts we document. While we cannot make definitive welfare statements, our findings do suggest that further work is warranted. Ultimately, the answers to these questions may suggest that changes are needed to how policymakers handle relabeling events.

2.0 FDA drug relabeling

The pharmaceutical industry in the U.S. is highly regulated and drug candidates undergo rigorous clinical testing prior to being submitted to the FDA for approval. During this rigorous process possible risks and side effects of a drug candidate are identified. This information becomes part of the FDA approved label and drug insert that accompanies a newly approved drug. Unfortunately, some side effects do not become known until after a drug has been approved. To help with the reporting and collection of

these adverse events the FDA founded MedWatch in 1993. Healthcare professionals or consumers (patients) can voluntary report to Medwatch. In more recent times this adverse events data has been made available via FAERS.[4]

During the post-approval time period the FDA monitors adverse reporting along with results from post-approval studies and peer-reviewed literature. Negative safety-related information is scrutinized and the FDA can form an investigation team to determine if a safety label update is needed. If the FDA believes a safety label change is warranted the manufacturer is notified and is required to report back to the FDA within a predetermined period. The agency works privately with a manufacturer to determine which type of safety label change will be made. At the end of the process the FDA will publish this information online while allowing firms additional time to change actual printed material. [5] Prior to 2016 product safety data was available via MedWatch but has since shifted to the FDA Drug Safety Label Change database.

The main safety labeling changes that the FDA issues include: *adverse reaction*, *precaution*, *warning*, *contraindication*, and *box warning*. These classifications serve to inform physicians and consumers of possible health concerns that have been clinically identified, anticipated to occur, or associated with unapproved uses. A box or "black box" warning is the most severe of type of label change and is intended to communicate potentially severe health risks resulting from taking a drug. For example, in 2004 the FDA issued box warnings for all anti-depressant drugs over the concern that these drugs could lead to suicide in patients younger than 18 years of age.[6] While these changes can be sensational, most changes are much less so. For example, Topomax® was the target of a "precaution and warning" label change due to an elevated risk of kidney stones.[7] As this example demonstrates, drugs can undergo several types of relabeling simultaneously. Additionally, a drug that has been relabeled can undergo additional safety label changes in the future, if warranted.

3.0 Adverse regulatory events

We draw on several strands of literature starting with the economics of regulation. Early work in this area theorizes on the impact of regulation on consumer and firm behavior (*e.g.*, Stigler, 1971; Peltzman, 1976; Migue, 1977). Brown *et al*. (1964) argued that regulation could be viewed as an information transmission process. As consumers receive new information they are able to update and change their behavior. Subsequent work built on this idea to show how information influences consumer perception of product quality (Zeithaml, 1988) and how behavior changes with positive information (Nelson, 1970). In contrast, Hartley (1994) showed how negative product information led to decreased

[4] http://www.nber.org/data/fda-adverse-event-reporting-system-faers-data.html.
[5] https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm250783.pdf
[6] https://www.medpagetoday.com/psychiatry/depression/210
[7] https://www.accessdata.fda.gov/scripts/cder/safetylabelingchanges/index.cfm?event=searchdetail.page&DrugNameID=1063

sales. More broadly, Oberholzer-Gee and Mitsunari (2006) examined how non-related negative events, in their case the release of pollution information, decreased property values. In our context, the process of relabeling in the pharmaceutical industry can be viewed as an information transmission process that could subsequently impact physician (or consumer) behavior.

Our paper also draws on studies in health economics where scholars have explored the implementation of regulatory procedures on public health (*e.g.*, Gruenspecht and Lave, 2006). In the case of the pharmaceutical industry, Dranove (2011) stresses the importance of quality certification for efficient and optimal regulation. For drugs, this certification comes in the form of the FDA approval process. This process can be divided into pre- and post-approval stages. The pre-approval stage includes clinical testing and provides the first line of defense to ensure safety and efficacy of products. This creates a tension, however, for regulators between length of trials and getting new drugs to market. For example, adverse events have been increasing (Moore *et al*, 2007) and have been associated with declines in pre-approval times (Olson, 2003). This makes post-approval safety monitoring critically important. In recent times FAERS has served as an important source of data for updating safety labeling information (Wysowski and Swartz, 2005).

Based on these prior studies, adverse safety related information should improve physician (or consumer) awareness about the potential safety of a drug and lead to changes in behavior. Presumably physicians (or consumers) may shift away from a drug given a safety concern. A number of studies focused on specific therapeutic markets support this association (*e.g.*, Dranove and Olsen, 1994; Smalley *et al*., 2000; Cheah *et al*., 2007; Olfson and Marcus, 2008; Tekin and Markowitz, 2008; Bunniran *et al*., 2009; Dorsey *et al*., 2010; Kales *et al.*, 2011; Dusetzina *et al*., 2012; Briesacher *et al*., 2013; Lu *et al*., 2014). Prior work also documents that this association could be differential; new drugs tend be impacted more than existing drugs (Wilkinson *et al*., 2004) and geographic variation could cause the usage of a warned drug to be different (Shah *et al*., 2010).

What remains unknown from this batch of prior work is what constitutes a rational, medically appropriate response? One might expect consumers, through their physicians, to be switched to other drugs as the severity of events increase. Moreover, consumers may just decide to stop treatment altogether and leave the market. Importantly, these studies are unable to distinguish between these two possibilities.[8] Prospect theory (Tversky and Kahneman, 1992; Kahneman and Tverksy, 1979) provides one explanation as to why consumers may ultimately leave the market. These consumers, when confronted with new information about a drug, may vastly overestimate the probability of a negative event within their own weighting function. They may also incorrectly attribute these same negative

[8] Given the breadth and depth of our data this is a distinction we will be able to analyze. Many prior studies lack sufficient data to capture all the intra- and inter-market substitution patterns that are possible.

effects to substitute drugs that physicians prescribe. As a result, they conclude that the benefit does not outweigh the risk and exit the market.

There is experimental evidence that reinforces these negative actions. For example, Bunniran *et al.* (2009) study blame and trust due to pharmaceutical products withdrawn as a result of safety related concerns. They found that consumers taking the withdrawn drug or those taking another drug within the same class were highly likely to blame pharmaceutical companies and the FDA. After an event trust in both institutions was also fairly low. These declines provide one plausible explanation as to why consumers may formulate and attribute the negative effects described by Tversky and Kahneman (1992) and Kahneman and Tverksy (1979) to a focal drug or a substitute.

The way in which consumers appear to be responding to adverse safety events suggests that there are potential welfare implications. On the one hand, with better information consumers can be more accurately and effectively treated, which should enhance welfare. On the other hand, however, if consumers are irrationally switched to inferior drugs (in terms of treatment) or choose to leave the market all together, this could diminish welfare. Given the tension between moving drugs through clinical trials in a timely manner and post-approval safety events, these welfare implications are important for regulators and for the implementation of policy.

More broadly, the answer contributes additional evidence to the literature on welfare effects of regulation. By its nature, regulation should be welfare enhancing but ample evidence exists that this may not always be the case. For example, the milk industry was regulated in the 1960s but that regulation was shown to be a detriment to welfare (Kessel, 1967). Bartel and Thomas (1987, 1985) found that the Occupational Safety and Health Administration did not have a significant impact on national injury rates. Similarly, Ter-Martirosyan and Kwoka (2010) found that incentive regulation caused quality degradation in the U.S. electricity industry. In healthcare, mandatory prescription regulation has not significantly improved health outcomes (Peltzman, 1987) and in hospitals regulation of quality standards has inflated costs and diminished patient welfare (Sloan and Steinwald, 1980).

Finally, regulation can also have unintended consequence from spillovers. For example, toy recalls due to safety reasons tend to cause negative industry-wide spillover effects for similar types of toys (Freedman and Lederman, 2009). In our setting, such spillovers would manifest in the drugs within the same market or related market as the focal drug that is relabeled due to a safety concern. Following this literature, we intend to capture both the direct and indirect (*i.e.*, spillovers) effects of drug relabeling due to adverse safety concerns.

4.0 Empirical strategy and data

4.1. Empirical strategy

We exploit FDA relabeling events to estimate a difference-in-differences (diff-in-diffs) specification. As we discussed above, the relabeling process involves private interaction between the FDA and focal firm and remains unknown to consumers and physicians prior to formal action. We use two groups of observations. The first group (treated) includes drugs sold in the U.S. Because FDA relabeling events only affects drugs sold in the U.S., our treated group is exposed to treatment in the post-relabel period but not in the pre-relabel period. The second group (control) is comprised of the same drugs as those in the treated group but sold in the U.K. Our identification strategy relies on the fact that the control group is not exposed to treatment in either period (see Figure 1).[9] Importantly, the FDA does not have regulatory jurisdiction over drugs sold in the U.K.

Using a traditional diff-in-diffs approach, we estimate the following model (controls are omitted for simplicity):

$$(1) \qquad Y_{it} = \beta_0 + \beta_1 Relabel_t + \beta_2 US_i + \beta_3 (Relabel_t \, x \, US_i) + \beta_4 Price_{it} + \beta_5 (Lagged\ promotion\ stock_{it}) + \varepsilon_{it}$$

where Y_{it} is demand (*i.e.*, drug sales). $Relabel_t$ is a dummy variable for the post-treatment period represented by drug relabeling events and captures aggregate factors that would cause changes in Y_{it} even in the absence of the treatment. US_i is a dummy variable and captures possible differences between the treatment and control groups. This base specification is estimated at differing levels of aggregation so it will include varying sets of fixed effects. The coefficient of interest across all models is β_3 and it represents the impact induced by drug relabeling events on U.S. drugs relative to U.K. drugs.

4.2. Data

Our sample consists of all drugs sold in both the U.S. and U.K. during 2003 to 2009 as identified by IMS MIDAS™. Relabeling data for drugs sold in the U.S. was collected from the FDA MedWatch database and we restricted that data to those drugs that experienced a first-instance of a drug relabel.[10] Relabeling data for drugs sold in the U.K. was gathered from Datapharm's electronic Medicines Compendium that covers all drugs approved by the U.K. Medicines and Healthcare Products Regulatory Agency (MHRA).[11] In order to create a clean control group we further restricted our treated drugs to include only those that experienced a relabel in the U.S. but no relabel in the U.K. within eight quarters of the U.S. relabel. Table 1 provides the distribution of relabel activity in the U.S. and U.K. For those drugs

[9] It is possible that drugs in the U.K. can undergo relabeling by the U.K. Medicines and Healthcare Products Regulatory Agency. We discuss this issue in Section 4.2.

[10] It is possible to have multiple different types of relabeling activity at the same time. This is not a concern for our baseline models. However, when we examine the variation across types of relabeling activity we include those observations in each type of relabeling activity. We focus on four types of relabeling events: precaution, adverse reaction, warning and box warning. There was only one first-instance of a contraindication that met our sample criteria. It was excluded from the final sample; our results do not change with this exclusion.

