

OFF-LABEL USE OF PHARMACEUTICALS: A DETECTION
CONTROLLED ESTIMATION APPROACH*

W. DAVID BRADFORD[†]
 JOHN L. TURNER[†]
 JONATHAN W. WILLIAMS[‡]

We identify the rate of off-label use of prescription drugs in the United States during 1993-2008 using Detection Controlled Estimation. We find that the rate rises from 29.9% to 38.3% during this period. Off-label prescribing increases when there are fewer FDA-approved alternatives and a patient's insurance has less restrictive formularies and lower copayments. The temporal increase in off-label use coincides with a surge in settlements of Department of Justice (DOJ) lawsuits for off-label marketing. Those drugs targeted by the DOJ have 4.6% higher rates of off-label use initially, but the rate decreases 10.2% after the company becomes aware of the suit. The welfare effects of such regulatory intervention are unclear because we find off-label prescribing patterns by physicians that are consistent with enhancement of patient welfare.

I. INTRODUCTION

SINCE 1962, THE U.S. FOOD AND DRUG ADMINISTRATION (FDA) has restricted the marketing of a drug to just the set of 'on label' indications for which the drug is approved. However, physicians may prescribe any approved drug for any indication.¹ In the market for pharmaceuticals, which accounted for \$321.3 billion in sales in the United States in 2010 (2.2% of U.S. GDP),² 'off label' use is common and potentially desirable. On one hand, the best treatment for a patient's particular indication may require

*For helpful comments, we thank Guy David, Paul Greenburg, John Rizzo, Matthew Rosenberg, Geoffrey Sanzenbacher, Tamar Sisitsky, Marta Wosinska, and participants at the UGA/GSU/Emory Health Economics Conference (2013), the Bates-White Life Sciences Symposium (2013), the ASHEcon Conference (2014), the AEA Meetings (2015), the Society for Economic Measurement Conference (2015), and the Innovation and Behavior in Health Markets Conference (2016). Andrey Minaev provided excellent research assistance. Remaining errors are our own.

[†]Authors' affiliations: University of Georgia, Athens, Georgia, U.S.A.
e-mail: bradfowd@uga.edu jltturner@uga.edu

[‡]University of North Carolina at Chapel Hill, Chapel Hill, North Carolina USA
e-mail: jonwms@unc.edu

¹For clarity, we use the term 'indication' instead of 'diagnosis,' 'condition,' 'disease,' etc., throughout the paper.

²Source: 2010 IMS National Sales Perspectives.

using a drug off-label. In addition, applying FDA-approved drugs to new uses may be a particularly cost-effective type of innovation because these drugs have already passed safety benchmarks in clinical trials. On the other hand, ineffective off-label use is socially wasteful. In a few cases, off-label use has also led to patients' being physically harmed (O'Reilly and Dalal [2003]).

Not surprisingly, off-label use is enormously controversial in the clinical and policy communities and among federal regulators (Salbu [1999]; Klein and Tabarrok [2004]; Stafford [2008]). Despite the implications for the industry, no prior research supports broad, systematic, and trend-based analysis of off-label use. Indeed, no economics papers analyze it empirically, and virtually nothing is known about its welfare consequences. Further, there is a lack of empirical evidence on the effect of regulatory interventions intended to deter off-label promotion, or guidance on the most-effective role for regulators with respect to off-label use. In this paper, we start to fill these gaps.

We apply Detection Controlled Estimation (Feinstein [1990]) to a comprehensive data set of patient prescriptions from the National Ambulatory Medical Care Survey (NAMCS) during 1993-2008 to identify the incidence of off-label use and to test for what drives it. We find that in the most recent years, more than one in three prescriptions is written for off-label uses. This rate is lowest in the earliest years, and rises from 29.9% in 1993 to 38.3% in 2008, a 28.1% increase. The three years with the highest frequency are 2006, 2007, and 2008. Our yearly estimates are higher than the estimate by Radley *et al.* [2006] of a 21% rate of off-label use in 2001. Notably, our study considers a sample representative of all prescribed drugs, whereas Radley *et al.* [2006] consider just 160 drugs comprising about 56% of prescriptions written in 2001.

In contrast to Radley *et al.* [2006], we also capture data that allow us to test for economic drivers of off-label use. Analyzing the effect of patient characteristics, we find patterns that are consistent with choices we expect rational, fully-informed patients to make. For example, after controlling for a host of drug and patient characteristics, we show that an increase in the number of drugs that have been approved to treat a patient's set of diagnoses leads to a reduction in the probability that a physician prescribes off-label. We also find that when patients face lower out-of-pocket costs, physicians tend to prescribe off-label more often and patients get a higher percentage of off-label drugs. Relative to patients with no insurance, patients with insurance (of any type) are more likely to be prescribed off-label. Of those with insurance, those with Medicaid are the most likely to be prescribed off-label. Specifically, relative to those with private insurance, the probability of a physician's prescribing off-label is about 2.3% higher when a patient is insured through Medicaid. This may be in response to the relatively weak prior authorization programs at the

state level for Medicaid since 1990 (Dranove [1989]; Huskamp [2003]) or the very low copayments (relative to privately insured patients) that Medicaid recipients pay. Those on Medicare are prescribed off-label at a similar rate to those on private insurance, which is consistent with the privatized nature of 'Medigap' policies that were in place over much of our sample time period (Oliver *et al.* [2004]; Rowland [2001]). Together, these substitution patterns are consistent with predictions from our simple theoretical model of drug choice, where off-label prescribing by physicians may enhance patient welfare. We also evaluate the relationship between actual off-label use and government scrutiny of alleged off-label marketing. Between 1996-2012, the U.S. Department of Justice (DOJ) sued numerous manufacturers for off-label marketing of drugs, and has publicly disclosed settlements totaling over \$16 billion. We first show that the targeted drugs (from these settlements) have 4.6% higher average rates of off-label use. This suggests that DOJ suspicion of off-label marketing activities is associated with higher (and otherwise unexplained) rates of off-label use. However, the effect that the regulatory intervention itself has on subsequent off-label use is theoretically ambiguous. While the suit and settlement should deter further promotion of off-label uses, the legal process itself often receives considerable attention in the news and may actually facilitate information acquisition by prescribing doctors on off-label uses. We find that, on average, the net effect is to reduce off-label prescribing by 10.2% from the level prior to the firm's being aware of the DOJ suit. The effect of the intervention is increasing in the time since the investigation began. This is consistent with lags in the effects of the elimination of alleged off-label promotional activities.

Studying off-label use retrospectively using prescription data presents two significant challenges. First, we must classify uses as on-label and off-label. Unfortunately, no existing archive tracks a drug's FDA-approved uses across time. We use annual issues of the *Physician's Desk Reference* (PDR) to build yearly matches between drugs and their approved (on-label) indications. To match the non-standardized indications in the PDR to the list of *International Classification of Disease - 9th Revision* (ICD-9) codes from NAMCS prescription records, we rely on a professional clinical documentation specialist employed at a major academic medical center hospital. We treat a drug as being on-label for an indication if it has the same active ingredient as one of the drugs identified in the PDR as being on-label for that indication.

Second, the way that NAMCS prescription data (and data in nearly every other survey or retrospective data set) are recorded almost guarantees false detection of off-label use. For one thing, indications recorded on survey forms are limited by the number of available fields. For another, physicians often base their reports on administrative claims that seek to maximize

reimbursement.³ Perhaps most importantly, patients frequently visit their physicians about one indication and receive a prescription for another. For example, suppose a person with chronic hypertension visits the physician because he has the flu. If the physician does not record hypertension as an indication on the NAMCS form, but does record a (convenience-driven) re-fill prescription for an ACE inhibitor, then naive inspection of the prescription record would classify the use of the ACE inhibitor as off-label. This issue is problematic in nearly every other clinical, administrative or retrospective data source.

To overcome these problems, we appeal to Detection Controlled Estimation (DCE), first used by Feinstein [1989, 1990, 1991] to study tax evasion and regulation of U.S. nuclear power plants. Intuitively, this procedure constructs a model that separately predicts the probability of on-label use and the probability of whether it is detected, with estimation done via maximum likelihood. Crucially, identification requires a subset of variables that affects only the probability of on-label use and a subset of variables that affects only the probability of detection. For the latter category, we rely on the plausible assumption that changes to the NAMCS survey form affects detection but do not affect a physician's decision to prescribe on-label. The form permits up to five prescriptions during 1993-94, up to six from 1995-2002 and up to eight from 2003-2008. Since the maximum number of prescriptions and indications is often binding, and an increase in the number of prescriptions relative to the number of indications tends to increase the likelihood that on-label use is not detected, these changes vary the rate of detection exogenously. In using DCE, a cautious interpretation of our results treats the estimates of off-label use as upper bounds. Because we get close to the limiting case in our data, we interpret our estimates as tight upper bounds.

More work is required to understand fully the efficiency and welfare properties of off-label prescribing. For example, our paper does not identify either the dollar costs or treatment gains of off-label uses relative to on-label alternatives. We view our results, which identify the overall rate with which the practice occurs and identify some determinants of physicians' decisions to prescribe off-label, as an important first step in understanding the phenomenon of off-label prescribing. A greater understanding of these issues can be helpful to identify the most-effective role for regulatory bodies in prescription drug markets.

³In a study providing advice to physicians on how to increase revenue through careful attention to reimbursement intricacies, Heidelbaugh *et al.* [2008] states 'Be sure to list active and acute medical indications discussed during the visit... rather than those that are stable...'. Again, this could cause a prescription to appear off-label in claims data, even if it were actually prescribed for an approved use.

The remainder of the paper is as follows. Section II provides some background on the practice of off-label use in the pharmaceutical industry, including a literature review. Section III presents a theoretical model of off-label prescribing to assist in the interpretation of our empirical findings. We discuss the data and empirical model in Sections IV and V respectively. Section VI discusses our results and Section VII concludes.

II. FDA OVERSIGHT AND THE NATURE OF OFF-LABEL USE

The Food and Drug Administration Act of 1906 created the FDA, initially just to set manufacturing standards. In the wake of the Elixir Sulfanilamide episode, Congress passed the Federal Food, Drug and Cosmetic Act of 1938.⁴ This prevents a new drug's introduction without FDA certification that the drug is safe. The law also led to the modern arrangement where by many drugs are available only with a physician's prescription (Temin [1979]).⁵

In 1962, Congress passed the Kefauver-Harris Amendments to the Food, Drug and Cosmetic Act. This grants the FDA the authority to certify a drug's efficacy, in addition to safety, before a firm can sell it. Initially, supporters of this law argued the law would prevent firms from aggressively marketing drug products, with dubious effectiveness, to physicians who would then write prescriptions for their patients. However, the Thalidomide episode of 1961-62 created a sense of urgency that also helped facilitate passage of the law (Harris [1964]). While the 1938 law kept Thalidomide off the U.S. market, some physicians had already received the drug for experimental purposes. The FDA did not heavily regulate this type of distribution. Reports of birth defects in babies born to European mothers who had taken Thalidomide raised concerns that pharmaceutical firms might harm patients by moving their products to market too quickly (Peltzman [1973]). This series of events contributed significantly to the emergence of the current regulatory environment, which has important implications for the incentives for firms to gain approval for particular indications.