[11] https://www.medicines.org.uk/emc/

that were subsequently relabeled in the U.K. the average time until relabel was 12.95 months after the relabel event in the U.S. This was shorter than 18.5 months documented by Pfistermeister et al (2013) for a limited sample of psychiatric drugs.[12] Importantly, we could find no evidence that drug relabeling in the U.S., on average, systematically impacted contemporaneous physician prescription patterns in the U.K. (see Figure 1). This further validates our U.K. sample as a clean control for causal estimates in our study.

Next, we gathered quarterly drug-level sales, detailing (promotions), and price data from IMS MIDAS™. Sales or quantity data is standardized by IMS into a 'standard unit' that equates pills, tablets and liquids. The data for both the U.S. and U.K. includes both hospital and retail channels. IMS MIDAS™ includes all branded and generic drugs and covers every therapeutic category. Detailing or direct-to-physician promotion data is available for all approved drugs. Financial variables from the U.K. have been converted by IMS to U.S. dollars and all financial variables have been converted to real 2009 dollars using a GDP deflator.[13] Descriptive statistics are presented in Table 2.

Note that drugs are approved for use within 4-digit anatomical therapeutic chemical (ATC) markets. The ATC classification is controlled by the World Health Organization and was designed to categorize drugs into different groups according to the organ or systems that they treat.[14] There are four different levels of classification ranging from the most aggregate (1-digit ATC) to most disaggregate (4-digit ATC). For example, 1-digit ATC market N comprises drugs for the nervous system. Within ATC N there are seven 2-digit ATC markets that contain 19 3-digit ATC markets. Each of these 3-digit ATC markets, in turn, contains 4-digit ATC markets. An advantage of our data is that it is available at the 4-digit ATC market level and can be aggregated as needed.

4.2.1. Dependent variable

Our baseline dependent variable is focal drug sales (quantity) in standard units as determined by IMS. Sales are aggregated across varying dosages to the drug level since a relabeling event will impact the drug similarly across dosage types. We define *Sales* as the natural logarithm of quarterly focal drug sales plus one. In addition to the baseline focal drug level, we will consider two additional aggregate models. First, we consider sales of all drugs within a focal drug's 4-digit ATC market. These drugs can be reasonably viewed as substitutes. For example, both anti-viral drugs Invirase® and Norvir® are contained in the 4-digit ATC market J5AE (protease inhibitors). Importantly, this aggregation allows us to capture intra-market substitution by physicians.

[12] In Appendix Table 1 we extend the time frame for our baseline model from eight quarters to 12 and 16 quarters; our results remain robust to these longer time frames.
[13] It is critical to note that the price data within IMS MIDAS™ is a wholesale price. It does not include adjustments as a result of back-end rebate payments or any other discounts that may be offered to insurance or prescription benefit companies.
[14] For a more detailed discussion: https://www.whocc.no/atc_ddd_methodology/purpose_of_the_atc_ddd_system/

Second, we move up one more level of aggregation to the 3-digit ATC market. At this level of analysis we capture all drugs within multiple 4-digit ATC markets but contained within the same 3-digit ATC market.[15] For example, the two 4-digit ATC markets J5AE (protease inhibitors) and J5AE (nucleotide reverse transcriptase inhibitors) are contained within the 3-digit ATC market J5A (direct acting antivirals). This level of aggregation allows us to capture inter-market substitution by physicians.[16]

4.2.2. Independent variables

Our sample includes drugs that were sold both in the U.S. and U.K. We define *U.S.* as a dummy variable that equals one if the drug was sold in the U.S., zero otherwise. In order to implement our diff-in-diffs strategy, we define a dummy variable (*Relabel*) that equals one for all observations after a drug's first relabeling event, zero otherwise. *Relabel* encompasses four types of events: precaution, adverse reaction, warning and box warning.

Prior work has demonstrated the importance of detailing on physician prescription behavior (*e.g.*, Datta and Dave, 2017; Manchanda and Honka, 2005) and reducing price elasticity (Windmeijer *et al*, 2006; Rizzo, 1999). However, contemporaneous detailing is a function of current sales, which creates a reverse causal relationship. To resolve this issue we use lagged promotion stock as studies have shown that promotions have a carry-over effect (*e.g.*, Zhao *et al*, 2011). Prior promotion expenditures should not be impacted by contemporaneous sales. As such we define *Lagged promotion stock* as the discounted sum of the prior three quarters detailing expenditures. We follow the literature (Leone, 1995) and use a 70 percent discount rate, however our baseline results are not sensitive to inclusion or variation in the discount rate.[17]

Focusing on the Type-2 diabetes market and the case of box warnings, Macher and Wade (2016) found that affected firms took strategic actions with respect to promotions to mitigate losses from the relabeling event. *Lagged promotion stock* in the focal drug-level models will capture these effects. They also found that competitors take advantage of these adverse events by increasing promotion activity in order to try to steal market share. *Lagged promotion stock* in the aggregated models at the 4-digit and 3-

[15] These markets can be explored at: https://www.whocc.no/atc_ddd_index/?code=J05A.

[16] As a robustness check, and as a method to verify we have captured all reasonable substitution patterns, on average, we aggregate markets up one more level to the 2-digit ATC market. At this level of aggregation we capture all 3-digit ATC markets contained within a 2-digit ATC market. Each of those 3-digit ATC markets will include 4-digit ATC markets. For example, let's consider the 2-digit ATC market J04 (antimycobaterials). It contains two 3-digit ATC markets, J04A (drugs for treatment of tuberculosis) and J04B (drugs for treatment of lepra). The 3-digit ATC market J04A contains six 4-digit ATC markets: J04AA (aminosalicylic acid and derivatives), J04AB (antibiotics), J04AC (hydrazides), J04AD (thiocarbamide derivatives), J04AK (other drugs for the treatment of tuberculosis), and J04AM (combinations of drugs for the treatment of tuberculosis). The 3-digit ATC market J04B contains one 4-digit ATC market, J04BA (drugs for the treatment of lepra). Like our 3-digit ATC market level of analysis this 2-digit ATC market level of analysis can also be viewed as capturing inter-market substitution.

[17] Following Leone (1995) we vary the discount rate between 50 and 70 percent.

digit ATC market-level will control for these competitive dynamics. These latter two models will also capture and control for any affected firm promotion response.

Next, we control for several drug and market characteristics that may influence sales or demand. First, we define *Vintage* as a measure of elapsed time, in quarters, from introduction. Drugs that have been on the market longer have time to build up brand loyalties with consumers and physicians even though they may become 'outdated' as newer treatments come to market. Finally, we include count variables for the *Number of brands* and *Number of generics*. The former controls for the intra-market substitution possibilities or the ability of physicians to switch patients to another drug within the same therapeutic market (often referred to 'me-to' drugs). The latter controls for cross-molecular substitution or the insurance companies ability to attempt to influence physicians to switch patients to a generic of another branded drug within a focal drug therapeutic market (Branstetter *et al*, 2016 and 2014).[18]

4.2.3. Endogeneity of price

As indicated above, for those drugs that have multiple dosages sold by the same firm we aggregate the data together to the drug-level. We define *Price* by dividing drug-level revenues by the number of drug-level SU sold. It is important to note that we are capturing wholesale price and this does not include any unmeasured discounting (rebates) by pharmaceutical companies, which is not commercially available. This price variable, however, will be highly correlated with ultimate consumer price and as such will be endogenous.[19] To address this concern we follow the literature (*e.g.*, Nevo, 2001) and use the mean and median price of other drugs in closely related markets as instruments for the focal drug's price. Specifically, we use the mean and median price of other drugs within the same 2-digit ATC market. For example, if our focal drug is a MAO-inhibitor (4-digit ATC market C02KC) we take the mean and median price of drugs in the broader 2-digit ATC market, C02 (anti-hypertensives). Drugs within the same 2-digit ATC should, on average, be correlated due to similar marginal costs but uncorrelated with the focal drug's unobserved product characteristics. The instruments pass the usual tests and are reported in the bottom panel of each table.

5.0 Empirical findings

5.1 Impact of drug relabeling on demand

In Table 3 we present empirical results from Equation 1. Model 1 presents estimates at the focal drug level, Model 2 presents estimates at the 4-digit ATC market level and Model 3 presents estimates at

[18] As an example, assume there are three branded drugs within a market, Brand A, Brand B and Brand C along with two generic drugs, Generic B and Generic C. Cross-molecular substitution refers to an insurance company trying to convince physicians to shift patients from Brand A to either Generic B or Generic C, often by providing economic incentives to consumers.

[19] A significant body of prior research on the pharmaceutical industry uses earlier versions of the IMS Health data that we employ here. Like us, these prior researchers do not directly observe retail sales or prices.

the 3-digit ATC market level. Model 1 can be viewed as testing the casual impact of drug relabeling on aggregate focal drug demand while Model 2 captures intra-market drug substitution. In other words, Model 2 helps us understand if physicians switch consumers to another drug in the same 4-digit ATC market. An example of such a substitution would be a switch from the anti-viral Invirase® to Norvir®. Finally, Model 3 captures inter-market drug substitution. In this case, physicians switch patients to another drug in a different 4-digit ATC market but within the same 3-digit ATC market. In the prior example, both Invirase® and Norvir® are in the 4-digit ATC market J5AE (protease inhibitors). In the current example, a physician would be switching a patient from either of those two drugs to Retrovir®, which is in the 4-digit ATC market J5AF (nucleotide reverse transcriptase inhibitors). All three drugs are treatments for HIV and both 4-digit ATC markets, J5AE and J5AF, are contained within the 3-digit ATC market J5A (direct acting antivirals).

The dependent variable across all three models is *Sales* and includes our full set of controls. In Model 1 we include drug and time fixed effects while in Models 2 and 3 we include market and time fixed effects. *Price* is instrumented in all models and pass the usual test statistics, which are reported at the bottom of the table. Standard errors are clustered at the 2-digit ATC market level. The coefficient of interest is the interaction term (*Relabel * U.S.*); it is negative and statistically significant across models. In Model 1 we find a 16.9 percent decline in aggregate focal drug sales caused by the first instance of a drug relabel. When we aggregate within 4-digit ATC markets in Model 2 we find a 5.1 percent decline in aggregate sales. Importantly, this model accounts for demand of the focal drug that was absorbed by other drugs *within* that same 4-digit ATC market. In other words, physicians engaged in intra-market substitution and switched patients to another drug within the same therapeutic market. From the previous example, this would be a switch from Invirase® to Norvir® within the 4-digit ATC market J5AE.[20]

This is not the only substitution that can take place. It is possible that physicians can engage in inter-market substitution and switch consumers to another drug in a different 4-digit ATC market but still within the same 3-digit ATC market. Again, in the above example, this would be a switch from Invirase® (4-digit ATC market J5AE) to Retrovir® (4-digit ATC market J5AF) which are both in 3-digit ATC market J5A. In Model 3 we find a 4.7 percent decline in sales for drugs within a 3-digit ATC market that experienced a relabel. Critically, the result in Model 3 implies that after controlling for competitor actions

[20] In Appendix Tables 1 and 2 we test alternative treatment periods. First, in Appendix Table 1 we consider time periods of three (Model 2) and four years (Model 3) before and after a drug relabeling. Our base model (Model 1, Table 3) is included as Model 1 for comparative purposes. Second, in Appendix Table 2 we widen the treatment window around the actual drug relabel. As a reminder, our baseline model excludes the quarter when a relabeling event occurred. In Model 1 and Model 2 we increase that exclusion to one and two quarters, respectively, before the quarter of relabel. This increase in exclusion will help if information leaks prior to announcement. All of the robustness results are consistent with our main findings in Table 3.

and capturing intra- and inter-market substitution patterns aggregate demand still declined by 4.7 percent. This decline plausibly represents consumers that fall out of the market.