II(i). *Drug Development*

Drug development begins with the isolation of a new molecule. A researcher then tests to determine whether the molecule is biologically active and to

⁴In the 1930's, the S.E. Massengill Company sold the antibiotic sulfanilamide first as tablets and capsules. They then developed a liquid version by dissolving sulfanilamide in diethylene glycol. In September, 1937, Massengill marketed this liquid as 'Elixir Sulfanilamide.' Unbeknownst to their chemists, diethylene glycol is toxic. Over 100 people died. See Temin [1979].

⁵As Temin [1979] discusses, the new law did not directly implement the prescription system, but the FDA moved quickly to use the law to establish this system.

identify the nature of that action. Once action is determined, usually in animal models, the researcher (typically backed by a pharmaceutical manufacturer) files an investigational new drug application (IND) to begin human trials. These clinical trials follow a strict three-phase process in which the applicant must prove safety and efficacy.⁶ If successful in these trials, the applicant submits a New Drug Application (NDA) to the FDA; if it is approved, the NDA allows the applicant to sell the drug in the U.S. and specifies a list of approved indications. Firms may market their product for the approved (i.e., on-label) indications only.

The process from IND to NDA is long, risky and costly. Typically, it takes a decade or more (DiMasi *et al.* [2003]), and only around 9% of drugs for which an IND is filed achieve an NDA (DiMasi [2001]). Because firms typically seek molecule patents at the moment of initial discovery, the process often consumes half or more of the life of the core patent covering the molecule. Given failure rates, direct costs of drug discovery and testing, and the opportunity cost of capital that must be devoted to the effort, estimated average costs of achieving an NDA range from \$800 million to \$1 billion (DiMasi *et al.* [2003]; Adams and Brantner [2006]).

Because of the incentives in the clinical-trials process, firms do not typically include a broad list of approved indications on their drugs' labels. Part of the reason is that clinical trials center on demonstrating efficacy for a given indication, and the results of one approval (NDA) are often not transferable to another. To add a new indication for an existing drug, a manufacturer must go through the same tedious, costly clinical-trials process needed to achieve the original marketing approval.⁷

The FDA grants a variety of marketing exclusivities, some of which affect strongly the incentives to pursue NDA's and new indications. First, a drug approved with a new molecule may earn a new chemical entity (NCE) exclusivity. For five years, the FDA prevents any other firm from marketing the drug, regardless of patent protection.

⁶In Phase I, healthy humans are given the compound to establish safety. Phase II tests involve a small number of volunteers with the disease for which the manufacturer is ultimately seeking approval. Phase II aims to establish some measurable treatment effect. Once a minimal efficacy standard is met, the study proceeds to Phase III, which requires that much larger pools of volunteers with the disease be tested to identify the treatment effect more precisely. All three phases are generally double blind.

⁷As the USGAO has described it: 'If, after FDA has approved a drug, evidence arises of its safety and effectiveness in treating conditions or patient groups other than those named on the label, then the drug's manufacturer (or any other interested party) can submit a new application to have the label changed. This application, known as an 'efficacy supplement,' is similar to the original application in that it must contain evidence demonstrating to FDA's satisfaction that the product is both safe and effective for the treatment of the new condition' (USGAO [1996]: page 2).

More importantly, the FDA also grants a three-year marketing exclusivity for a label change that recognizes a new indication (NI). In principle, the NI exclusivity grants a significant incentive to endure clinical trials to modify a drug's label. However, this benefit depends crucially on generic competition. If one or more generic manufacturers have gained approval to sell a drug, then the benefits of obtaining an NI may spill over to generic manufacturers. Intuitively, a firm that obtains an NI receives only a three-year exclusive right to *advertise* its product for the new use. But since a physician can prescribe a drug for any reason, if one version of a drug is advertised for a use, physicians may prescribe all other versions at nearly the same rate.

Prior to 1984, generic competition was less of a concern. Because generic applicants had to go through clinical trials to sell drugs bioequivalent to approved drugs, and because there was tremendous uncertainty about whether testing a drug infringed on patents, generic products were a relatively small part of the approved drug portfolio in the U.S. In 1984, however, Congress passed the Hatch-Waxman Act to clarify the rules for, and increase the rate of, generic entry.⁸ Now, a generic applicant may receive an Abbreviated New Drug Application (ANDA), permitting entry, by demonstrating that its product is bioequivalent to the branded product. The cost of an ANDA is a small fraction of that of an NDA. Generic entry has expanded so rapidly, that generics now account for nearly two-thirds of all prescriptions written in the U.S. (Aitken *et al.* [2009]).

Practically speaking, obtaining a new indication for a drug (and getting the three-year NI marketing exclusivity) has high economic value when the drug is covered by an NCE, moderate economic value after the NCE expires but patents still partially block generic entry,⁹ and low economic value upon generic entry. Since some new uses emerge late during a drug's life cycle, manufacturers may often find it optimal not to pursue an NI exclusivity.

As an alternative, a manufacturer may pursue approval for a new drug. For example, Glaxo-Wellcome gained FDA approval for bupropion hydrochloride in 1985, called it Wellbutrin (IR formulation) and marketed it as an antidepressant.¹⁰ Some time afterward, Dr. Linda Ferry, a family

⁸Innovating firms can also earn (up to five years) patent life lost during the clinical-trials process.

⁹Under the Hatch-Waxman Act, a generic may file for an ANDA once four of the five NCE years have expired. It must then show the pioneering firm's patents are invalid or would not be infringed by the generic product.

¹⁰Initially, there were a significant number of seizures at the recommended dosage of 400-600 mg. Glaxo removed Wellbutrin from the market in 1986. After subsequently discovering that reducing the dose by about half sharply reduced the risk of seizures, Glaxo reintroduced Wellbutrin to the market in 1989 with a maximum dose of 450 mg/day. See <http://www.emedexpert.com/facts/bupropion-facts.shtml>.

practitioner in California, noticed that many of her patients taking Wellbutrin (for depression) showed decreased interest in smoking. She then convinced Glaxo to pursue clinical trials. After successful trials, Glaxo got approval in 1997 for an on-label indication for bupropion as a treatment for smoking cessation. However, Glaxo did not add smoking cessation to the label for Wellbutrin (and they never have). Rather, they got smoking cessation on the label for a new drug, Zyban (Perkins *et al.* [2008], pp. 113-114). Obfuscating the relationship between Wellbutrin and Zyban may mitigate the advertising spillover effect.

Numerous legal challenges around the practice recently led to a settlement between the FDA and Amarin Corporation. The settlement, albeit for a single drug, opens the door once again for off-label promotion if a drug has been approved by the FDA for another indication, and if the promoted off-label use is supported by clinical evidence. Although the settlement does not represent a legal precedent, it is clearly a relaxation of the current regulatory view of the practice, and may reflect a change in the rigor with which the FDA will pursue enforcement of off-label promotion.

II(ii). *Literature Review*

Given the FDA's regulatory architecture, there are a variety of reasons that drugs' labels do not cover all uses. Numerous papers discuss anecdotes of off-label use. Collectively, these anecdotes illustrate that off-label use may give a patient the highest-quality option, or may give a lower-quality (but cheaper) option.

For example, beta-blockers, such as metoprolol and propranolol, have been used for decades to treat hypertension, cardiac dysrhythmias, and other diseases. Clinicians have noted that beta-blockers also control physical sensations associated with anxiety (such as rapid heartbeat, tightness in the chest, and trembling), and that when patients do not feel these sensations, their psychological experience of anxiety is significantly reduced. As a result, these drugs are widely prescribed for situational and other forms of anxiety, to apparent great effect. Lin *et al.* [2006] estimate that 52% of prescriptions for beta-blockers were off-label from 1999-2002. Anxiety is not, however, an approved use of any beta-blocker. Such examples of effective re-purposing are not rare and off-label use is not a new phenomenon. In a 1991 survey, roughly one-third of cancer drugs were prescribed off-label and about one-half of patients were prescribed at least one drug off-label (USGAO [1991]).¹¹

¹¹Salbu [1999] and Klein and Tabarrok [2004] discuss a number of other examples of beneficial off-label use.

On the other hand, off-label use may introduce production complications that lower the quality of drugs used off-label relative to the best on-label options. Consider macular degeneration, which occurs when a person's retina is damaged and the person loses vision in the center of the visual field. It is quite common—approximately 10% of people 66-74 years old have macular degeneration.

Lucentis is FDA approved to treat macular degeneration. It sells for \$2,000 a dose, and a typical regimen is monthly injections for 12-24 months. Avastin is FDA approved to treat metastatic colon cancer (and some other cancers) but has been shown to be effective in treating macular degeneration. In treating macular degeneration, it sells for \$50 a dose because the dosage is far smaller for this than for treating cancer. Again, a typical treatment regimen is monthly injections for 12-24 months. For this use, however, Avastin must also be 'compounded' by pharmacies. There have been cases of bad batches of Avastin, and these have caused ocular infections and blindness.

This creates an obvious dilemma for ophthalmologists. Prescribe a \$2,000 medicine which is indicated but which many patients can't or won't pay for?¹² Or prescribe an off-label medicine which costs \$50 a dose but carries a risk of infection?¹³

There are also many cases, especially among the drugs involved in the DOJ settlements, where pharmaceutical firms are alleged to have marketed their drugs for off-label uses with uncertain clinical support (Stafford [2008]). In perhaps the best-known case, Warner-Lambert's drug Neurontin was initially approved for 'adjunctive therapy in the treatment of partial seizures...in patients above the age of 12 years,' and was later approved for patients 3-12 years old (2000) and for 'postherpetic neuralgia' (2004). Physicians have prescribed Neurontin for a number of off-label indications, such as bipolar disorder and neuropathic pain, for which the evidence for effect is at best equivocal (Mack [2003]). In 2004, Pfizer (which merged with Warner-Lambert in 2000) pled guilty to two felony counts of marketing a drug for unapproved uses, and paid \$430 million in civil and criminal fines. Some \$26.6 million went to whistleblower David Franklin, who started working for Warner-Lambert in 1996 (Evans [2009]).

¹²Note that for a typical Medicare reimbursement of 80%, that still leaves \$400 per dose of Lucentis uncovered.

¹³In addition, off-label use has led to clear cases of patients being harmed. During the 1990's, after Dr. Michael Weintraub showed that a group of 121 patients using a combination of the weight-loss drugs fenfluramine and phentermine lost an average of 30 pounds, this 'Fen-Phen' combination surged in popularity. Because neither drug's label discussed using the drugs in combination, this was an off-label use. Dr. Weintraub looked for side effects, but he assumed the drugs were safe (Kolata [1997]). Unfortunately, numerous patients suffered heart-valve damage (O'Reilly and Dalal [2003]).

In the closest related work, Radley *et al.* [2006] rely upon data from the *National Disease and Therapeutic Index* (NDTI), a survey of physicians from IMS.¹⁴ Their data include the 100 medications most commonly used in 2001, plus a random group of 60 additional medications. They find off-label use at about 21% of overall use, and additionally classify other uses according to whether there was any scientific support for the off-label use and differentiate rates of off-label use using a variety of drug characteristics. They do not study patient characteristics as drivers of off-label use, nor do they analyze the effect of regulatory intervention. We view our work characterizing economic effects of variables such as insurance status and DOJ scrutiny as complementing theirs. We also view our development of a technique to study off-label use with non-proprietary data as a useful contribution.

Numerous additional clinical papers study rates of off-label prescribing for narrower sets of drugs. Results vary widely and are virtually impossible to compare to each other. For example, in a survey of papers studying off-label use in pediatrics, Cuzzolin *et al.* [2003] identify 16 studies published during 1995-2001. Of these, 13 are prospective studies, two are retrospective, and one is prescription-event monitoring (a much longer study, Ten years). The number of patients varies from 40 to 24,337, while the number of prescriptions varies from 257 to 4,455.¹⁵ The percentage of off-label use varies from a low of 10.8% (McIntyre *et al.* [2000], ‘pediatric ambulatory’) to a high of 72% (Avenel *et al.* [2000], NICU). Finally, USGAO [1991] and Molitor [2012] examine some drivers of off-label drug use by cancer patients. USGAO [1991] shows, for example, no pattern of off-label use by age group and gender, while Molitor [2012] finds that over 20% of new cancer drug use within the Medicare population during 1998-2008 was off-label.

III. A SIMPLE MODEL OF DRUG CHOICE

We now present a highly-stylized model of vertical product differentiation. This helps us to understand the potential implications of the Lucentis/Avastin example from Section II(ii). It also helps us interpret our empirical results that follow in Section VI.

Assume consumers have a condition that may be treated by some pharmaceutical product with some maximum observable clinical treatment effect, which will generate a value of V for a consumer, if she receives the full benefit. Further, assume that any individual consumer’s response to the drug is idiosyncratic – she may receive the full measured benefit, or she

¹⁴These data are ‘nationally representative diagnostic and treatment data similar to that contained in the National Ambulatory Medical Care Surveys (NAMCS)’ (Radley *et al.* [2006])

¹⁵Some studies did not indicate numbers for both categories.

may not benefit at all. The magnitude of any individual consumer's realized treatment effect from the drug is her match quality. Consumers are arrayed along a distribution of responses to any given drug, so that each drug will be characterized by an average match quality, summarized by the parameter, θ . Average match quality for a drug is uniformly distributed on the $[0, 1]$ interval. A representative consumer's net (expected) utility of purchasing a drug of quality V for price P is $\theta V - P$. The consumer prefers to buy the drug rather than do nothing if $\theta > \frac{P}{V}$.

Let there be a continuum of markets for drugs. A market may be thought of as a medical condition (e.g., depression or angina) for which pharmacological treatments are available. In each market assume that there is one on-label option and one off-label option. To simplify the analysis, assume that markets may be characterized by outcomes – a high value-generating option and a low value-generating one. In fraction η of the markets, the on-label option yields value for the representative consumer of V and the off-label option generates value \underline{V} . In fraction $1-\eta$ of the markets, the on-label option generates a value of \underline{V} and the off-label option has yields a value of V .

Let prices for uninsured consumers be P and \underline{P} for the higher and lower average match quality drug, respectively. Insured consumers pay $P^0 < \underline{P}$ for either drug. Then uninsured consumers with $\theta > \frac{P-P}{V-\underline{V}}$ purchase the higher average match quality drug. Uninsured consumers with $\theta \in [\frac{P}{\underline{V}}, \frac{P-P}{V-\underline{V}}]$ purchase the lower match quality drug. Uninsured consumers with $\theta < \frac{P}{\underline{V}}$ purchase nothing. Insured customers always prefer the higher match quality drug, and all consumers with $\theta > \frac{P^0}{\underline{V}}$ buy this drug.

Thus, conditional on having insurance, the fraction of consumer choices that are the on-label option is just η , which is the probability the on-label option has the higher average match quality. For uninsured consumers, the fraction of consumer choices that are on-label is

$$Pr(OnLabel) = \frac{\eta \left(1 - \frac{\bar{P}-P}{V-\underline{V}}\right) + (1-\eta) \left(\frac{\bar{P}-P}{V-\underline{V}} - \frac{P}{\underline{V}}\right)}{1 - \frac{P}{\underline{V}}}.$$

It is easy to show that this is lower than η (the analogous probability of on-label use by consumers with insurance) if and only if $\eta > \frac{1}{2}$. Hence, off-label prescribing rises with insurance if and only if off-label options are more typically superior ($\eta < \frac{1}{2}$).

It may be, of course, more natural to expect $\eta > \frac{1}{2}$. If this is the case, then to get more off-label use under insurance, it is necessary to have more subscribers switch to an off-label drug when the off-label option is superior, than switch to an on-label drug when the on-label drug is superior.

One way to achieve this result is to let the higher average match quality drug's price be different in the η markets where the on-label option is superior. Specifically, let it be γP , where $\gamma > 1$ implies a higher price and $\gamma < 1$ implies a lower price, relative to the price of the off-label drug in the $1-\eta$ markets where the off-label drug is superior. It may be that $\gamma < 1$ if physicians are willing to cover extra costs for their patients in prescribing on-label options (i.e., ophthalmologists willing to cover the extra costs to allow their patients to purchase Lucentis, rather than take risks with Avastin) but unwilling to cover extra costs in prescribing off-label options. For example, when physicians give out free samples to their patients, they are lowering the out of pocket costs for patients for the prescribed treatment regimen. If physicians do this more often for on-label prescriptions than off-label, then $\gamma < 1$ could obtain.

Now for uninsured consumers, we have

$$Pr(OnLabel) = \frac{\eta \left(1 - \frac{\gamma \bar{P} - P}{\bar{V} - \underline{V}} \right) + (1 - \eta) \left(\frac{\bar{P} - P}{\bar{V} - \underline{V}} - \frac{P}{\underline{V}} \right)}{1 - \frac{P}{\underline{V}}}$$

Now suppose that $\gamma < 1$. Then, it is possible to have $\eta > \frac{1}{2}$ but still have higher rates of off-label prescribing with insurance. Intuitively, relatively few people consume the off-label drug when it is inferior, so relatively more people switch to (superior) off-label options when they obtain insurance. Thus, a positive relationship in our data between the frequency of off-label use and the generosity of a patient's insurance suggests that observed off-label use enhances welfare.

IV. DATA

Our data are drawn primarily from two main sources. The *National Ambulatory Medical Care Survey* (NAMCS) provides information on the prescriptions written for a representative set of U.S. physician office visits. NAMCS data are collected for 1993-2008. The *Supplement to the Physicians' Desk Reference* (PDR) is titled either the *PDR Companion Guide* or *PDR Guide to Drug Interactions, Side Effects, Indications, Contraindications*, depending on the year. This serves as physicians' primary reference on FDA-approved indications for each drug, for 1993-2008 as well.

No existing research combines these two sources.¹⁶ Moreover, because of the incentives facing firms to get indications on drug labels, the usually unanticipated timing of discoveries of useful ways to repurpose drugs, and the ways physicians' manuals (like the PDR) list sets of approved indications, there are multiple ways of classifying on-label and off-label use. To clarify, we first define two types of *on-label use*:

¹⁶Radley *et al.* [2006] use the PDR to classify off-label use.

1. A *drug-label use* occurs whenever a drug is used for an indication on the drug's FDA-approved label.
2. An *active-ingredient-label use* occurs whenever a drug is used for an indication on any label of any drug with the same active ingredient.

For any drug, the set of drug-label uses is a subset of the set of active-ingredient-label uses.

Next, we define two types of off-label use.

1. An *anticipated off-label use* is any active-ingredient-label use that is not a drug-label use.
2. An *unanticipated off-label use* is any use that is not an active ingredient use.

Legally, active-ingredient uses that are not drug-label uses are off-label uses. However, these off-label uses are natural candidates for prescribing physicians given that the use has been approved for some other drug with the same active ingredient. Hence, we say such off-label uses (e.g., Wellbutrin used for smoking cessation) are anticipated. For the purpose of our analysis, we define off-label use as *unanticipated off-label use*. All other uses of a drug will be defined as on-label.

In addition to NAMCS and PDR data, we collect a comprehensive list of settlements between the DOJ and pharmaceutical companies for alleged promotion of off-label uses.

IV(i). *NAMCS*

The NAMCS is an annual survey, conducted by the Center for Disease Control and Prevention (CDC). The multi-state design generates a sample of patient visits to community-based (non-institutional, non-clinic) physician offices. One big advantage of NAMCS is that the sample is specifically drawn to be representative of the population of visits in the U.S. Each physician is randomly assigned a one-week participation window and uses an abstraction form to characterize approximately 30 office visits. This results in detailed quantitative descriptions of around 35,000 office visits each year. Visit weights are assigned so that national estimates can be produced.¹⁷

¹⁷This use of the NAMCS weights to estimate prescribing volume has been validated in the literature; see Pincus *et al.* [1998], Thomas *et al.* [2006], and Iizuka [2004]. Note that the NAMCS visit weights are calculated to impute prescriptions to the national and annual level. Given the weight construction we could allocate annual prescription estimates to the month by dividing the weighted estimate by 12; allocating prescriptions to the region could be accomplished by re-weighting on the basis of the region's proportion of the national population. In either case, any biases from not explicitly adjusting our prescription counts will be isolated in the intercept term, and will not affect the marginal effects of interest.

Another advantage is that NAMCS data record the physician's prescribing decisions, and are not based upon which prescriptions are actually filled. From 1993-94, the physician may record up to five prescriptions. This maximum increases to six from 1995-2002, and to eight for 2003-08. The physician may record up to three indications. The NAMCS data code indications according to the ICD-9 classification.

We extract data on all visits from 1993 to 2008. After we keep those visits with at least one prescription, the resulting data set contains information on 549,092 prescriptions from 249,146 office visits during 1993 to 2008. In addition to prescriptions and indications, we extract information on visit payment source, physician specialty, census region, and the patient's race, age, gender, and insurance coverage.

HIGHLIGHTS OF PRESCRIBING INFORMATION

These highlights do not include all the information needed to use CREON[®] safely and effectively. See full prescribing information for CREON.

**CREON (pancrelipase) Capsules, Delayed Release for Oral Use
Initial U.S. Approval: _____**

-----INDICATIONS AND USAGE-----

CREON Capsules is a pancrelipase which is a combination of porcine-derived lipases, proteases, and amylases indicated for the treatment of exocrine pancreatic insufficiency due to cystic fibrosis or other conditions.

Figure 1

The Current FDA Label for Creon (Pancrelipase)

Notes: This figure shows the FDA label for the drug Creon

IV(ii). PDR

Each prescription drug available for sale in the U.S. has an FDA-mandated label, which lists (among many other things) the indications for which the FDA has approved use. Figure 1 is an extract from the first part of the Creon label. Its label shows the drug-label use (treatment of exocrine pancreatic insufficiency due to cystic fibrosis, chronic pancreatitis, pancreatectomy, or other conditions) in specific populations. Identifying the history of FDA-approved indications for a drug is surprisingly difficult. Although the FDA maintains the current FDA-approved indications for a drug in the Orange Book, these listings do not include the history of how the label reached its current state.

The annual *Supplement* to the PDR fills this void. In particular, the 'Indications' section lists FDA-approved drugs for each indication, giving physicians up-to-date information on a drug's approved uses, interactions,

side effects, and contraindications. We obtain each edition of the PDR from 1993 to 2008 and capture all drug-indication combinations.¹⁸ Figure 2 provides an example of the *Supplement's* formatting from 1997. For the indication 'Pancreatic cystic fibrosis,' there are seven associated FDA-approved drugs: Cotazym Capsules, Creon, KuZyme HP Capsules, Pancrease Capsules, Pancrease MT Capsules, Viokase, and Zymase Capsules. PDR records also include the active ingredient for each drug—in this case, Pancrelipase for all drugs—and the manufacturers.