It is important to recall the process that is involved with these types of substitutions. Only a physician can switch a consumer to another drug. While we can detect *ex post* that a substitution has occurred, we do not know what precipitated the move.[21] There are several possibilities. First, consumers could become informed of the relabel and push a physician to switch them. Second, physicians could independently learn about the relabel and decide to proactively switch a consumer. Third, physicians could learn about the relabel through detailing, either by the affected company or by a competitor and then decide to switch a consumer to another drug. These three are not mutually exclusive and there is some evidence to support the third explanation (Macher and Wade, 2016).

Given that our data is at the standard unit level we do not know exactly how many consumers this represents because prescription patterns will differ across drugs and consumers. We can, however, calculate a conservative, lower bound if we assume that the loss was for chronic conditions that require daily uptake. Under this assumption, we can multiply the decline in aggregate demand from Model 3 by average sales over the two-year sample period prior to the relabeling event. This translates into an estimated decline of 7.97 million standard units or slightly over 265,000 30-day prescriptions. If all of these prescriptions were for chronic conditions then this translates into approximately 11,000 consumers that fall out of the market.[22] Again, this is likely to be a conservative, lower bound estimate because not every prescription is for a chronic condition requiring a daily dose. As the number of prescriptions for acute conditions increase so would the number of consumers that fall out of the market.

These results have interesting welfare implications but even with our data, the best we can do is conjecture about them. First, if consumers that should be medicated move from the treated to untreated population this would be a detriment to consumer welfare. As suggested above, prospect theory provides one explanation as to why consumers may fall out of the market. These consumers, when confronted with new information about their current drug, may vastly overestimate the probability of a negative event within their own weighting function. They may also incorrectly attribute these same negative effects to the substitute drugs that physicians prescribe. As a result, they conclude the benefit does not outweigh the risk and exit the market.[23] Second, if the consumers that leave the market were those that were only marginally benefitting from a drug or maybe should not have been prescribed a drug to begin with, then

[21] This would require data on why physicians switched or changed a prescription. It would also require us to have prescription level data as opposed to what we have at the standard-unit level.

[22] Average quarterly sales (21.2 million) x 4.7% = 0.99 million standard units x 8 quarters = 7.97 million standard units. Next, 7.97 million divided by 30 = 265,707 30-day prescriptions. Finally, 265,707 divided by 24 months = 11,071 chronic patients.

[23] Consumers could leave the market for financial reasons too. For example, the new drug that a physician wants to switch to may be too expensive or not covered by insurance.

the negative impact on welfare may be muted. For example, ample evidence exists that suggests many medications are overprescribed (*e.g.*, Lembke *et al*., 2018; Sacarny *et al*., 2016; Forgacs, 2008; Price *et al*., 1986). Unfortunately, no data exists to determine which type of consumer leaves the market and why. As such, the overall impact on welfare is most likely bounded between these two extremes.

5.2 Heterogeneous impacts across relabeling intensity

Relabeling intensity varies across therapeutic markets (see Appendix Table 3). In Tables 4 and 5 we explore how these differential intensities impact aggregate demand. We divide our data into two sub-samples and define 'low-intensity markets' and 'high-intensity markets'.[24] In Table 4, low-intensity markets are defined as those 4-digit ATC markets where there was only one relabeling event over our sample period. In contrast, in Table 5, we define high-intensity markets as those 4-digit ATC markets where more than one relabeling event occurred over the sample period. In Table 4, Model 1 the coefficient on the interaction term (*Relabel * U.S.*) is negative and statistically significant at the one percent level. We find a decline of 10.8 percent in aggregate demand for focal drugs in these low-intensity markets. Interestingly, however, in Model 2 and Model 3 the interaction is not statistically significant. This suggests that intra-market substitution absorbed the decline in aggregate focal drug demand. In other words, in these markets physicians were successfully able to switch consumers to another drug *within* that same 4-digit ATC market. To the extent that consumer or physician concerns are warranted due to a relabeling event, this is the expected outcome.

In high-intensity markets, on the other hand, results are more complex. Across all models in Table 5 the interaction term is negative and statistically significant. In Model 1 aggregate focal drug demand declined by 18.9 percent decline while in Model 2 aggregate demand declined by 6.0 percent for drugs within a focal drug's 4-digit ATC market. As before Model 2 represents intra-market substitution or consumers being switched to other drugs *within* the same 4-digit ATC market. Shifting to the 3-digit ATC market that incorporates inter-market substitution patterns, Model 3, aggregate demand declined by 5.0 percent. Critically, this 5.0 percent decline in aggregate demand represents consumers that fall out of the market.

In Appendix Tables 4 and 5 we redefine low-intensity and high-intensity markets as those markets in the bottom and top quartile of relabeling activity.[25] Results remain robust with those reported in Tables 4 and 5. In low-intensity markets, Appendix Table 4, Model 1 aggregate focal demand declined by 10.3 percent. The interaction was not significant in Model 2 nor Model 3 again suggesting that intra-market substitution absorbed the entire decline. For the high-intensity markets, Appendix Table 5, Model 1 aggregate focal drug demand declined by 20.1 percent. In Model 2, which incorporates intra-market

[24] At the 4-digit ATC market-level there are 61 markets categorized as low-intensity and 76 as high-intensity.
[25] At the 4-digit ATC market-level there are 35 markets in the bottom quartile and 36 markets in the top quartile.

substitution patterns, aggregate demand declined by 13.0 percent. Finally, in Model 3 that incorporates inter-market substitution, aggregate demand declined by 8.3 percent. Again, this 8.3 percent decline in aggregate demand represents consumers that fall out of the market.

5.3. Heterogeneous impacts across levels of relabeling severity

As discussed in Section 2.0 the severity of drug relabeling spans from precaution (least serious) through box warnings (most serious). Table 6 explores whether the aggregate demand response we document varies across this continuum of severity. We split the data into three sub-samples representing precaution (Model 1), adverse reaction (Model 2) and warning/box warning (Model 3). The categorization continues to be based on the first time a drug is relabeled and allows us to isolate out the effects of any potential prior relabeling activity. Drugs that have multiple types of relabeling are counted individually in each category.[26] Across all models the interaction remains negative and statistically significant. As expected, we see an increasingly negative aggregate demand response as severity increases; aggregate demand declines by 15.6 percent, 20.3 percent and 36.3 percent in Models 1, 2 and 3, respectively.

The increasing decline in aggregate demand as severity increases should not be surprising; physicians appear to be switching consumers to other drugs as new potential risks reveal themselves. Notwithstanding the general decline, the results in Model 1 are unexpected. This appears to be a rather strong aggregate demand response given the limited severity of the relabel. If the response is medically warranted or if physicians believe there may be future problems with a relabeled drug, then we should see intra-market substitution absorb this decline.[27] We examine this in Table 7 where we split the sample and combine the two least severe relabeling events (*i.e.*, precaution and adverse reaction) together. Again, across the models we find a negative and statistically significant coefficient on our interaction of interest. At the focal drug level, Model 1, aggregate demand declined by 14.7 percent while at the 4-digit ATC market level, which incorporates intra-market substitution, aggregate demand declined by 5.1 percent. At the 3-digit ATC market level, Model 3, which accounts for inter-market substitution aggregate demand declined by 4.0 percent and represents the consumers that leave the market.

In Table 6 Model 3, aggregate demand declined by 36.3 percent for drugs that received either a warning or box warning. This response should not be surprising given the severity of the relabeling event. In Table 8, we combine warnings and box warnings and examine their intra- and inter-market substitution patterns. Across all three models in Table 8 our coefficient on the interaction term is negative and

[26] For example, if a relabeling event included both a precaution and an adverse reaction it would be included both as a precaution and adverse reaction individually.

[27] The average probability that a drug that has received a precaution receives another relabel is 72.2%. As such, physicians may be pre-emptively switching patients to another drug. However, in this case we should see the entirety of aggregate demand decline of a focal drug absorbed by intra-market substitution.

statistically significant. At the 4-digit ATC market level that incorporates intra-market substitution patterns (Model 2), aggregate demand declined by 10.0 percent. At the 3-digit ATC market level that accounts for inter-market substitution patterns (Model 3), aggregate demand declined by 8.3 percent and again represents consumers that leave the market. As the severity of the relabeling event increases (Table 7, Model 3 versus Table 8, Model 3) the percentage of consumers that choose to leave the market increases as well.[28] Importantly, given the substitution patterns captured within Model 3, consumers appear to viewing potential substitutes in the same negative manner as the focal drug.

Finally, we combine the intensity levels of relabeling activity from the prior section and examine how it impacts the heterogeneity of relabeling severity that we considered in this section. In Appendix Tables 8 and 9 we replicate Tables 7 and 8 for low-intensity markets. Results are consistent with our prior findings (Table 4 and Appendix Table 4). In Appendix Tables 8 and 9 we see declines in aggregate focal demand (Model 1) of 6.6 and 45.0 percent, respectively. Results in Models 2 and 3 are not statistically significant, suggesting that the entire decline in aggregate focal drug demand was absorbed by intra-market substitution.

In Appendix Tables 10 and 11 we replicate Tables 7 and 8 for high-intensity markets. Again, results are consistent with our prior findings for high-intensity markets (Table 5 and Appendix Table 5). For relabeling events that involved precaution or adverse warnings in high intensity markets, aggregate demand declined by 17.3 percent (Appendix Table 10, Model 1). At the 4-digit ATC market (Model 2) that incorporates intra-market substitution patterns, aggregate demand declined by 5.9 percent. Finally, at the 3-digit ATC market level (Model 3) that incorporates inter-market substitution patterns, aggregate demand declined by 4.8 percent. This again represents consumers that leave the market. The most significant declines are in high-intensity markets with warnings or box warnings (Appendix Table 11). Aggregate demand declined by 34.3 percent at the focal drug level (Model 1), 10.4 percent at the 4-digit ATC market level (Model 2), and 15.8 percent at the 3-digit ATC market level (Model 3). Unlike low-intensity markets where intra-market substitution absorbed the decline in aggregate focal drug demand, in high-intensity markets we see significant movement by consumers out of the market.

6.0 Robustness

6.1. Variation across market concentration and market size

It may be possible that variation in market size or the level of competition within markets may differentially influence physician prescribing behavior or consumer behavior. For example, business or

[28] In Appendix Tables 6 and 7 we consider alternative time periods. First, in Appendix Table 6 we consider three and four years before and after a relabeling event (as opposed to two years in our baseline model). Second, our baseline model excludes the quarter in which a relabeling event occurred. In Appendix Table 7 we exclude one and two quarters prior to the relabeling event (along with the quarter of the event). In both tables and across all models our results remain robust to our baseline findings.

general news stories may enhance physician or consumer awareness about a drug. We examine these issues in Appendix Table 12. In Models 1 and 2 we separate markets into the bottom and top quartiles of sales while in Models 3 and 4 we create a HHI index and separate markets into the bottom and top quartiles, respectively. Across all models we find a negative and significant coefficient on our interaction term. Aggregate demand declined by 9.5 percent and 19.8 percent in the bottom and top sales quartiles (Models 1 and 2), respectively. However, when we consider the bottom and top quartiles of HHI, the difference becomes negligible. In Models 3 and 4, aggregate demand declined by 22.8 percent and 21.3 percent, respectively. Thus, we appear to see some variation in response across market sizes but not across levels of competition.

6.2 Cross-molecular substitution

Drugs vary in their level of cross-molecular substitution (Branstetter *et al*., 2016 and 2014) or rates at which one drug can be substituted for another.[29] Thus we examine two classes of drugs that, according to our discussions with physicians and prior research, should exhibit significantly different cross-molecular substitution. The first market we will consider is ATC N (nervous system), which is comprised of seven 2-digit ATC therapeutic markets: anesthetics (N01), analgesics (N02), antiepileptics (N03), anti-Parkinson (N04), psycholeptics (N05), psychoanaleptics (N06) and other nervous system drugs (N07). Within these 2-digit ATC markets we have additional 3-digit and 4-digit ATC markets. For example, within N06 resides anti-depressants (N06A) and anti-dementia (N06D) drugs. In general, ATC N should exhibit lowers of cross-molecular substitution (Branstetter *et* al., 2016). In Appendix Table 13 we find a decline in aggregate focal drug demand of 21.4 percent (Model 1), however, the coefficient of interest is not significant in Model 2 or Model 3. These markets experience greater declines in aggregate demand, in percentage terms, than we saw for the overall sample, however, the entire decline is absorbed by intra-market substitution. That is, physicians successfully switch consumers to other drugs within the same 4-digit ATC market.

The second market that we consider is ATC J (anti-infectives), which is comprised of six 2-digit ATC markets: anti-bacterials (J01), anti-mycotics (J02), anti-mycobaterials (J04), anti-virals (J05), immune sera and immunoglobulins (J06), and vaccines (J07). The 2-digit ATC market J01 includes 10 different 3-digit ATC markets comprising various classes of anti-bacterials; for example, tetracyclines (J01A) and beta-lactam anti-bacterials/penicillins (J01C). In general, these ATC markets should exhibit greater rates of CMS than ATC N. In Appendix Table 14 we find a focal decline in aggregate demand of 24.2 percent (Model 1). In these markets, however, we also see declines of 13.8 percent and 13.5 percent

[29] Note that we are not referring to generic substitution of a branded product; for example, substituting Generic Drug A for Branded Drug A. These are natural substitutions that will take place based on availability of generic drugs and state substitution laws. With CMS, as discussed in Footnote 18, we are referring to the substation from Branded Drug A to Branded Drug B or from Branded Drug A to Generic Drug B.

in the 4-digit (Model 2) and 3-digit (Model 3) ATC markets, respectively. Thus, in markets where we observe greater rates of CMS, we also observe significant numbers of consumers leaving the market.

While we explore only two markets that represent varying levels of CMS we do find interesting results that are suggestive that the degree to which drugs are substituted ultimately impacts consumer response and their choice to leave the market. Generally, drugs within ATC N tend to be for chronic conditions as opposed to drugs within ATC J that tend to be for more acute conditions. Thus, consumers may be more willing to drop out of less-chronic markets. Interestingly, the rate at which consumers leave the market appears to be the opposite of what scholars have found with respect to drug adherence rates (*e.g.*, Haynes *et al*., 2002; Jackevicius *et al*., 2002; Cramer *et al*., 2003). In one study using data from Kaiser, primary non-adherence for anti-infectives was found to be 2.9 percent (Shin *et al*., 2012). Our combined results appear to suggest that for acute conditions primary non-adherence is lower but when confronted with negative information consumers leave the market at a higher rate than those suffering chronic conditions. In chronic markets, primary non-adherence is greater but those consumers are less likely to completely abandon the market. Overall, our results suggest that physicians, firms and policy makers need to worry about not only non-adherence but also how consumers respond to negative news about drugs.

7.0. Discussion and conclusions

We are not the first to analyze the impacts of drug relabeling in the U.S. However, unlike prior studies we are able to do so in a more comprehensive and causal manner. Given the breadth of our data we are able to incorporate all plausible intra- and inter-market substitution patterns across a larger number of markets along with competitor actions. This allows us to estimate not only the causal impact of a relabeling event on a focal drug but also quantify consumers that ultimately fall out of the market. In our baseline regressions (Table 3) we find a decline in aggregate focal drug demand of 16.9 percent at the first instance of relabeling. As Figure 1 demonstrates there is a significant decline in aggregate demand due to relabeling. Critically, after accounting for intra- and inter-market substitution patterns and potential competitor reactions we find a decline in aggregate demand of 4.7 percent that represents consumers that leave the market (Table 3, Model 3). As a lower bound, this translates into approximately 11,000 consumers with chronic conditions. We believe this is the first evidence to suggest that drug relabeling is causing consumers to leave the market.

These baseline results have potential welfare implications as consumers shift from the treated to untreated population. On the one hand, if these are consumers that should be treated but are no longer treated this will be a detriment to welfare. Adding to this negative effect will be those consumers that may have been switched to another drug that turns out to be less effective. On the other hand, there is ample evidence to suggest that some types of drugs are overprescribed so if the consumers that leave the market

should never have been treated in the first place then these negative welfare impacts will be dampened. These declines, however, would have to be matched against potential gains from consumers that didn't suffer a negative medical consequence because they were switched as a result of a relabeling event. Ultimately, we are unable to make a definitive statement about overall welfare. Nonetheless, given the significant aggregate demand responses that we have documented, further exploration is warranted.

In addition to our baseline results, we find increasing impacts across all levels of relabeling severity (Table 6). Consistent with prior literature (*e.g.*, Dorsey *et al*., 2010) we find the greatest impact for the most severe type of relabel. Less intuitive, however, is why we see such a significant demand response for the least severe relabel (*i.e.*, precaution). Conditional on receiving a precaution, there is a significant probability that a drug will be relabeled again in the future. So it is plausible that consumers or physicians are preemptively switching to other drugs. Given the particularities of the drug market in the U.S., consumers can only switch to different drugs if a physician changes their prescription.[30] After accounting for intra- and inter-market substitution (Table 7) we still find a 4.0 percent decline in aggregate demand, which represents consumers falling out of the market. As we conjectured in the paper, prospect theory provides one explanation for why this may be occurring; consumers may be over-estimating future negative events and projecting those negative events onto substitute products. As a result, in their minds the expected costs of taking substitute products are not worth the expected benefits.

We exploit other variation in our data. For example, we break markets into "low-intensity" and "high-intensity" markets based on the level of relabeling activity within a particular 4-digit ATC market. In the case of low-intensity markets (Table 4 and Appendix Table 4) and low-intensity markets across types of relabeling (Appendix Tables 8 and 9), we find that the entire decline in aggregate focal demand was absorbed by intra-market substitution. That is patients were all successfully switched to other drugs within the same 4-digit ATC market. In contrast, in the case of high-intensity markets (Table 5 and Appendix Table 5) and high-intensity markets across types of relabeling (Appendix Tables 9 and 10) we find not only declines in aggregate focal drug demand but also find that consumers leave the market. This split is an important caveat to prior work, especially the work focused on box warnings (*e.g.*, Dorsey *et al*., 2010; Olfson *et al*., 2008; Jacoby *et al*, 2005) because it suggests the impacts are more nuanced. Ultimately, the impacts of these kinds of warnings depend on the type of market in which they occurred.

Our results are robust to alternative treatment windows and to the exclusion of additional quarters prior to treatment. In addition to being sensitive to relabeling intensity, our results also vary across different levels of CMS. In markets were we expect low levels of cross-molecular substitution we find strong declines in aggregate focal drug demand (Appendix Table 13, Model 1). The entire decline,

[30] Here we are not referring to state substitution laws that allow for the shift from Brand A to Generic A without physician approval. Instead, we are talking about the movement from Brand A to either Brand B or Generic B.

however, is absorbed by intra-market substitution. That is, on average, consumers are switched to other drugs within the same 4-digit ATC market. We see the opposite in markets with higher levels of CMS. In that market (Appendix Table 14) we again see similar declines in aggregate focal drug demand. However, in this market we still see aggregate demand decline by 13.5 percent after accounting for intra- and inter-market substitution patterns. This result can be viewed as consumers that leave the market. This finding complements existing literatures on drug adherence. Low CMS markets tend to be chronic conditions with higher levels of non-adherence while higher CMS markets tend to be acute conditions with lower levels of non-adherence. Thus, in low CMS markets physicians need to worry about adherence while in high CMS markets they need to worry about consumers leaving the market if exposed to negative product news.

A significant body of work has focused on elasticity and brand loyalty within the pharmaceutical industry. These issues are critical, for example, for pricing strategies and how firms respond to competitors and structure end of life strategies of branded products. Our findings suggest that firms should also be concerned with the magnitude of consumer (and physicians) response to adverse news from relabeling events. While some of these shifts may be medically warranted, others may be due to competitor behavior (Macher and Wade, 2016), physicians responding defensively, consumers acting irrationally or some combination of these. Given that we control for focal firm detailing activity our findings suggest that this may not be enough to stem the decline in aggregate demand.

Our work is not without limitations and we are left with important questions whose answers will allow policymakers to reassess how negative product information is conveyed to consumers, physicians and the market. First, we need to understand why physicians are switching patients and whether those reasons vary across variation in relabeling severity. Are physicians practicing defensive medicine and shifting patients out of potential liability concerns? Are consumers requesting to be switched? Or, are physicians succumbing to promotion influences by competitor firms (*e.g.*, Macher and Wade, 2016)? Second, we don't know if consumers that get switched are moved to 'inferior' treatments for their respective condition. Not all drugs are perfect substitutes so to the extent that there is variation in the medical level of substitution this could have welfare implications for those consumers. Third, with our data we don't know which types of consumers leave the market. Were these consumers that should be treated? Or were they only marginally benefiting? Finally, if there is a behavioral component to consumer response, can anything be done by firms or on the policy front to 'nudge' them back into the market? We leave these questions for future work.

References

Bartel, A. P., Thomas, L. G., 1985. Direct and indirect effects of regulation: A new look at OSHA's impact. Journal of Law and Economics, 28(1): 1-26.

Bartel, A. P., Thomas, L. G., 1987. Predation through regulation: The wage and profit effects of the occupational safety and health administration and the environmental protection agency. Journal of Law and Economics, 30(2): 239-264.

Branstetter, L., Chatterjee, C., Higgins, M.J., 2014. Starving (or fattening) the golden goose: Generic entry and incentives for early-stage pharmaceutical innovation. NBER working paper 20532.