Note that this list represents anticipated off-label uses as being on-label. For example, Creon is approved for 'treatment of exocrine pancreatic insufficiency due to cystic fibrosis or other conditions,' i.e., pancreatic cystic fibrosis. But Viokase is approved just for 'treatment of exocrine pancreatic insufficiency in adults due to chronic pancreatitis or pancreatectomy,' i.e., not pancreatic cystic fibrosis. Use of Viokase for pancreatic cystic fibrosis, like Wellbutrin for smoking cessation, is an active-ingredient use that is an anticipated off-label use.

Pancreatic cystic fibrosis	
Cotazym Capsules (Pancrelipase)	
Organon	1866
Creon (Pancrelipase) Solvay	2714
Ku-Zyme HP Capsules	
(Pancrelipase) Schwarz	2547
Pancrease Capsules (Pancrelipase)	
McNeil Pharmaceutical	1589
Pancrease MT Capsules	
(Pancrelipase) McNeil	
Pharmaceutical	1589
Viokase (Pancrelipase) Robins	2251
Zymase Capsules (Pancrelipase)	
Organon	1889

Figure 2

An Example of On-Label Drugs for an Indication in the PDR

Notes: This figure shows an example of a list of FDA-approved drugs for pancreatic cystic fibrosis. Source: Supplement to the *Physician's Desk Reference*, 'Indications' section

¹⁸To extract this information from over 1,000 pages formatted in this way, we remove the *Supplement's* binding and use a high resolution scanner to create images of each page. We use *Able2Extract Professional*, software that uses Optical Character Recognition (OCR) technology, to extract the information to a raw text format to parse into a useable form. Following the extraction and parsing processes, we search for and remove the (surprisingly small) number of errors in the resulting files. The consistency in the *Supplement's* format and few changes in the text descriptors for indications and drug names make this feasible. For example, to catch errors in the text describing an indication, we sort the entire panel (all years) by indication and flag subtle changes to the text descriptor. We then compare the flags to the original image files. We use a similar cleaning process for drug names.

We format the data so that the unit of observation is an indication-year-ingredient (IYI) combination. For the text from Figure 2, the IYI is 'pancreatic cystic fibrosis - 1997 - pancrelipase.' There are 178,823 total IYIs, composed of 3,587 unique indication descriptors and 3,342 unique active ingredients.

IV(iii). *Matching NAMCS to PDR*

Drugs and indications are each defined differently in the NAMCS and the PDR. Hence, matching drugs to indications for the purposes of identifying off-label use requires careful interpretation of language from both definitions. For drugs, we use a word-based matching algorithm to match the text for each active ingredient in the PDR to the NCHS Multum codes used by the NAMCS to identify active ingredients. This corrects for abbreviations and is virtually one-to-one.¹⁹

Matching indications in the PDR to indications in the NAMCS is far more complicated. The PDR uses terms for indications that do not map directly to the ICD-9 codes used on NAMCS prescription records. In fact, processing the language in these descriptors and matching them to ICD-9 codes is sufficiently complicated that we rely on a Clinical Documentation Specialist, or *coder*, from a large academic medical center.²⁰ The coder manually matches each PDR indication to the appropriate ICD-9 code(s).²¹

This process is often not straightforward. Some of the language in the PDR indications requires further research for additional information to establish an appropriate match. In other situations where the PDR indication is not specific enough to map to a single 5-digit ICD-9 code, we match to a less-specific (four- or three-digit) code corresponding to a broader range of indications. For example, 'Pneumonia, Community Acquired' and 'Pneumonia Nosocomial' are listed as separate indications in the PDR. 'Pneumonia Nosocomial' is a more complex pneumonia with an easily identified ICD-9 code for Gram Negative Pneumonia, 482.83. 'Pneumonia, Community Acquired' is less specific and is coded as a Pneumonia, NOS (not otherwise specified), 486. In other situations, a descriptor requires more than one ICD-9 code to adequately describe the indication. In these situations, the coder lists all relevant codes.

Hence, matching PDR indications to NAMCS indications is a many-to-many match. Our conservative matching approach also leads us to

¹⁹The few instances where the match is not one-to-one are due to combinations.

²⁰This person's primary role at the medical center is to look for missing, vague, or incomplete physician documentation and to query the physician for the specificity needed to ensure that the coders can capture the most accurate indication, indication or symptoms for the patient.

²¹Our coder was provided a spreadsheet with the 3,587 unique text descriptors from the PDR for 1993-2008, as well as the *ICD-9-CM Manual* Volumes 1, 2, 3 published by HCPro, Inc [2011].

undercount off-label use. Alternatively, we could make subjective decisions regarding the closest matching indication. We regard a conservative approach as most sensible here.

Having matched PDR indications to NAMCS indications and PDR active ingredients to NAMCS active ingredients, we reformat our data so that the unit of observation is a prescription. For each observation, we then list the characteristics of the patient, the date of the visit and the complete list of indications assigned for the visit. For each prescription, we cycle through the indications to determine whether the set of approved active ingredients for those indications includes the prescribed drug. To be conservative in our estimates of off-label use, we convert the NAMCS indications to the three-digit ICD-9 level before querying the set of approved drugs. For example, 410.01 would be matched to any 410 code.²²

We set the *on-label indicator* to 1 if the prescribed drug's active ingredient is among the set of approved active ingredients for at least one of the indications listed for the visit. This serves as the dependent variable in our analysis. As we discuss in more detail in Section IV, this variable cannot be used to measure directly the amount of on-label use because of imperfect detection.

IV(iv). *DOJ Settlements*

To collect information on settlements between the DOJ and pharmaceutical companies, we rely on a number of sources. For many of the settlements, the details are directly available on the DOJ's website (*justice.gov*). Many of the settlements also received substantial popular-press coverage (*New York Times*, *Wall Street Journal*, etc, mainly due to the implications for financial markets), which provided additional detail and confirmation of settlement details from other sources.²³ Table I lists the details of each settlement: drug name, company being sued, date at which the company was aware of the DOJ investigation, and the size of the settlement. We use all drugs in settlements between 2000 and 2008 (i.e., those with asterisks in Table I).

From this data, we construct three variables. The first is an indicator that equals one if a drug is investigated by the DOJ for off-label promotion, and zero otherwise. This indicator helps identify whether the DOJ successfully targeted drugs with high rates of off-label use that cannot otherwise be explained. The next is an indicator that equals one for investigated drugs after the DOJ announced the investigation, and zero otherwise. This indicator captures the net effect of the investigation on alleged promotional activities by the firm and physicians' awareness of the drug's off-label uses. Finally, we construct a discrete variable which equals the number of years since the DOJ announcement that it had begun an investigation, and zero before the

²²Our results are nearly identical regardless of whether we match using a three-digit or four-digit ICD-9 code.

²³Our list includes all those considered by Kesselheim *et al.* [2011].

TABLE I
DOJ SETTLEMENTS ASSOCIATED WITH OFF-LABEL USE

Drug Name(s)	Company	DOJ	
		Aware	Liability (\$Millions)
Neurontin	Davis	8/1996	430.0
*Serostim	Serono	9/2000	704.0
*Temodar/Intron A	Schering-Plough	6/2001	435.0
*Topamax	Ortho-McNeil Pharma.	8/2003	81.5
*Actiq/Gabitril/Provigil	Cephalon	11/2003	425.0
*Advicor/Niaspan	Kos Pharma.	3/2004	41.0
*Actimmune	InterMune	7/2004	36.9
*Zonegran	Forest Laboratories	7/2004	203.0
*Zonegran	Ian/Eisai	7/2004	11.0
*Lidoderm	Endo Pharma.	7/2005	192.7
*Tobi	Novartis	1/2006	72.5
*Trisenox	Cell Therapeutics	2/2006	10.5
*Xyrem	Orphan Medical	2/2006	20.0
*Seroquel	AstraZeneca	3/2006	520.0
*Abilify	Bristol-Myers Squibb	5/2006	515.0
*Kadian	Alpharma	9/2006	42.5
*Zyprexa	Eli Lilly and Co.	12/2006	1,415.0
*Bextra/Geodon/Zyvox/Lyrica	Pfizer	6/2007	2,300.0
*Botox	Allergan	6/2007	600.0
*Depakote	Abbott Laboratories	10/2007	1,500.0
*NovoSeven	Novo Nordisk	10/2008	25.0
Keppra	UCB	1/2009	34.0
Levithroid/Celexa	Novartis	2/2009	313.0
Protonix	Wyeth	5/2009	55.0
Detrol	Pfizer	1/2010	14.5
Risperdal/Invega/Natrecor	Johnson & Johnson	1/2010	2,205.0
Rapamune	Wyeth	5/2010	490.9
Trileptal	Novartis	8/2010	422.5
ChloraPrep	CareFusion	9/2010	40.1
Aranesp/Enbrel/Neulasta	Amgen	8/2011	762.0
Paxil/Wellbutrin/Advair/ Lamictal/Zofran	GlaxoSmithKline	10/2011	3,000.0
Megace ES	Par Pharmaceutical	3/2012	22.5

(Continued)

TABLE I (CONTINUED)

Drug Name(s)	Company	DOJ	Liability
		Aware	(\$Millions)
Total			16,949.9

Notes: Information on the settlements were collected from a variety of sources, including but not limited to DOJ press releases at justice.gov, various law school sites (Harvard, Virginia, etc.), and sites of state attorney general offices (e.g., ag.ny.gov). It is important to note that in many of the settlements, the magnitude of the settlements may represent simultaneously agreeing to resolve other litigation matters. Further, firms often do not admit guilt. For example, in the case of Serostim, Merck Serono denies both the DOJ's claims and liability but settled 'solely to avoid the further expense, inconvenience, burden and uncertainty of these lawsuits.' Settlements marked with an asterisk (*) are used in estimation.

investigation began. We use this variable to measure whether the effect of DOJ intervention changed once the alleged illegal activity ended, which would suggest that the effect of the promotional activities diminishes with time.²⁴

IV(v). *Descriptive Statistics*

Table II shows the demographic composition of the individuals and physician specialties associated with prescriptions written in our sample from the NAMCS. Just over 58% of prescriptions are written for female patients, 8.9% are for African-Americans, and 7.5% are for Hispanics. The average prescription is written for a patient who is 51.8 years of age. Over 93% of prescriptions are written for patients who have some form of insurance—47.5% with private insurance, 31.4% with Medicare, 10.2% with Medicaid, and 4.1% with some other form of insurance. Just over 88% of the prescriptions in the data are written during visits with established patients (i.e., it is not the first time the patient visited the physician). The physician specialties with the greatest frequency are internal medicine (13.2%) and cardiology (11.2%). The midwest census region accounts for the largest proportion of prescriptions at 34.4%, while the east census region (omitted region) accounts for the lowest, only 20.3%.²⁵

²⁴Note that we drop Neurontin from the analysis. We have almost no observations in the before-DOJ-announcement period, and if included it would account for almost half of the number of prescriptions for all DOJ-investigated drugs. It was also the first drug targeted for off-label marketing, so it is likely that drug manufacturers had different information about the implications of DOJ scrutiny for Neurontin than for other drugs.

²⁵While this sample is representative of the population we wish to study, it is not representative of the U.S. population. In comparison with the 2000 U.S. Census, our sample differs slightly. It over-represents females (58.8% vs. 50.1%), patients with Medicare (31.3% vs 13.4%), patients with Medicaid (10.4% vs. 10.2%), the Midwest (34.4% vs. 21.7%) and Northeast (20.3% vs. 17.9%). It under-represents African-Americans (8.9% vs. 12.3%), Hispanics (7.5% vs. 12.5%), the South (22.8% vs. 37.1%) and the West (22.5% vs. 23.3%). The source for census data on gender and race/ethnicity is <https://www.census.gov/census2000/demoprofile.html>; the source for census data on health insurance is <https://www.census.gov/prod/2001pubs/p60-215.pdf>; the source for census data on U.S. regions is <https://www.census.gov/popclock/>.

TABLE II
 PATIENT AND PHYSICIAN VARIABLES

	mean	sd
Patient age	51.84	23.22
Patient is female	0.580	0.494
Patient is Hispanic	0.075	0.263
Patient is African-American	0.089	0.284
Patient is other (non-White) race	0.093	0.290
Established patient	0.883	0.321
Patient has Medicare	0.314	0.464
Patient has Medicaid	0.102	0.303
Patient has private insurance	0.475	0.499
Patient has other insurance	0.041	0.198
Physician specialty is internal medicine	0.132	0.339
Physician specialty is pediatrics	0.064	0.245
Physician specialty is OB/GYN	0.023	0.150
Physician specialty is cardiology	0.112	0.316
Physician specialty is psychiatry	0.071	0.256
Physician specialty is neurology	0.061	0.238
Physician has other specialty	0.306	0.461
Southern Census region	0.228	0.420
Mid-West Census region	0.344	0.475
West Census region	0.225	0.418
Observations	549,092	

Notes: These statistics reflect the demographic composition and insurance status of patients in our NAMCS sample of prescriptions during 1993-2008, as well as the specialty and location of the physician treating them.

Table III shows that the most commonly prescribed drugs are central nervous system agents (19.6%), cardiovascular agents (18.5%), and anti-infective agents (10.5%). The most common indications are associated with circulatory disorders (25.4%) and respiratory disorders (17.7%). The frequencies of the indications do not sum to one, because all indications from a visit are listed for each prescription from the same visit.

On average, a patient is prescribed 3.9 prescriptions per visit. However, the standard deviation is quite large. Across all visits, a physician has an average of 71.4 FDA-approved alternatives for the entire set of (up to three) visit-specific indications, defined at the three-digit ICD-9 level, reported in NAMCS. The median number of approved prescription alternatives (i.e., unique Multum codes) per indication is 27.7. We expect that when the

on-label alternatives for treating a patient are limited (i.e., a small number of approved drugs for the complete set of observed indications), a physician will be more likely to prescribe off-label.

The bottom panel of Table III shows that about 4.5% of our observations include a drug targeted by DOJ, and about 0.8% of observations occur after the firm was aware of this targeting. For many of the cases, we have at least a few years after the DOJ sued the company. This permits us to examine whether the effect of the DOJ's action increases over time, as would be expected if it were successful at eliminating promotional activities whose persistent effects diminish with time, resulting in a lower rate of off-label use.

The NAMCS survey form changes significantly over our sample. From 1995 to 2002 (40.2% of our sample), the form allows six prescriptions to be reported, up from only five in 1993 and 1994. From 2003 onward, the form allowed a total of eight prescriptions to be listed (48.5% of our sample). The form never increases the maximum allowable number of indications, three, during our period of study. Over 61.7% of prescriptions appear on a form with two indications listed and 33.5% appear on a form with three indications.

Table IV highlights how the NAMCS form limitations affect detection. The numbers in the table are conditional means of the on-label indicator, by numbers of prescriptions and indications. If the form limitations have identifying power, then our ability to detect on-label use should be highest for those visits with a low number of prescriptions and a high number of indications (top right corner of the table), and lowest for those visits with a high number of prescriptions but a low number of indications (lower left corner of the table). These are precisely the patterns we observe. The conditional means of this indicator rise virtually monotonically up the columns and across the rows, and plateau in the top-right corner of Table IV. Thus there is a group of observations for which the NAMCS form limitations appear to have significantly less impact on our ability to detect on-label use.

Hence, time-varying form limitations are very useful for identification purposes. Each of the exogenous increases in the maximum number of prescriptions that can be recorded was not accompanied by a similar increase in the maximum number of recorded indications during a visit. These form limitations result in only 33.5% of all prescriptions in our data appearing to be written for an FDA-approved use, or 66.5% off-label use. In the next section, we discuss how we correct for detection error and identify the rate of on-label use.

V. THE ECONOMETRIC MODEL

We cannot use our naïve indicator of on-label use directly because of *false negatives*. When a prescription is on-label but the indication is not listed on the NAMCS survey form, true on-label use is unobserved. To identify the

TABLE III
PRESCRIPTION AND INDICATION VARIABLES

	mean	sd
Prescription observed to be on-label	0.335	0.472
Number of drugs approved to treat diagnoses	71.39	69.24
Number of prescriptions written during visit	3.868	2.245
Two indications recorded on NAMCS form	0.618	0.486
Three indications recorded on NAMCS form	0.335	0.472
NAMCS form allows 6 listed medications	0.402	0.490
NAMCS form allows 8 listed medications	0.485	0.500
Anti-infectives	0.105	0.306
Cardiovascular agents	0.185	0.388
Central nervous system agents	0.196	0.397
Coagulant modifiers	0.018	0.134
Gastrointestinal agents	0.044	0.206
Hormonal agents	0.063	0.242
Miscellaneous agents	0.024	0.153
Respiratory agents	0.083	0.276
Topical agents	0.094	0.292
Psychological agents	0.081	0.272
Immunologic agents	0.028	0.166
Metabolic agents	0.079	0.270
ICD9 code for infectious and parasitic disease	0.035	0.184
ICD9 code for neoplasm	0.039	0.193
ICD9 code for endocrine disorders	0.154	0.361
ICD9 code for mental disorders	0.130	0.336
ICD9 code for nervous system disorders	0.135	0.342
ICD9 code for circulatory system disorders	0.254	0.435
ICD9 code for respiratory system disorders	0.177	0.382
ICD9 code for digestive system disorders	0.058	0.234
ICD9 code for genitourinary system disorders	0.068	0.252
ICD9 code for skin disorders	0.073	0.261
ICD9 code for musculoskeletal system disorders	0.122	0.328
ICD9 code for ill-defined disorders	0.138	0.344
ICD9 code for injury and poisoning	0.046	0.209
Manufacturer ever investigated by DoJ	0.045	0.207

(Continued)

TABLE III (CONTINUED)

	mean	sd
After DoJ announced investigation for DoJ manufacturers	0.008	0.083
Years since DoJ announced investigation for DoJ manufacturers	0.011	0.185
Observations	549,092	

Notes: These statistics reflect information on the prescriptions and indications in our NAMCS sample during 1993–2008. The top portion of the table provides information on the number of prescriptions and indications, while the middle and bottom portions provide information on the types of drugs and indications, respectively.

rate of on-label use, we adopt the DCE methodology introduced by Feinstein [1990].²⁶

Let a physician's decision, of whether to write prescription i for a drug that is on-label, be summarized by a stochastic latent variable,

$$(1) \quad Y_{1i}^* = x_{1i}\beta_1 + \varepsilon_{1i},$$

where ε_{1i} is independently and identically distributed $N(0, 1)$. The binary outcome of this decision-making process, which is not observed, is then

$$(2) \quad Y_{1i} = \begin{cases} 1 & \text{if } Y_{1i}^* > 0 \\ 0 & \text{otherwise.} \end{cases}$$

With the normality assumption, the probability that the physician prescribes on-label, conditional on x_{1i} , is

$$(3) \quad Pr(Y_{1i} = 1) = \Phi(x_{1i}\beta_1),$$

where $\Phi(x_{1i}\beta_1)$ is the standard normal cumulative distribution function (CDF). The variables in the linear index, $x_{1i}\beta_1$, alter the likelihood of a physician's writing a prescription for an approved indication.

The difficulty is that we do not observe this binary outcome and cannot infer our outcome of interest (i.e., $\Phi(x_{1i}\beta_1)$) by estimating a Probit model. Instead, we observe an indicator that takes on a value of one when a prescription is both on-label and we can detect it in the data. Conversely, the indicator takes on a value of zero when the prescription is actually off-label or when we have a false negative. By specifying a model for detection, along with the equation characterizing the physician's choice to prescribe on-label, DCE allows us to distinguish between these two cases.

Consider cases when on-label prescribing actually occurs and let

$$Y_{2i} = \begin{cases} 1 & \text{if on-label use is detected} \\ 0 & \text{otherwise.} \end{cases}$$

²⁶DCE has been applied previously in a health context by Kleit and Ruiz [2003] and Bradford *et al.* [2001], and a non-health context by Helland [1998], Feinstein [1989, 1991].

TABLE IV
FORM LIMITATIONS AND DETECTION, MEAN Y_{2i}

	(1)	(2)	(3)	(4)
	One Indications	Two Indications	Three Indications	Average
One prescription	0.379	0.449	0.463	0.409
Two prescriptions	0.319	0.455	0.470	0.392
Three prescriptions	0.272	0.414	0.465	0.371
Four prescriptions	0.207	0.365	0.429	0.335
Five prescriptions	0.164	0.308	0.390	0.302
Six prescriptions	0.114	0.256	0.361	0.277
Seven prescriptions	0.086	0.202	0.318	0.234
Eight prescriptions	0.069	0.151	0.277	0.206
Average	0.267	0.365	0.388	0.335
Observations	209,986	155,231	183,875	549,092

Notes: These statistics reflect the mean of Y_{2i} , the indicator for whether on-label use both occurred and is detected, by number of prescriptions and indications recorded by the physician on the NAMCS survey form.

Thus, when $Y_{2i} = 1$, on-label use occurs (i.e., $Y_{1i} = 1$) and is detected. This is precisely the naïve on-label indicator we construct. Further, assume the distribution of the underlying latent variable for the Y_{2i} indicator is given by

$$Y_{2i}^* = x_{2i}\beta_2 + \varepsilon_{2i}.$$

We assume ε_{2i} is distributed standard normal and the linear index, $x_{2i}\beta_2$, includes those variables that alter the probability of detecting off-label use.

The probability of detecting true on-label use is

$$Pr(Y_{2i} = 1) = Pr(Y_{2i} = 1 | Y_{1i} = 1) * P(Y_{1i} = 1) + Pr(Y_{2i} = 1 | Y_{1i} = 0) * P(Y_{1i} = 0).$$

Assuming there are no *false positives*, i.e., all detected on-label use is truly on-label, then $Pr(Y_{2i} = 1 | Y_{1i} = 0) = 0$ and $Pr(Y_{2i} = 1 | Y_{1i} = 1) = Pr(Y_{2i} = 1)$.²⁷ With the normality assumption,

$$Pr(Y_{2i} = 1) = \Phi(x_{1i}\beta_1)\Phi(x_{2i}\beta_2).$$

The likelihood function is then

$$(4) \quad \mathcal{L} = \prod_i [\Phi(x_{1i}\beta_1)\Phi(x_{2i}\beta_2)]^{Y_{2i}} [1 - \Phi(x_{1i}\beta_1)\Phi(x_{2i}\beta_2)]^{1-Y_{2i}}.$$

²⁷Given the presence of false negatives, this assumption is required for the detection model to be identified. In our application, false positives can arise if there are errors in matching NAMCS indications to the PDR data. We hired a professional medical records coder to do the matching to minimize this possibility. The other source of false positives arises when a patient is actually prescribed a drug off-label for one observed indication but the drug is approved for another of the patient's observed indications. We believe the frequency of such events is rare, and would lead to a conservative estimate of off-label use.