Branstetter, L., Chatterjee, C., Higgins, M.J., 2016. Regulation and welfare: Evidence from Paragraph-IV generic entry in the pharmaceutical industry. NBER working paper 17188. RAND Journal of Economics 47(4), 857-890.

Briesacher, B. A., Soumerai, S. B., Zhang, F., Toh, S., Andrade, S. E., Wagner, J. L., Shoaibi, A., Gurwitz, J. H., 2013. A critical review of methods to evaluate the impact of FDA regulatory actions. Pharmacoepidemiology and Drug Safety, 22(9): 986-994.

Brown, G. W., Lu, J. Y., Wolfson, R. J., 1964. Dynamic modeling of inventories subject to obsolescence. Management Science, 11(1): 51-63.

Bunniran, S., McCaffrey, D. J., Bentley, J. P., Bouldin, A. S., 2009. Pharmaceutical product withdrawal: Attributions of blame and its impact on trust. Research in Social and Administrative Pharmacy, 5(3): 262-273.

Cheah, E. T., Chan, W. L., Chieng, C. L. L., 2007. The corporate social responsibility of pharmaceutical product recalls: An empirical examination of US and UK markets. Journal of Business Ethics, 76(4): 427-449.

Cramer, J., Rosenheck, R., Kirk, G., Krol, W., Krystal, J., 2003. Medication compliance feedback and monitoring in a clinical trial: Predictors and outcomes. Value in Health, 6(5): 566-573.

Datta, A., Dave, D., 2017. Effects of physician□directed pharmaceutical promotion on prescription behaviors: Longitudinal evidence. Health Economics, 26(4): 450-468.

Dorsey, E. R., Rabbani, A., Gallagher, S. A., Conti, R. M., Alexander, G. C., 2010. Impact of FDA black box advisory on antipsychotic medication use. Archives of Internal Medicine, 170(1): 96-103.

Dranove, D., 2011. Health care markets, regulators, and certifiers. *Handbook of Health Economics*, 2: 639-690. Elsevier.

Dranove, D., Olsen, C., 1994. The economic side effects of dangerous drug announcements. The Journal of Law and Economics, 37(2): 323-348.

Dusetzina, S. B., Higashi, A. S., Dorsey, E. R., Conti, R., Huskamp, H. A., Zhu, S., Garfield, C. F., Alexander, G. C., 2012. Impact of FDA drug risk communications on health care utilization and health behaviors: A systematic review. Medical Care, 50(6): 466-478.

Forgacs, I., Loganayagam, A., 2008. Overprescribing proton pump inhibitors. British Medical Journal 336: 2.

Freedman, S., Kearney, M., Lederman, M., 2012. Product recalls, imperfect information, and spillover effects: Lessons from the consumer response to the 2007 toy recalls. Review of Economics and Statistics, 94(2): 499-516.

Gruenspecht, H. K., Lave, L. B., 2006. The economics of health, safety, and environmental regulation. *Handbook of Industrial Organization*, 2(7): 1507-1550.

Hartley, R. F. 1994. *Management Mistakes & Successes*: John Wiley and Sons.

Haynes, R. B., McDonald, H. P., Garg, A. X., 2002. Helping patients follow prescribed treatment: Clinical applications. The Journal of the American Medical Association, 288(22): 2880-2883.

Jackevicius, C. A., Mamdani, M., Tu, J. V., 2002. Adherence with statin therapy in elderly patients with and without acute coronary syndromes. The Journal of the American Medical Association, 288(4): 462-467.

Jacoby, J. L., Fulton, J., Cesta, M., Heller, M., 2005. After the black box warning: Dramatic changes in ED use of droperidol. The American Journal of Emergency Medicine, 23(2): 196.

Kahneman, D., Tverksy, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica, 47: 263-291.

Kales, H. C., Zivin, K., Kim, H. M., Valenstein, M., Chiang, C., Ignacio, R. V., Ganoczy, D., Cunningham, F., Schneider, L. S., Blow, F. C., 2011. Trends in antipsychotic use in dementia 1999-2007. Archives of General Psychiatry, 68(2): 190-197.

Kessel, R. A., 1967. Economic effects of rederal Regulation of milk markets. Journal of Law and Economics, 10: 51-78.

Lembke, A., Papac, J., Humphreys, K., 2018. Our other prescription drug problem. New England Journal of Medicine, 378(8): 693-695.

Leone, R. P., 1995. Generalizing what is known about temporal aggregation and advertising carryover. Marketing Science, 14(3): G141-G150.

Lu, C. Y., Zhang, F., Lakoma, M. D., Madden, J. M., Rusinak, D., Penfold, R. B., Simon, G., Ahmedani, B. K., Clarke, G., Hunkeler, E. M., Waitzfelder, B., 2014. Changes in antidepressant use by young people and suicidal behavior after FDA warnings and media coverage: Quasi-experimental study. British Medical Journal, 348: 3596.

Macher, J., Wade, J., 2016. The 'black box' of strategy: Competitive responses to and performance responses to adverse regulatory events. https://ashecon.confex.com/ashecon/2016/webprogram/Paper4635.html

Manchanda, P., Honka, E., 2005. The effects and role of direct-to-physician marketing in the pharmaceutical industry: An integrative review. Yale Journal of Health Policy, Law and Ethics, 5(2): 785-822.

Migue, J. L., 1977. Controls versus subsidies in the economic theory of regulation. Journal of Law and Economics, 20(1): 213-221.

Moore, T. J., Cohen, M. R., Furberg, C. D., 2007. Serious adverse drug events reported to the Food and Drug Administration, 1998-2005. Archives of Internal Medicine, 167(16): 1752-1759.

Nelson, P., 1970. Information and consumer behavior. Journal of Political Economy, 78(2): 311-329.

Nevo, A., 2001. Measuring market power in the ready-to-eat cereal industry. Econometrica, 69(2): 307-342.

Oberholzer-Gee, F., Mitsunari, M., 2006. Information regulation: Do the victims of externalities pay attention? Journal of Regulatory Economics, 30(2): 141-158.

Olfson, M., Marcus, S. C., 2008. A case-control study of antidepressants and attempted suicide during early phase treatment of major depressive episodes. Journal of Clinical Psychiatry, 69(3): 425-432.

Olson, M. K., 2003. Pharmaceutical policy change and the safety of new drugs. Journal of Law and Economics, 45(S2): 615-642.

Peltzman, S., 1976. Toward a more general theory of economic regulation. Journal of Law and Economics, 19(2): 211-240.

Peltzman, S., 1987. The health effects of mandatory prescriptions. Journal of Law and Economics, 30(2): 207-238.

Price, D., Cooke, J., Singleton, S., Feely, M., 1986. Doctors' unawareness of the drugs their patients are taking: A major cause of overprescribing? British Medical Journal, 292(6513): 99-100.

Qureshi, Z., Seoane-Vazqquez, E., Rodriguez-Monguio, R., Stevenson, K., Szeinbach, S., 2011. Market withdrawl of new molecular entities approved in the United States from 1980 to 2009. Pharmacoepidemiology & Drug Safety, 20(7): 772-777.

Rizzo, J., 1999. Advertising and competition in the ethical pharmaceutical industry: The case of antihypertensive drugs. The Journal of Law and Economics, 42(1): 89-116.

Sacarny, A., Yokum, D., Finkelstein, A., Agrawal, S., 2016. Medicare letters to curb overprescribing of controlled substances had no detectable effect on providers. Health Affairs, 35(3): 471-479.

Shah, N. D., Montori, V. M., Krumholz, H. M., Tu, K., Alexander, G. C., Jackevicius, C. A., 2010. Responding to an FDA warning—geographic variation in the use of rosiglitazone. New England Journal of Medicine, 363(22): 2081-2084.

Shin, J., McCombs, J. S., Sanchez, R. J., Udall, M., Deminski, M. C., Cheetham, T. C., 2012. Primary nonadherence to medications in an integrated healthcare setting. The American Journal of Managed Care, 18(8): 426-434.

Sloan, F. A., Steinwald, B., 1980. Effects of regulation on hospital costs and input use. Journal of Law and Economics, 23(1): 81-109.

Smalley, W., Shatin, D., Wysowski, D. K., Gurwitz, J., Andrade, S. E., Goodman, M., Chan, K. A., Platt, R., Schech, S. D., Ray, W. A., 2000. Contraindicated use of cisapride: Impact of food and drug administration regulatory action. The Journal of the American Medical Association, 284(23): 3036-3039.

Stigler, G. J., 1971. Theory of economic regulation. Bell Journal of Economics, 2(1): 3-21.

Tekin, E., Markowitz, S., 2008. The relationship between suicidal behavior and productive activities of young adults. Southern Economic Journal: 300-331.

Ter-Martirosyan, A., Kwoka, J., 2010. Incentive regulation, service quality, and standards in U.S. electricity distribution. Journal of Regulatory Economics, 38(3): 258-273.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty, 5(4): 297-323.

Wilkinson, J. J., Force, R. W., Cady, P. S., 2004. Impact of safety warnings on drug utilization: Marketplace life span of cisapride and troglitazone. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 24(8): 978-986.

Windmeijer, F., de Laat, E., Douven, R., Mot, E., 2005. Pharmaceutical promotion and GP prescription behavior. Health Economics, 15(1): 5-18.

Wong, C., Siah, K., Lo, A., 2018. Estimation of clinical trial success rates and related parameters. *Forthcoming* Biostatistics: https://doi.org/10.1093/biostatistics/kxx069

Wysowski, D. K., Swartz, L., 2005. Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002: The importance of reporting suspected reactions. Archives of Internal Medicine, 165(12): 1363-1369.

Zeithaml, V. A., 1988. Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. Journal of Marketing, 52(3): 2-22.

Zhao, X., Li, Y., Flynn, B. B., 2013. The financial impact of product recall announcements in China. International Journal of Production Economics, 142(1): 115-123.

Figure 1. Focal drug demand in U.S. and U.K. surrounding relabel events. The figure shows the sales quantity of focal drugs in the U.S. (treated) and U.K. (control) before and after relabeling. The relabeling event is set at *t*=0 where time horizon is in quarters and labeled on the x-axis. Sales are shown over eight quarters before and after the quarter of relabeling. The left y-axis are sales of drugs in the U.S., the right y-axis are sales of the same drugs in the U.K. Drug sales are in standardized units (millions) as determined by IMS Health and natural logarithms are taken.

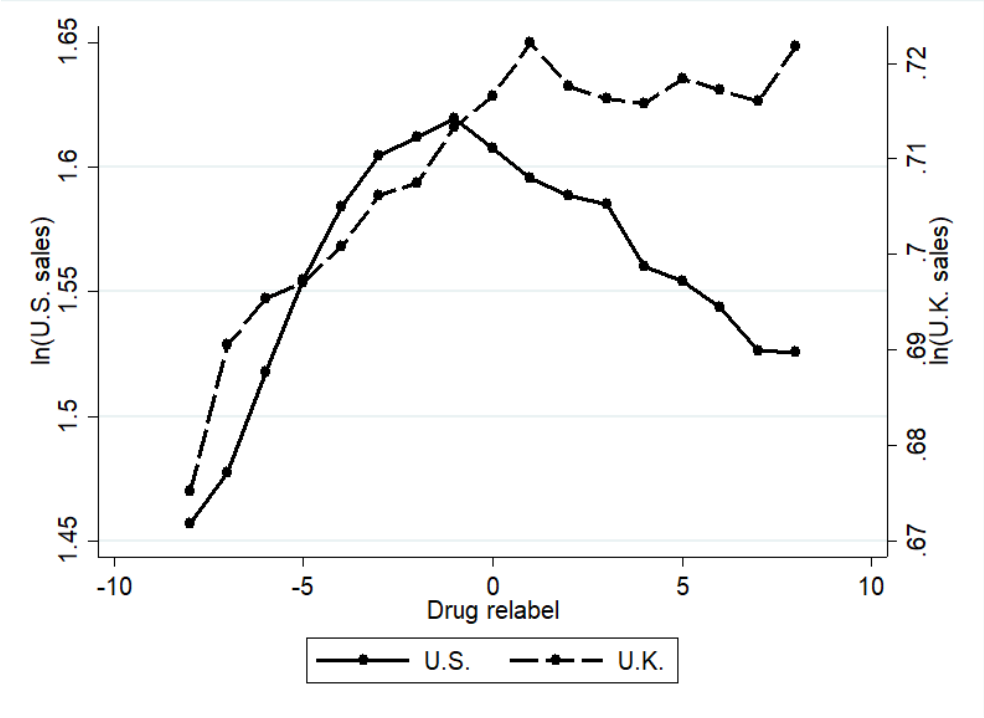

Table 1. Distribution of relabel activity between the U.S. and U.K. Our sample consists of drugs sold both in the U.S. and U.K. In order to create a clean control window we excluded drugs if they were relabeled in the U.K. within eight quarters of a U.S. relabel. This table shows the variation in relabeling types across the U.S. and U.K. for our sample. Within the imposed restrictions the average elapsed time between relabeling in the U.S. and U.K. is 12.95 quarters.

Relabeling Type	U.S.	U.K.	Average time (Quarters)
Precaution	226	166	13.48
Adverse Reaction	176	134	11.83
Warning	161	115	12.06
Box Warning	53	35	9.40
Label Changes	251	180	12.95

Table 2. Descriptive statistics. Sales (quantity) are measured in million of standardized units. IMS Health has converted financial variables for U.K. drugs to U.S. dollars. All financial variables have been converted to real 2009 U.S. dollars using a GDP deflator.

Variable	N	Mean	Median	Std. Dev.	Min	Max
U.S.	6,519	0.54	1.00	0.50	0.00	1.00
Sales (standard units)	6,519	13.87	0.88	45.81	0.00	577.85
Promotion	6,519	1.73	0.02	5.40	0.00	63.18
Lagged promotion stock	6,519	6.84	0.68	15.46	0.00	135.17
Price	6,519	91.78	2.36	357.43	0.01	5352.50
Relabel	6,519	0.27	0.00	0.45	0.00	1.00
Precaution	6,519	0.22	0.00	0.42	0.00	1.00
Adverse reaction	6,519	0.16	0.00	0.37	0.00	1.00
Warning	6,519	0.11	0.00	0.31	0.00	1.00
Box warning	6,519	0.03	0.00	0.17	0.00	1.00
Vintage	6,519	23.53	24.00	11.53	1.00	56.00
Number of brands	6,519	7.98	6.00	6.31	0.00	32.00
Number of generics	6,519	13.70	5.00	23.70	0.00	149.00

Table 3. Effects of relabeling on demand. Dependent variable is the natural logarithm of sales, ln(*Sales*). The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.108***	0.058**	0.023
	(0.032)	(0.024)	(0.020)
U.S.	0.712***	1.796***	1.375***
	(0.034)	(0.040)	(0.036)
Relabel * U.S.	-0.185***	-0.052*	-0.048**
	(0.025)	(0.028)	(0.023)
ln(Price)	-0.610***	-1.158***	-0.544***
	(0.053)	(0.049)	(0.049)
ln(Lagged promotion stock)	0.742***	0.186***	0.151***
	(0.015)	(0.009)	(0.007)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	6,519	5,946	4,946
Adjusted R^2	0.531	0.765	0.820
First stage F-statistic	37.12	64.79	26.56
Hansen J-statistic	2.12	0.15	2.621
Hansen J p-value	0.145	0.698	0.105
Marginal effects:			
Relabel * U.S.	**-0.169**	**-0.051**	**-0.047**

Table 4. Effects of relabeling on demand: Low-intensity markets. Dependent variable is the natural logarithm of sales, ln(*Sales*). Low-intensity markets are defined as those 4-digit ATC markets where there was only one relabeling event over our sample period. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.096*	-0.008	-0.071
	-0.052	-0.094	-0.089
U.S.	0.350***	1.640***	0.842***
	-0.056	-0.094	-0.085
Relabel * U.S.	-0.114***	-0.012	-0.015
	-0.04	-0.07	-0.064
ln(Price)	-0.522***	-1.073***	-0.445***
	-0.055	-0.046	-0.072
ln(Lagged promotion stock)	0.691***	0.153***	0.166***
	-0.034	-0.021	-0.019
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	1,576	1,576	749
Adjusted R^2	0.655	0.561	0.638
First stage F-statistic	79.67	150.21	16.34
Hansen J-statistic	1.235	0.477	0.092
Hansen J p-value	0.267	0.490	0.761
Marginal effects:			
Relabel * U.S.	**-0.108**		

Table 5. Effects of relabeling on demand: High-intensity markets. Dependent variable is the natural logarithm of sales, ln(*Sales*). High-intensity markets are defined as those 4-digit ATC markets where there was more than one relabeling event over our sample period. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.118***	0.072***	0.026
	-0.037	-0.022	-0.02
U.S.	0.904***	1.874***	1.539***
	-0.038	-0.034	-0.043
Relabel * U.S.	-0.210***	-0.062**	-0.051**
	-0.03	-0.028	-0.023
ln(Price)	-0.846***	-1.137***	-0.539***
	-0.099	-0.081	-0.076
ln(Lagged promotion stock)	0.706***	0.180***	0.129***
	-0.019	-0.008	-0.006
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	4,943	4,370	4,197
Adjusted R^2	0.526	0.833	0.852
First stage F-statistic	17.61	48.22	33.22
Hansen J-statistic	1.501	0.218	0.778
Hansen J p-value	0.221	0.640	0.378
Marginal effects:			
Relabel * U.S.	**-0.189**	**-0.060**	**-0.050**

Table 6. **Heterogeneous impacts across levels of relabeling severity.** Dependent variable is the natural logarithm of sales, ln(*Sales*). Data is split into three sub-samples representing precaution (Model 1), adverse reaction (Model 2) and warning/box warning (Model 3). The categorization is based on the first time a drug is relabeled and allows us to isolate out the effects of any potential prior relabeling activity. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

DV = ln(Sales)	Model 1 Precaution	Model 2 Adverse reaction	Model 3 Warning/Box
Relabel	0.111***	0.143***	0.176***
	(0.030)	(0.040)	(0.050)
U.S.	0.838***	0.749***	0.607***
	(0.039)	(0.041)	(0.051)
Relabel * U.S.	-0.170***	-0.227***	-0.451***
	(0.027)	(0.036)	(0.059)
ln(Price)	-0.759***	-0.528***	-0.709***
	(0.069)	(0.039)	(0.070)
ln(Lagged promotion stock)	0.725***	0.790***	0.756***
	(0.017)	(0.022)	(0.026)
Controls	Y	Y	Y
Drug fixed effect	Y	Y	Y
Market fixed effect	N	N	N
Time fixed effect	Y	Y	Y
N	5,183	3,166	2,236
Adjusted R^2	0.517	0.579	0.430
First stage F-statistic	29.39	65.76	37.76
Hansen J-statistic	0.81	5.821	1.451
Hansen J p-value	0.368	0.055	0.228
Marginal effects:			
Relabel * U.S.	**-0.156**	**-0.203**	**-0.363**

Table 7. Effects of precaution/adverse reaction relabeling on demand. Dependent variable is the natural logarithm of sales, ln(*Sales*). Sample includes the combination of precaution and adverse reaction. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.068**	0.067***	0.021
	(0.032)	(0.024)	(0.021)
U.S.	0.659***	1.779***	1.440***
	(0.039)	(0.042)	(0.033)
Relabel * U.S.	-0.159***	-0.052*	-0.041*
	(0.035)	(0.028)	(0.023)
ln(Price)	-0.569***	-1.175***	-0.700***
	(0.053)	(0.045)	(0.027)
ln(Lagged promotion stock)	0.808***	0.186***	0.151***
	(0.018)	(0.009)	(0.007)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	6,310	5,722	4,955
Adjusted R^2	0.407	0.768	0.812
First stage F-statistic	49.62	70.28	42.90
Hansen J-statistic	0.065	0.090	2.254
Hansen J p-value	0.799	0.765	0.133
Marginal effects:			
Relabel * U.S.	**-0.147**	**-0.051**	**-0.040**

Table 8. Effects of warning/box warning relabeling on demand. Dependent variable is the natural logarithm of sales, ln(*Sales*). Sample includes the combination of warning and box warning. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

DV = ln(Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.176***	0.020	0.005
	(0.050)	(0.034)	(0.029)
U.S.	0.607***	1.821***	1.545***
	(0.051)	(0.048)	(0.042)
Relabel * U.S.	-0.451***	-0.105**	-0.087**
	(0.059)	(0.044)	(0.036)
ln(Price)	-0.709***	-0.997***	-0.533***
	(0.070)	(0.059)	(0.028)
ln(Lagged promotion stock)	0.756***	0.174***	0.117***
	(0.026)	(0.012)	(0.012)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	2,236	2,189	1,991
Adjusted R^2	0.430	0.834	0.812
First stage F-statistic	37.76	96.64	753.48
Hansen J-statistic	1.451	0.824	1.653
Hansen J p-value	0.228	0.364	0.199
Marginal effects:			
Relabel * U.S.	**-0.363**	**-0.100**	**-0.083**

Appendix Table 1. Baseline results across alternative time periods. Dependent variable is the natural logarithm of sales, ln(*Sales*). Model 1 replicates Model 1, Table 3. The time period of analysis is extended to three years (12 quarters) in Model 2 and four years (16 quarters) in Model 3. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

DV = ln(Sales)	Model 1 2 years	Model 2 3 years	Model 3 4 years
Relabel	0.108***	0.113***	0.132***
	(0.032)	(0.027)	(0.025)
U.S.	0.712***	0.696***	0.711***
	(0.034)	(0.027)	(0.026)
Relabel * U.S.	-0.185***	-0.235***	-0.279***
	(0.025)	(0.022)	(0.021)
ln(Price)	-0.610***	-0.499***	-0.506***
	(0.053)	(0.033)	(0.036)
ln(Lagged promotion stock)	0.742***	0.760***	0.772***
	(0.015)	(0.013)	(0.012)
Controls	Y	Y	Y
Drug fixed effect	Y	Y	Y
Market fixed effect	N	N	N
Time fixed effect	Y	Y	Y
N	6,519	9,229	11,842
Adjusted R^2	0.497	0.498	0.484
First stage F-statistic	37.12	36.23	39.12
Hansen J-statistic	2.120	4.310	3.982
Hansen J p-value	0.145	0.635	0.679
Marginal effects:			
Relabel * U.S.	**-0.168**	**-0.209**	**-0.243**