As Feinstein [1990] points out, certain conditions must be satisfied to identify the parameters of the model via maximum likelihood. In particular, note that when $\Phi(x_{1i}\beta_1)\Phi(x_{2i}\beta_2)$ is high, it is unclear whether this is due to a high rate of on-label prescribing, or a high rate of detection. For example, if x_{1i} and x_{2i} include all the same variables, then nothing differentially shifts the probability of on-label prescribing and the probability of detection, so β_1 and β_2 are not separately identified. But if there are some variables in the group determining on-label prescribing but not in the group determining detection (or vice-versa), then each probability will vary significantly conditional on a fixed value for the other probability and the model is identified. Ideally, we would include a variable in x_{2i} that pushed the probability of detection to one, and identified the rate of on-label prescribing in a simple discrete choice model. Without this variable, the parametric assumptions (i.e., normality and linear indices) are required for identification. In Section VI, we discuss how close we get to this ideal setting.

The common elements in x_{1i} and x_{2i} move both the probability of on-label use and the probability of detection. For this group of elements, we include patient characteristics such as age, gender and race. For example, a patient's age may lead physicians to take a more cautious or aggressive approach in prescribing off-label. Also, certain types of patients may be more likely to have a chronic indication (e.g., African-Americans have a higher incidence of diabetes), which decreases the probability of observing the relevant indication for the visit, lowering the probability of detection. Those variables only in x_{1i} move the probability of on-label use but do not affect the probability of detection. These include the number of FDA-approved drugs for the visit-specific indications, plus indicators for physician specialty, categories of indications (e.g., mental disorders), the patient's insurance status and census region and year. Those variables only in x_{2i} move the probability of detection but do not affect the probability of prescribing off-label. These include the number of prescriptions written on the visit, plus indicators for whether there were two or three indications and for the maximum number of prescriptions recordable on the NAMCS survey form.

Thus our strategy for separately identifying β_1 and β_2 relies largely on exploiting the limitations, and changes in those limitations, of the NAMCS form during our sample period, along with some visit-specific information. Recalling Table IV, we expect an increase in the maximum allowable number of prescriptions to decrease our chances of detecting on-label use, but not to affect the physician's actual decision to prescribe off-label. In Section VI, we show that our results are robust even if the set of variables that are excluded from the detection equation is varied substantially.

With estimates $\hat{\beta}_1$ and $\hat{\beta}_2$, we recover the rate of off-label prescribing as the complement of the rate of on-label prescribing, $1 - \Phi(x_{1i}\hat{\beta}_1)$. This can be calculated at the individual prescription or population level. However, as

TABLE V
DCE MODEL FOR DETECTION OF ON-LABEL PRESCRIBING

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Exclusive</i>						
# RXs written	-0.12*** (-99.61)	-0.12*** (-100.05)	-0.12*** (-98.33)	-0.13*** (-100.36)	-0.13*** (-100.36)	-0.13*** (-100.36)
2 Indications	0.016** (2.52)	0.015** (2.47)	0.016*** (2.60)	0.024*** (3.83)	0.024*** (3.85)	0.024*** (3.84)
3 Indications	0.068*** (12.34)	0.067*** (12.27)	0.068*** (12.26)	0.048*** (8.54)	0.048*** (8.59)	0.048*** (8.58)
Form allows 6 RX	-0.078*** (-9.76)	-0.074*** (-9.64)	-0.075*** (-9.34)	-0.069*** (-8.34)	-0.068*** (-8.32)	-0.068*** (-8.32)
Form allows 8 RX	-0.072*** (-12.65)	-0.065*** (-11.81)	-0.075*** (-13.00)	-0.044*** (-7.70)	-0.045*** (-7.81)	-0.044*** (-7.78)
<i>Common</i>						
Patient age	-0.0032*** (-27.15)	-0.0032*** (-27.07)	-0.0032*** (-27.07)	-0.0035*** (-26.00)	-0.0035*** (-26.05)	-0.0035*** (-26.06)
Female	-0.046*** (-9.25)	-0.046*** (-9.20)	-0.047*** (-9.30)	-0.021*** (-4.06)	-0.021*** (-4.06)	-0.021*** (-4.05)
Hispanic	0.028*** (3.04)	0.027*** (2.91)	0.029*** (3.15)	0.0043 (0.45)	0.0041 (0.43)	0.0040 (0.42)
African-American	0.13*** (15.01)	0.12*** (14.89)	0.13*** (14.96)	0.099*** (11.56)	0.099*** (11.52)	0.099*** (11.54)
Other (non-White)	-0.049*** (-5.42)	-0.052*** (-5.74)	-0.049*** (-5.34)	-0.051*** (-5.50)	-0.051*** (-5.50)	-0.051*** (-5.48)
Established patient	0.12*** (15.47)	0.12*** (15.41)	0.12*** (15.24)	0.086*** (10.28)	0.085*** (10.27)	0.085*** (10.27)
<i>Additional Controls</i>						
Detection Model	No	No	No	Yes	Yes	Yes
MD-Specialty	No	No	No	Yes	Yes	Yes
On-Label RX Model	No	No	No	Yes	Yes	Yes
Drug Class	Yes	Yes	Yes	Yes	Yes	Yes
ICD-9 Indications	Yes	Yes	Yes	Yes	Yes	Yes
MD-Specialty	Yes	Yes	No	No	No	No
Time-Squared	No	Yes	No	No	No	No
Year	Yes	No	No	No	No	No
Year/MD-Specialty	No	No	Yes	Yes	Yes	Yes
Observations	545,102	545,102	545,102	545,102	545,102	545,102

Notes: These estimates are for the model of detection. The top panel of the table reports estimates for variables exclusive to the model of detection, while the middle panel reports estimates for variables that are included in both models. The bottom panel indicates which controls are included in the specification of the two models. All additional controls are indicator variables except for the quadratic time trend. Numbers in parentheses are t-statistics. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

noted by Feinstein [1990], even when valid exclusion restrictions give consistent estimates of $\Phi(x_{1i}\hat{\beta}_1)$, the coefficient estimates on those variables common to x_{1i} and x_{2i} may not be precisely estimated if variables excluded from the detection linear index do not vary the probability of detection enough. For this reason, we interpret these coefficients with caution, despite the finding that our exclusion restrictions provide a strong source of identifying variation.

VI. RESULTS

It is simplest to discuss our estimates as two separate models, one describing detection and the other physician behavior. Tables V and VI present the coefficient estimates for selected variables in the model of detection and physician behavior, respectively. The top panel of each table contains the estimates for variables that are exclusive to each model, while the middle panel of the respective tables presents estimates for variables common to both models. Because identification of the coefficients for variables common to both models is driven in part by the parametric structure of the model, we are hesitant to draw strong conclusions about them. The bottom panels of Tables V and VI describe the controls and exclusion restrictions for each model.

Tables V and VI each have six columns. Our baseline specifications without controls for regulatory interventions are in the first four columns of the respective tables. Column 1 presents the results from a parsimonious specification, restricting temporal trends to be the same for all physician specialties in the model of physician behavior. Column 2 presents results from a specification that is nearly identical to the first, the only difference being that the model of physician behavior includes a quadratic polynomial to allow for temporal trends rather than year indicators.²⁸ Column 3 presents estimates from a more general model of physician behavior that allows for differential trends by including a full set of interactions of year and physician specialty indicators. The estimates in Column 4 are from a specification that is similar to Column 3, except that the model of detection also includes physician-specialty indicators. Since the estimates in each column of Tables V and VI are nearly identical, we emphasize the results from Column 3.

It is useful to first consider the results for the detection model. The coefficient estimates in Table V for the detection model are consistent with our

²⁸The coefficients on the year indicators and form-change indicators are separately identified, since the form-change indicators represent only a subset of the year indicators. That is, the rate of detection does not vary in years in which the survey is unchanged, but the rate of off-label use may vary. This specification with a quadratic polynomial is included to demonstrate this point.

TABLE VI
DCE MODEL FOR TRUE ON-LABEL PRESCRIBING

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Exclusive</i>						
# drugs approved	0.046*** (64.29)	0.046*** (64.34)	0.045*** (67.35)	0.034*** (83.01)	0.034*** (82.93)	0.034*** (82.96)
Medicare	-0.082*** (-4.71)	-0.075*** (-4.34)	-0.083*** (-4.74)	-0.084*** (-4.87)	-0.083*** (-4.83)	-0.084*** (-4.85)
Medicaid	-0.144*** (-7.03)	-0.13*** (-6.71)	-0.14*** (-7.02)	-0.13*** (-6.21)	-0.13*** (-6.17)	-0.13*** (-6.19)
Private insurance	-0.077*** (-5.51)	-0.073*** (-5.23)	-0.082*** (-5.82)	-0.077*** (-5.43)	-0.077*** (-5.43)	-0.078*** (-5.47)
Other insurance	-0.088*** (-3.36)	-0.068*** (-2.62)	-0.11*** (-4.02)	-0.069*** (-2.60)	-0.068*** (-2.59)	-0.069*** (-2.60)
Drug DOJ Settlement					-0.15*** (-4.98)	-0.16*** (-5.19)
After aware of DOJ					0.26*** (3.03)	0.16*** (4.48)
Years aware of DOJ						
<i>Common</i>						
Patient age	-0.0029*** (-8.35)	-0.0029*** (-8.39)	-0.0030*** (-8.65)	-0.0015*** (-4.42)	-0.0015*** (-4.44)	-0.0015*** (-4.38)
Female	-0.072*** (-6.33)	-0.072*** (-6.36)	-0.069*** (-6.08)	-0.052*** (-4.66)	-0.052*** (-4.68)	-0.052*** (-4.72)
Hispanic	-0.042** (-2.02)	-0.042** (-2.02)	-0.027 (-1.31)	-0.0034 (-0.17)	-0.0031 (-0.15)	-0.0028 (-0.14)
African-American	-0.10*** (-5.40)	-0.10*** (-5.47)	-0.095*** (-4.92)	-0.068*** (-3.56)	-0.067*** (-3.50)	-0.067*** (-3.53)
Other (non-White)	-0.015 (-0.71)	-0.020 (-0.92)	0.017 (0.79)	0.025 (1.21)	0.026 (1.23)	0.025 (1.20)
Established patient	-0.033* (-1.95)	-0.031* (-1.80)	-0.044*** (-2.60)	-0.025 (-1.50)	-0.025 (-1.48)	-0.024 (-1.47)
<i>Additional Controls</i>						
Detection Model						
MD-Specialty	No	No	No	Yes	Yes	Yes
On-Label RX Model						
Drug Class	Yes	Yes	Yes	Yes	Yes	Yes

	(1)	(2)	(3)	(4)	(5)	(6)
ICD-9 Indications	Yes	Yes	Yes	Yes	Yes	Yes
MD-Specialty	Yes	Yes	No	No	No	No
Time-Squared	No	Yes	No	No	No	No
Year	Yes	No	No	No	No	No
Year/MD-Specialty	No	No	Yes	Yes	Yes	Yes
Observations	545,102	545,102	545,102	545,102	545,102	545,102

Notes: These estimates are for the model of physician on-label prescribing. The top panel of the table reports estimates for variables exclusive to the model of physician behavior, while the middle panel reports estimates for variables that are included in both models. The bottom panel indicates which controls are included in the specification of the two models. All additional controls are indicator variables except for the quadratic time trend. Numbers in parentheses are t-statistics. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

expectations. Both the number of prescriptions for a visit and the form changes to the NAMCS decrease the probability of detecting on-label use, and each coefficient is statistically significant at the 1% level. The marginal effects for the NAMCS-form changes are also quite large, demonstrating that they represent a strong source of identification. The mean probability of detection decreased by 4.5% following the first form change, and a further 9.8% following the second change, resulting in an overall reduction of 14.3%. Further, the indicators for whether there are two or three indications for a visit are positive and statistically significant, such that for a fixed number of prescriptions an increase in the number of indications increases our rate of detection. The results in Table V are similar in all six columns.²⁹ Collectively, the precise and intuitive results in Table V give us confidence that we successfully identify the rate of detection.