Appendix Table 2. Baseline results across alternative treatment periods. Dependent variable is the natural logarithm of sales, ln(*Sales*). In our baseline model specification the quarter of relabel is excluded from analysis. Model 1 drops the quarter of relabel (t = 0) and the quarter prior (t = -1) Model 2 drops the quarter of relabel (t = 0) and the two quarters prior (t = -1, -2). Dropping prior quarters controls for any possible leakage of information. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 T = 0, -1	Model 2 T = 0, -1, -2
Relabel	0.118***	0.124***
	(0.036)	(0.041)
U.S.	0.671***	0.673***
	(0.032)	(0.034)
Relabel * U.S.	-0.203***	-0.210***
	(0.027)	(0.028)
ln(Price)	-0.472***	-0.470***
	(0.040)	(0.042)
ln(Lagged promotion stock)	0.739***	0.745***
	(0.016)	(0.016)
Controls	Y	Y
Drug fixed effect	Y	Y
Market fixed effect	N	N
Time fixed effect	Y	Y
N	6,074	5,661
Adjusted R^2	0.501	0.498
First stage F-statistic	48.86	44.81
Hansen J statistic	3.791	3.457
Hansen J p-value	0.150	0.178
Marginal effects:		
Relabel * U.S.	**-0.183**	**-0.189**

Appendix Table 3. Variation in relabeling activity. Count of first instance relabeling activity across type and over time. If a drug has multiple relabel types they are each counted below.

ATC	Relabel Type	2003	2004	2005	2006	2007	2008	2009
A	Precaution	6	13	17	19	24	26	26
	Adverse Reaction	4	7	12	17	21	26	29
	Warning	1	3	4	6	7	12	13
	Boxed Warning	0	0	0	0	3	5	5
	Total Relabel	6	14	19	22	28	31	33
B	Precaution	4	5	7	7	7	8	8
	Adverse Reaction	4	7	9	9	9	9	9
	Warning	2	3	6	6	7	7	7
	Boxed Warning	0	0	0	0	1	1	3
	Total Relabel	4	7	9	9	9	10	11
C	Precaution	10	15	20	26	30	36	39
	Adverse Reaction	11	15	19	22	22	24	26
	Warning	9	11	15	17	20	23	24
	Boxed Warning	0	0	1	2	3	4	4
	Total Relabel	16	23	27	31	34	40	42
D	Precaution	1	5	7	7	7	7	7
	Adverse Reaction	0	3	4	4	5	5	5
	Warning	0	0	1	3	3	3	3
	Boxed Warning	0	0	0	2	2	3	3
	Total Relabel	1	5	7	8	8	8	8
G	Precaution	5	7	14	15	16	17	17
	Adverse Reaction	2	4	10	11	11	12	13
	Warning	0	0	1	1	5	9	11
	Boxed Warning	0	0	0	0	0	0	0
	Total Relabel	5	7	14	15	17	18	18
H	Precaution	1	5	5	6	6	7	7
	Adverse Reaction	1	4	5	6	6	7	7
	Warning	0	2	2	3	3	4	4
	Boxed Warning	0	0	0	0	0	0	0
	Total Relabel	1	5	5	6	6	7	7
J	Precaution	18	44	50	56	59	64	68
	Adverse Reaction	12	30	40	45	46	53	57
	Warning	6	16	22	30	41	53	56
	Boxed Warning	2	8	9	10	13	15	16
	Total Relabel	22	49	55	59	62	69	72

Appendix Table 3. Variation in relabeling activity (continued).

ATC	Relabel Type	2003	2004	2005	2006	2007	2008	2009
L	Precaution	14	25	31	38	44	46	52
	Adverse Reaction	11	19	30	35	41	46	51
	Warning	7	10	17	29	38	41	47
	Boxed Warning	4	5	6	16	21	21	25
	Total Relabel	17	30	39	43	52	54	57
M	Precaution	8	17	20	23	24	24	24
	Adverse Reaction	5	11	15	15	17	18	19
	Warning	1	6	9	14	16	18	18
	Boxed Warning	0	1	3	10	11	11	12
	Total Relabel	10	19	21	23	24	24	24
N	Precaution	12	24	32	40	45	47	51
	Adverse Reaction	7	14	18	27	35	38	39
	Warning	7	11	22	30	38	40	44
	Boxed Warning	0	0	9	14	16	19	19
	Total Relabel	14	28	36	44	52	55	56
P	Precaution	1	1	2	2	2	2	2
	Adverse Reaction	1	1	1	1	1	1	1
	Warning	0	0	0	0	0	0	0
	Boxed Warning	0	0	0	0	0	0	0
	Total Relabel	1	1	2	2	2	2	2
R	Precaution	3	5	6	6	9	11	11
	Adverse Reaction	3	3	4	4	5	7	7
	Warning	2	3	3	3	4	5	6
	Boxed Warning	1	1	1	1	2	2	2
	Total Relabel	6	8	8	8	10	12	12
S	Precaution	6	8	9	11	11	11	11
	Adverse Reaction	0	1	3	4	4	4	4
	Warning	0	0	0	0	0	0	0
	Boxed Warning	0	0	0	0	0	0	0
	Total Relabel	6	8	10	12	12	12	12
T	Precaution	2	2	3	3	5	5	5
	Adverse Reaction	0	1	1	1	1	1	1
	Warning	0	0	0	1	2	2	3
	Boxed Warning	0	0	0	0	1	1	1
	Total Relabel	2	3	4	4	5	5	6

Appendix Table 3. Variation in relabeling activity (continued).

ATC	Relabel Type	2003	2004	2005	2006	2007	2008	2009
V	Precaution	1	2	2	3	3	4	5
	Adverse Reaction	1	2	2	2	3	3	3
	Warning	1	1	1	1	3	3	3
	Boxed Warning	0	0	0	0	0	0	0
	Total Relabel	1	2	2	3	4	4	5

ATC Therapeutic Code Definition:

A: Alimentary tract and metabolism
B: Blood and blood forming organs
C: Cardiovascular system
D: Dermatological
G: Genitourinary system and sex hormones
H: Systemic hormonal preparations, excluding sex hormones
J: Anti-infectives
L: Anti-neoplastic and immunomodulating agents
M: Musculoskeletal system
N: Nervous system
P: Anti-parasitic products
R: Respiratory system
S: Sensory organs
T: Diagnostic agents
V: Various

Appendix Table 4. Effects of relabeling on demand: Low-intensity markets (bottom quartile). Dependent variable is the natural logarithm of sales, ln(*Sales*). Low-intensity markets are defined as those 4-digit ATC markets in the bottom quartile of relabeling activity over our sample period. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.077	0.140**	0.007
	(0.054)	(0.059)	(0.051)
U.S.	0.386***	1.013***	1.092***
	(0.051)	(0.117)	(0.080)
Relabel * U.S.	-0.109***	-0.104	-0.018
	(0.041)	(0.067)	(0.054)
ln(Price)	-0.564***	-1.068***	-0.740***
	(0.129)	(0.097)	(0.047)
ln(Lagged promotion stock)	0.971***	0.321***	0.070***
	(0.024)	(0.028)	(0.016)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	1,264	1,202	1,170
Adjusted R^2	0.725	0.726	0.599
First stage F-statistic	12.81	38.02	81.74
Hansen J-statistic	0.558	0.017	0.590
Hansen J p-value	0.455	0.897	0.442
Marginal effects:			
Relabel * U.S.	**-0.103**		

Appendix Table 5. Effects of relabeling on demand: High-intensity markets (top quartile).
Dependent variable is the natural logarithm of sales, ln(*Sales*). High-intensity markets are defined as those 4-digit ATC markets in the top quartile of relabeling activity over our sample period. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.157***	0.172***	0.088***
	(0.051)	(0.063)	(0.028)
U.S.	0.812***	1.809***	1.574***
	(0.048)	(0.128)	(0.055)
Relabel * U.S.	-0.224***	-0.139**	-0.087***
	(0.043)	(0.064)	(0.033)
ln(Price)	-0.628***	-1.209***	-0.814***
	(0.051)	(0.056)	(0.028)
ln(Lagged promotion stock)	0.695***	0.259***	0.165***
	(0.036)	(0.026)	(0.012)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	1,903	1,669	1,533
Adjusted R^2	0.508	0.692	0.882
First stage F-statistic	100.28	147.89	1054.4
Hansen J-statistic	0.147	1.459	0.066
Hansen J p-value	0.701	0.227	0.798
Marginal effects:			
Relabel * U.S.	**-0.201**	**-0.130**	**-0.083**

Appendix Table 6. Heterogeneous relabeling severity across alternative time periods. Data split across types of relabeling activity and across alternative periods of analysis. Models 1-3 are extended to three years or 12 quarters before and after a relabeling event. Models 4-5 are extended to four years or 16 quarters before and after a relabeling event.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
		3 years			4 years	
DV = ln(Sales)	Precaution	Adverse Reaction	Warning/Box	Precaution	Adverse Reaction	Warning/Box
Relabel	0.075***	0.108***	0.036	0.083***	0.119***	0.063**
	(0.026)	(0.030)	(0.034)	(0.024)	(0.028)	(0.031)
U.S.	0.731***	0.770***	0.566***	0.708***	0.740***	0.741***
	(0.035)	(0.032)	(0.034)	(0.027)	(0.030)	(0.044)
Relabel * U.S.	-0.207***	-0.237***	-0.285***	-0.238***	-0.270***	-0.213***
	(0.023)	(0.027)	(0.041)	(0.021)	(0.026)	(0.039)
ln(Price)	-0.700***	-0.473***	-0.372***	-0.611***	-0.470***	-0.698***
	(0.063)	(0.039)	(0.045)	(0.043)	(0.039)	(0.072)
ln(Lagged promotion stock)	0.773***	0.720***	0.698***	0.795***	0.762***	0.703***
	(0.013)	(0.014)	(0.016)	(0.011)	(0.013)	(0.015)
Controls	Y	Y	Y	Y	Y	Y
Drug fixed effect	Y	Y	Y	Y	Y	Y
Market fixed effect	N	N	N	N	N	N
Time fixed effect	Y	Y	Y	Y	Y	Y
N	8,711	6,428	5,943	10,877	8,059	7,587
Adjusted R^2	0.487	0.558	0.433	0.505	0.549	0.418
First stage F-statistic	40.01	42.48	35.42	44.13	39.64	35.92
Hansen J-statistic	0.978	1.781	2.76	4.699	0.843	1.261
Hansen J p-value	0.323	0.619	0.430	0.454	0.974	0.261
Marginal effects:						
Relabel * U.S.	**-0.186**	**-0.211**	**-0.247**	**-0.211**	**-0.236**	**-0.192**

Appendix Table 7. Heterogeneous relabeling severity across alternative treatment periods. Data split across types of relabeling activity and across alternative exclusion windows. Models 1-3 exclude the quarter of a relabeling event and prior quarter from the analysis. Models 4-6 exclude the quarter of a relabeling event and two prior quarters.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
		T = 0, -1			T = 0, -1, -2	
DV = ln(Sales)	Precaution	Adverse Reaction	Warning/Box	Precaution	Adverse Reaction	Warning/Box
Relabel	0.124***	0.155***	0.179***	0.134***	0.166***	0.184***
	(0.038)	(0.050)	(0.054)	(0.043)	(0.056)	(0.056)
U.S.	0.847***	0.745***	0.468***	0.848***	0.758***	0.471***
	(0.042)	(0.044)	(0.047)	(0.044)	(0.047)	(0.048)
Relabel * U.S.	-0.191***	-0.255***	-0.456***	-0.203***	-0.273***	-0.455***
	(0.029)	(0.039)	(0.061)	(0.031)	(0.041)	(0.062)
ln(Price)	-0.758***	-0.515***	-0.408***	-0.749***	-0.510***	-0.398***
	(0.073)	(0.040)	(0.044)	(0.075)	(0.041)	(0.046)
ln(Lagged promotion stock)	0.725***	0.804***	0.738***	0.734***	0.807***	0.730***
	(0.018)	(0.024)	(0.026)	(0.018)	(0.025)	(0.027)
Controls	Y	Y	Y	Y	Y	Y
Drug fixed effect	Y	Y	Y	Y	Y	Y
Market fixed effect	N	N	N	N	N	N
Time fixed effect	Y	Y	Y	Y	Y	Y
N	4,552	2,786	1,975	4,240	2,599	1,839
Adjusted R^2	0.491	0.552	0.374	0.490	0.548	0.365
First stage F-statistic	48.17	55.94	36.8	21.06	51.52	33.89
Hansen J-statistic	0.934	2.399	2.316	0.57	1.641	3.03
Hansen J p-value	0.334	0.301	0.314	0.45	0.44	0.22
Marginal effects:						
Relabel * U.S.	**-0.173**	**-0.225**	**-0.366**	**-0.165**	**-0.238**	**-0.365**

Appendix Table 8. Effects of precaution/adverse selection relabeling: Low-intensity markets. Dependent variable is the natural logarithm of sales, ln(*Sales*). Sample includes the combination of precaution and adverse reaction along with low-intensity markets. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.075*	0.024	-0.063
	(0.039)	(0.088)	(0.086)
U.S.	0.333***	1.559***	0.829***
	(0.052)	(0.101)	(0.088)
Relabel * U.S.	-0.068*	-0.020	0.011
	(0.035)	(0.072)	(0.065)
ln(Price)	-0.287***	-1.044***	-0.584***
	(0.053)	(0.045)	(0.036)
ln(Lagged promotion stock)	0.804***	0.165***	0.178***
	(0.034)	(0.021)	(0.020)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	1,564	1,564	737
Adjusted R^2	0.622	0.558	0.638
First stage F-statistic	76.21	160.56	204.09
Hansen J-statistic	0.248	0.977	2.462
Hansen J p-value	0.618	0.323	0.117
Marginal effects:			
Relabel * U.S.	**-0.066**		

Appendix Table 9. Effects of warning/box warning relabeling: Low-intensity markets. Dependent variable is the natural logarithm of sales, ln(*Sales*). Sample includes the combination of warning and box warning along with low-intensity markets. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln (Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.228*	-0.251**	-0.007
	(0.117)	(0.098)	(0.118)
U.S.	0.174**	1.842***	1.051***
	(0.072)	(0.078)	(0.095)
Relabel * U.S.	-0.597***	0.163	0.021
	(0.148)	(0.113)	(0.120)
ln(Price)	-0.310***	-0.812***	-0.449***
	(0.077)	(0.048)	(0.039)
ln(Lagged promotion stock)	0.596***	0.183***	0.181***
	(0.049)	(0.022)	(0.029)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	654	762	430
Adjusted R^2	0.746	0.707	0.638
First stage F-statistic	18.28	410.61	580.23
Hansen J-statistic	0.708	1.67	1.276
Hansen J p-value	0.400	0.196	0.259
Marginal effects:			
Relabel * U.S.	**-0.450**		

Appendix Table 10. Effects of precaution/adverse selection relabeling: High-intensity markets. Dependent variable is the natural logarithm of sales, ln(*Sales*). Sample includes the combination of precaution and adverse reaction along with high-intensity markets. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.035	0.071***	0.019
	(0.041)	(0.023)	(0.021)
U.S.	0.773***	1.889***	1.630***
	(0.047)	(0.034)	(0.032)
Relabel * U.S.	-0.190***	-0.061**	-0.049**
	(0.045)	(0.027)	(0.024)
ln(Price)	-0.658***	-1.176***	-0.781***
	(0.071)	(0.072)	(0.033)
ln(Lagged promotion stock)	0.803***	0.174***	0.133***
	(0.022)	(0.007)	(0.007)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	4,746	4,158	4,218
Adjusted R^2	0.360	0.844	0.846
First stage F-statistic	31.25	54.03	405.85
Hansen J-statistic	0.011	0.218	1.095
Hansen J p-value	0.915	0.641	0.295
Marginal effects:			
Relabel * U.S.	**-0.173**	**-0.059**	**-0.048**

Appendix Table 11. Effects of warning/box warning relabeling: High-intensity markets. Dependent variable is the natural logarithm of sales, ln(*Sales*). Sample includes the combination of warning and box warning along with high-intensity markets. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln (Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.212**	0.093***	0.041
	(0.087)	(0.034)	(0.031)
U.S.	0.670***	1.778***	1.675***
	(0.070)	(0.052)	(0.041)
Relabel * U.S.	-0.420***	-0.110***	-0.172***
	(0.091)	(0.040)	(0.038)
ln(Price)	-0.891***	-0.581***	-0.594***
	(0.098)	(0.068)	(0.039)
ln(Lagged promotion stock)	0.791***	0.132***	0.114***
	(0.034)	(0.014)	(0.011)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	1,582	1,409	1,561
Adjusted R^2	0.425	0.902	0.862
First stage F-statistic	41.69	60.91	652.71
Hansen J-statistic	0.674	1.030	0.101
Hansen J p-value	0.412	0.310	0.751
Marginal effects:			
Relabel * U.S.	**-0.343**	**-0.104**	**-0.158**

Appendix Table 12. Effects of relabeling on demand across market size and concentration.
Dependent variable is the natural logarithm of sales, ln(*Sales*). Models 1-2 split the sample across the bottom and top quartile of sales within a 4-digit ATC market. Models 3-4 split the sample across the bottom and top quartile of market concentration or HHI within a 4-digit ATC market. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01

	Model 1	Model 2	Model 3	Model 4
	Bottom Quartile	Top Quartile	Bottom Quartile	Top Quartile
DV = ln(Sales)	Sales		HHI	
Relabel	0.055**	0.058	0.151*	0.175***
	(0.027)	(0.070)	(0.085)	(0.054)
U.S.	0.204***	1.116***	-0.029	0.890***
	(0.024)	(0.061)	(0.059)	(0.049)
Relabel * U.S.	-0.100***	-0.221***	-0.259***	-0.240***
	(0.025)	(0.055)	(0.063)	(0.043)
ln(Price)	-0.221***	-1.280***	0.173	-0.681***
	(0.036)	(0.125)	(0.119)	(0.046)
ln(Lagged promotion stock)	0.459***	0.721***	1.057***	0.685***
	(0.031)	(0.026)	(0.047)	(0.036)
Controls	Y	Y	Y	Y
Drug fixed effect	Y	Y	Y	Y
Market fixed effect	N	N	N	N
Time fixed effect	Y	Y	Y	Y
N	1,369	1,972	1,300	1,930
Adjusted R^2	0.241	0.504	0.332	0.431
First stage F-statistic	21.26	54.41	26.57	113.48
Hansen J-statistic	1.927	4.276	0.054	0.002
Hansen J p-value	0.165	0.233	0.817	0.961
Marginal effects:				
Relabel * U.S.	**-0.095**	**-0.198**	**-0.228**	**-0.213**

Appendix Table 13. Effects of relabeling on market ATC N. Dependent variable is the natural logarithm of sales, ln(*Sales*). The sample includes only those drugs within the 1-digit ATC market N or nervous system. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln (Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.100	0.017	-0.010
	(0.092)	(0.040)	(0.042)
U.S.	0.289***	1.117***	1.026***
	(0.109)	(0.067)	(0.057)
Relabel * U.S.	-0.241***	-0.022	0.039
	(0.074)	(0.053)	(0.054)
ln(Price)	-0.236	1.744***	1.372***
	(0.311)	(0.148)	(0.155)
ln(Lagged promotion stock)	0.961***	0.170***	0.188***
	(0.032)	(0.014)	(0.010)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	1,148	1,027	949
Adjusted R^2	0.602	0.852	0.859
First stage F-statistic	92.95	243.58	156.74
Hansen J-statistic	0.342	0.286	0.839
Hansen J p-value	0.559	0.593	0.360
Marginal effects:			
Relabel * U.S.	**-0.214**		

Appendix Table 14. Effects of relabeling on market ATC J. Dependent variable is the natural logarithm of sales, ln(*Sales*). The sample includes only those drugs within the 1-digit ATC market J or anti-infectives. The unit of analysis in Model 1 is the focal drug level, Model 2 is the 4-digit ATC market (ATC4) level and Model 3 is the 3-digit ATC market (ATC3) level. *Price* is instrumented in all models with relevant tests reported in the table. Controls include *Vintage*, *Number of brands*, and *Number of generics*. The models are log-linear, as such the marginal effects are calculated using the equation $\exp^{(\beta-1)}$ where β is the respective coefficient on our variable of interest, (*Relabel*U.S.*). Marginal effects for our variable of interest are reported in the lower panel. Standard errors are clustered at the 2-digit ATC market level. Constants are included in all specifications but omitted from the table. * p < 0.10, ** p < 0.05, *** p < 0.01.

DV = ln(Sales)	Model 1 Focal Drug	Model 2 ATC4 market	Model 3 ATC3 market
Relabel	0.174	0.292***	0.187***
	(0.128)	(0.076)	(0.040)
U.S.	1.549***	1.557***	1.888***
	(0.128)	(0.167)	(0.153)
Relabel * U.S.	-0.277***	-0.149**	-0.145***
	(0.103)	(0.072)	(0.044)
ln(Price)	-3.225***	-0.159	-1.216***
	(0.513)	(0.126)	(0.260)
ln(Lagged promotion stock)	0.541***	0.241***	0.137***
	(0.064)	(0.031)	(0.027)
Controls	Y	Y	Y
Drug fixed effect	Y	N	N
Market fixed effect	N	Y	Y
Time fixed effect	Y	Y	Y
N	1,126	1,108	1,022
Adjusted R^2	-0.107	0.781	0.882
First stage F-statistic	15.86	240.35	20.23
Hansen J-statistic	3.808	1.779	0.457
Hansen J p-value	0.149	0.182	0.499
Marginal effects:			
Relabel * U.S.	**-0.242**	**-0.138**	**-0.135**