Moving to the model of physician behavior, Table VI reports how the rate of actual on-label prescribing varies with the number of FDA-approved alternatives for the physician to treat a patient. We find that an increase in available alternatives significantly increases the probability that a physician prescribes on-label. The effect is also economically significant. Using the estimates from Column 3, increasing the observed number of on-label alternatives by 10% (about 7.2) decreases the probability of prescribing off-label by 4.5 percentage points, from 33.5% to 29.0%. This corresponds to a 13.1% reduction in the probability of off-label prescribing. This result is similar across all six columns of Table VI. This is, at a minimum, consistent with welfare-enhancing behavior on the part of physicians. In going beyond FDA-approved alternatives to find a good match between a patient and a drug, physicians are more likely to prescribe off-label when their choices are more limited.

Table VI also reports the estimates of the effect of insurance status on the physician's decision to prescribe on-label. Relative to the base case of a patient with no insurance (the omitted dummy variable), patients with any type of insurance are significantly more likely to be prescribed a drug off-label. We estimate marginal effects for insurance status by assuming a particular insurance status for every patient, calculating the probability of on-label prescribing, and comparing that to the probability of on-label prescribing when all patients have a different insurance status. If no patient has insurance, our estimates from Column 3 predict the proportion of prescriptions written off-label as 32.3%. If all patients have Medicaid, we estimate the proportion as 34.3%. This corresponds to a 5.9% increase in the probability that a prescription is written off-label for patients on Medicaid relative to no insurance. In comparing patients with Medicaid to those with private insurance, those on Medicaid are 2.3% more likely to be

²⁹The results are unchanged if the insurance indicators are added to the specification of the model of detection from Column 4, or if the ICD-9 indications are removed from the On-Label RX Model.

prescribed a drug off-label. Results are qualitatively similar when comparing to patients with Medicare.

The result that better insurance coverage leads to more off-label use, narrowly construed, is consistent with two hypotheses in our theoretical model. First, off-label options are more frequently superior, and insurance induces more switching to those superior off-label options. This seems unlikely, however, as off-label prescribing is less common even for patients with insurance. Second, if on-label options are more frequently superior, then it must be the case that relatively few patients without insurance are prescribed off-label when the off-label option is inferior. That is, in cases where there is a choice between a superior-but-more-expensive on-label option and an inferior-but-cheaper off-label option (e.g., Lucentis vs. Avastin), physicians prescribe the inferior option relatively infrequently.³⁰

Using the estimates of β_1 and β_2 from Column 3 of Tables V and VI, we construct implied estimates of the probability of detection and off-label use. Table VII gives the mean of the predicted probability of detection, $\Phi(x_{2i}\beta_2)$, for the taxonomy of prescription-indication combinations from Table IV. Over all observations, we estimate that the naive indicator is correct 50.3% of the time. The rate of detection falls significantly as the number of prescriptions increases, given any fixed number of indications. For example, on average, for visits with three indications and only one prescription, we detect on-label use 65.9% of the time. Yet this probability falls to only 29.6% when the maximum number of eight prescriptions is listed for a visit, along with three indications. This pattern holds across all three columns of Table VII. These probabilities provide the average adjustment applied to observations with certain combinations of number of prescriptions and indications.

For some prescriptions, the detection rate is over 80%, which means that we have a group of prescriptions detected to be on-label with a very high probability. The limiting case where the probability of detection is one is accomplished through the model's parametric assumptions. The limiting case where the linear index equals infinity does not exist in real-world retrospective data, but observing very high detection rates for a subset of prescriptions in our data gives us confidence in the DCE approach. A very cautious interpretation of results would treat our estimates of off-label use as tight upper bounds on the true rate since we get close to the limiting case in our data. Feinstein [1990] argues that this is how DCE models should be interpreted.³¹

³⁰Due to space considerations, we omit the remainder of coefficient estimates for the model of physician behavior: indicators for physician specialty, indication codes, and census region and year. These estimates are available from the authors upon request.

³¹For example, in Feinstein's [1991] study of tax evasion, it is necessary to assume that there is at least one auditor in the data that detects evasion with probability 1 in the absence of parametric assumptions.

TABLE VII
FORM LIMITATIONS AND DETECTION, MEAN OF $\Phi(x_{2i}\hat{\beta}_2)$

	(1)	(2)	(3)	(4)
	One Indication	Two Indications	Three Indications	Average
One prescription	0.646	0.646	0.659	0.648
Two prescriptions	0.596	0.600	0.615	0.601
Three prescriptions	0.546	0.548	0.566	0.552
Four prescriptions	0.491	0.494	0.513	0.500
Five prescriptions	0.437	0.440	0.463	0.449
Six prescriptions	0.376	0.383	0.408	0.394
Seven prescriptions	0.312	0.320	0.343	0.330
Eight prescriptions	0.265	0.271	0.296	0.284
Average	0.535	0.507	0.465	0.503
Observations	209,986	155,231	183,875	549,092

Notes: These estimates reflect the mean probability of detection, $\Phi(x_{2i}\hat{\beta}_2)$, by number of prescriptions and indications recorded by the physician on the NAMCS survey form. The estimates of $\hat{\beta}_2$ are drawn from column (3) of Table IV.

Figure 3 plots yearly rates of off-label use, calculated using the estimates from Column 3 of Table VI, along with a fitted polynomial. The rate of off-label prescribing fluctuates between around 29.9% and 31.4% between 1993 and 2002, and then rises consistently from 2003 to 2008. The predictions given by the estimates from the other columns of Table VI are nearly identical, never differing by more than 1% in any year. Overall, off-label use rises by nearly a third, from 29.9% in 1993 to 38.3% in 2008. Interestingly, this detected trend in off-label use also leads, by just a few years, the trend in off-label settlements, which accelerated from 2008 to 2011. Since our results cover essentially all drugs listed in the PDR, this gives us further confidence in our results over the 15-year period we study.

Figure 4(a) and (b) summarize the trends in off-label use by drug class; drug classes with the highest percentage increase in off-label use are in Figure 4(a), while drug classes with the lowest percentage increase in off-label use are in Figure 4(b).³² With the exception of a few drug classes (coagulant, gastrointestinal, and psychological), each ends our sample period with a higher rate of off-label use than it begins. Additionally, it is uniformly true that those classes with the lowest rates of off-label use in 1993

³²Again we focus on the predictions from Column 3 of Table VI, since the other specifications yield nearly identical results.

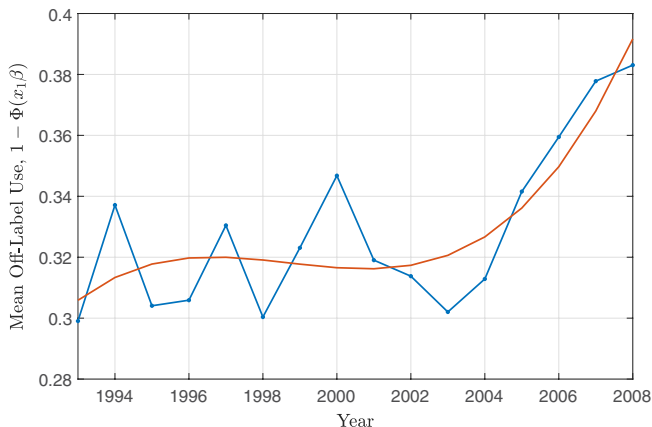


Figure 3
Trend in Off-label Use

Notes: Estimates of off-label use are from maximum likelihood estimation according to Equation (4), using the variables from column (3) of Tables V-VI. The fitted line is an 3rd-order polynomial.

had higher percentage increase in off-label use.³³ Thus the practice appears to be growing most quickly in drug classes where it had been less common.³⁴

We also evaluate the relationship between off-label use and government scrutiny of alleged off-label marketing. Column 5 of Tables V and VI reports model estimates when we include an indicator that equals one if a drug was part of a DOJ settlement and an indicator that equals one in the period after the firm became aware of the DOJ investigation. We find that off-label use is more likely if the prescribed drug is targeted by the DOJ for off-label marketing. To estimate the marginal effects, we set the DOJ indicator equal to one and the indicator variable for the time after the firm was aware of the DOJ scrutiny to 0 and predict off-label rates for drugs

³³In an online appendix, we also explore the possibility of classifying office visits where the patient has a chronic condition. We do this in two ways. First, we construct a dummy variable for the likely presence of a chronic condition. We classify prescriptions according to whether any of the 3-digit ICD9 codes are in a list where one of the 3, 4, or 5-digit ICD9 codes are classified as 'chronic' using an index constructed and validated by the Agency for Healthcare Research and Quality (AHRQ). If any are, then we label the prescription as being used for a chronic condition and re-estimate the models. Second, for a sub-set of years (1997-2008) the NAMCS asks the reason for the visit, including whether it was for a chronic condition. So, we also estimated a version of our models including that variable over the shorter time frame. In both cases our results change very little.

³⁴Radley *et al.* [2006] is the only other study, to our knowledge, that makes any attempt to estimate off-label use by drug class; however, we do not know the identity of the few drugs they study within each class, making the estimates impossible to compare.

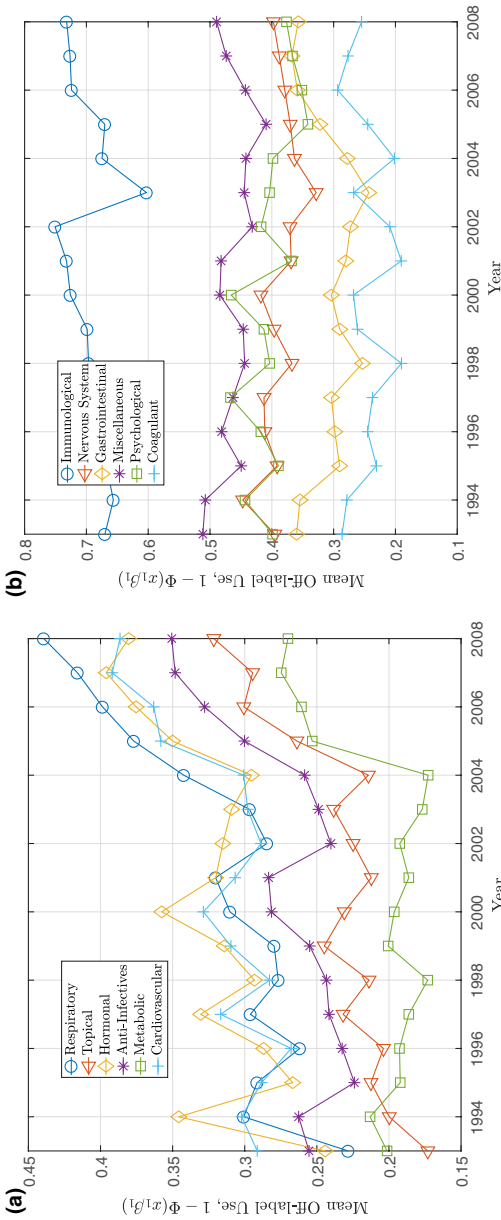


Figure 4
 Off-label Use by Drug Class
 Notes: Estimates of off-label use are from maximum likelihood estimation according to Equation (4), using the variables from column (3) of Tables V-VI.

that were under DOJ scrutiny for all time periods. We estimate that such prescriptions were 4.6% more likely to be written off-label compared to prescriptions from firms that were never scrutinized by the DOJ. We also find that DOJ investigations decrease the rate of off-label use. Using a similar approach for calculating the marginal effect, we estimate a 10.2% lower rate of off-label use *after* the manufacturer is aware that the DOJ is targeting the drug compared to before. On average, firm awareness of the scrutiny actually moves the off-label rate for the investigated firms below the off-label rate for firms that are never scrutinized. This suggests that tighter scrutiny of off-label marketing is effective at reducing the rate of off-label use. However, given our results that suggest patterns in off-label use seem consistent with enhancement of patient welfare, the social value of the interventions are unclear.

In Column 6, we include an alternative specification to explore the robustness of our results in Column 5 and provide further support for our conclusions. Specifically, rather than an indicator for the period after a firm became aware of the investigation, we include a variable that equals the number of years after the firm first became aware. If off-label marketing had an effect and that effect diminished once promotion ended, we would expect the rate of off-label (on-label) use to decrease (increase) in the number of years since the investigation began. Our results confirm this intuition, suggesting the investigations may be successful in mitigating problematic marketing of off-label uses.

VII. CONCLUSION

Off-label use is prevalent, controversial and under-studied. In this paper, we take the important step of identifying the incidence of off-label use from 1993 to 2008 and important factors that drive it. We are optimistic that our approach and these results will lead to further insights about this controversial subject and, hopefully, recommendations for policy.

For example, our finding that Medicaid patients are more likely to be prescribed off-label is important in light of recent policy developments. As of 2008, Medicaid expenditures on prescription drugs for certain low-income adults and children total \$15.2 billion dollars (USGAO [2010]: page 1); however, this program will be expanded significantly under the Patient Protection and Affordable Care Act of 2010. While our findings suggest that off-label prescribing patterns are consistent with the enhancement of patient welfare, additional economic analysis is required to ensure that the practice is desirable from a societal perspective and that tax dollars are spent efficiently. For instance, research that builds on our findings might study whether treatment outcomes from using drugs off-label justify potential costs (realized by any entity) in excess of on-label alternatives.

In addition, the DOJ continues to enforce FDA guidelines that ban promotion of off-label uses for drugs, but it does not know the extent to which this promotion induces physicians to prescribe off-label or the effect that strict enforcement has on patient welfare. Our finding that physicians tend to prescribe off-label when it is in the best interest of the patient suggests that an out-right ban on off-label use, and possibly even the current ban on promotion of off-label uses, has the *potential* to harm welfare. However, much more research is needed for clear and effective policy to be developed.

For example, little is known about how the current FDA regulatory architecture affects incentives for firms to invest in identifying new uses for existing drugs. The FDA provides three-year exclusivities to incentivize firms to incur the costs associated with seeking new indications for existing drugs, but this incentive may be negligible if the drug is already generic. Currently, firms face a difficult decision of whether to rely on clinicians to discover and promote these new uses, or incur the costs associated with adding the indication to a drug's label that will likely soon be, or already is, generic. Data from the FDA Orange Book on new chemical entities, associated patents and associated exclusivities exist can shed light on these questions.

Ideally, research on off-label use will pursue data that mitigate the detection problem further. For example, NAMCS data reflect single visits, but information from prior visits might be correlated with prescribed drugs. Also, some combinations of drugs might be common, such as high blood pressure and cholesterol medications. To push research in this direction will require considerable clinical expertise in developing classification schemes for relationships between and among indications and drugs.

REFERENCES

- Adams, C. P. and Brantner, V. V., 2006, 'Estimating the Cost of New Drug Development: Is It Really \$802 Million?' *Health Affairs*, 25, pp. 420–428.
- Aitken, M.; Berndt, E. R. and Cutler, D. M., 2009, 'Prescription Drug Spending Trends in the United States: Looking beyond the Turning Point,' *Health Affairs* 28, pp. 151–160.
- Avenel, S.; Bomkratz, A.; Dassieu, G.; Janaud, J. C. and Danan, C., 2000, 'The Incidence of Prescriptions without Marketing Product License in a Neonatal Intensive Care Unit,' *Pediatrics*, 7, pp. 143–147.
- Bradford, W.; Kleit, A.; Krousel-Wood, M. and Re, R., 2001, 'Testing Efficacy with Detection Controlled Estimation: An Application to Telemedicine,' *Health Economics*, 10, pp. 553–564.
- Cuzzolin, L.; Zaccaron, A. and Fanos, V., 2003, 'Unlicensed and Off-Label Uses of Drugs in Paediatrics: A Review of the Literature,' *Fundamental & Clinical Pharmacology*, 17, pp. 125–131.
- DiMasi, J., 2001, 'Risks in New Drug Development: Approval Success Rates for Investigational Drugs,' *Clinical Pharmacology and Therapeutics*, 69(5), pp. 297–307.

- DiMasi, J.; Hansen, R. and Grabowski, H., 2003, 'The Price of Innovation: New Estimates of Drug Development Costs,' *Journal of Health Economics*, 22, pp. 151–185.
- Dranove, D., 1989, 'Medicaid Drug Formulary Restrictions,' *Journal of Law and Economics*, 32, pp. 143–162.
- Evans, D., 2009, 'Pfizer Broke the Law by Promoting Drugs for Unapproved Uses,' November 9, <http://www.bloomberg.com/apps/news?pid=200g7000/85id=a4yvInYxCGoA.Nov.g>
- Feinstein, J., 1989, 'The Safety Regulation of U.S. Nuclear Power Plants: Violations, Inspections, and Abnormal Occurrences,' *Journal of Political Economy*, 97, pp. 115–154.
- Feinstein, J., 1990, 'Detection Controlled Estimation,' *Journal of Law and Economics*, 33, pp. 233–276.
- Feinstein, J., 1991, 'An Econometric Analysis of Income Tax Evasion and Its Detection,' *RAND Journal of Economics*, 22, pp. 14–35.
- Harris, R., 1964, *The Real Voice* (Macmillan, New York, New York, U.S.A.).
- Heidelbaugh, J.; Riley, M. and Habetier, J., 2008, '10 Billing & Coding Tips to Boost Your Reimbursement,' *Journal of Family Practice*, 57, pp. 724–730.
- Helland, E., 1998, 'The Enforcement of Pollution Control Laws: Inspections, Violations, and Self-Reporting,' *Review of Economics and Statistics*, 80, pp. 141–153.
- Huskamp, H. A., 2003, 'Managing Psychotropic Drug Costs: Will Formularies Work?' *Health Affairs*, 22(5), pp. 84–96.
- Iizuka, T., 2004, 'Prescribing Trends in Psychotropic Medications,' *Journal of Industrial Economics*, 52, pp. 349–379.
- Klein, D. and Tabarrok, A., 2004, 'Who Certifies Off-Label?' *Regulation*, pp. 60–63.
- Kesselheim, A.; Mello, M. and Studdert, D., 2011, 'Strategies and Practices in Off-Label Marketing of Pharmaceuticals: A Retrospective Analysis of Whistleblower Complaints,' *PLoS Medicine*, 8(4), pp. 1–9.
- Kolata, G., 1997, 'How Fen-Phen, A Diet "Miracle," Rose and Fell,' *The New York Times*, September 23.
- Kleit, A. and Ruiz, J., 2003, 'False Positive Mammograms and Detection Controlled Estimation,' *Health Services Research*, 38, pp. 1207–1228.
- Lin, H.; Phan, K. and Lin, S., 2006, 'Trends in Off-Label β -Blocker Use: A Secondary Data Analysis,' *Clinical Therapeutics*, 28, pp. 1736–1746.
- Mack, A., 2003, 'Examination of the Evidence for Off-Label Use of Gabapentin,' *Journal of Managed Care Pharmacy*, 9, pp. 559–568.
- McIntyre, J.; Conroy, S.; Avery, A.; Corns, H. and Choonara, I., 2000, 'Unlicensed and Off Label Prescribing of Drugs in General Practice,' *Archives of Disease in Childhood*, 83, pp. 498–501.
- Molitor, D., 2012, 'Physician Behavior and Technology Diffusion in Health Care,' MIT doctoral dissertation (Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.).
- Oliver, T. R.; Lee, P. R. and Lipton, H. L., 2004, 'A Political History of Medicare and Prescription Drug Coverage,' *The Milbank Quarterly*, 82, pp. 283–354.
- O'Reilly, J. and Dalal, A., 2003, 'Off-Label or Out of Bounds? Prescriber and Marketer Liability for Unapproved Uses of FDA-Approved Drugs,' *Annals of Health L*, 12, pp. 295–300.
- Peltzman, S., 1973, 'An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments,' *Journal of Political Economy*, 81, pp. 1049–1091.
- Perkins, K. A.; Conklin, C. A. and Levine, M. D., 2008, *Cognitive Behavior Therapy for Smoking Cessation: A Practical Guidebook to the Most Effective Treatments* (Routledge, New York, New York, U.S.A.).

- Pincus, H. A.; Tanielian, T. L.; Marcus, S. C.; Olfson, M.; Thompson, J. and Zito, J. M., 1998, 'Prescribing Trends in Psychotropic Medications,' *Journal of the American Medical Association*, 279, pp. 526–531.
- Radley, D.; Finkelstein, S. and Stafford, R., 2006, 'Off-label Prescriptions among Office-Based Physicians,' *Archives of Internal Medicine*, 166, pp. 1021–1026.
- Rowland, D. (Executive Director, Kaiser Commission on Medicaid and the Uninsured), 2001, 'Prescription Drug Coverage for the Medicare Population' - Testimony before the Subcommittee on Health, Committee on Energy and Commerce, The United States House of Representatives, February 15.
- Salbu, S. R., 1999, 'Off-Label Use, Prescription, and Marketing of FDA-Approved Drugs: An Assessment of Legislative and Regulatory Policy,' *Florida Law Review*, 51, pp. 18.
- Stafford, R., 2008, 'Regulating Off-Label Drug Use - Rethinking the Role of the FDA,' *New England Journal of Medicine*, 358, pp. 1427–1429.
- Temin, P., 1979, 'The Origin of Compulsory Drug Prescriptions,' *Journal of Law and Economics*, 22, pp. 91–105.
- Thomas, C. P.; Conrad, P.; Casler, R. and Goodman, E., 2006, 'Trends in the Use of Psychotropic Medications Among Adolescents, 1994 to 2001,' *Psychiatric Services*.
- United States General Accounting Office, 1991, 'Off-Label Drugs: Reimbursement Policies Constrain Physicians in Their Choice of Cancer Therapies,' September 27.
- United States General Accounting Office, 1996, 'Prescription Drugs: Implications of Drug Labeling and Off-Label Use,' September 12.
- United States General Accounting Office, 2010, 'Medicaid Outpatient Prescription Drugs,' December 15.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site: <http://wileyonlinelibrary.com/journal/joie> or via The Journal's website, <http://www.jindec.org>