Should We Prevent Off-Label Drug Prescriptions?

Empirical Evidence from France

Tuba Tunçel*

January 2021†

Abstract

After a drug obtains marketing authorisation, the usage depends on the regulation of off-label prescriptions for unapproved indications. As a consequence, drug pricing and firms' revenues depend on off-label use which in the long run play an important role in innovation incentives. Off label prescriptions represent more than 20% of drug spending and treatment choices in many developed economies. We investigate the impact of regulation of off-label prescriptions on physicians' behavior, patients' health, treatment costs and pharmaceutical firms' pricing with a structural model of demand and supply. Exploiting rich panel data on physicians' activities and office visits in France over a nine-year period, we use a model of prescription choice and health outcomes with unobserved patient-level heterogeneity. We identify the demand for on-label and off-label drugs and the effect of prescription choice on health outcomes. On the supply side, we use a Nash-in-Nash bargaining model between the government and the pharmaceutical companies that allows the partial identification of the marginal costs of drugs. Counterfactual simulations show that when we remove off-label drugs from the choice set of physicians, substitution to on-label drugs at constant prices would lead to an increase in the expenditure on prescription drugs. If we allow bargaining adjustment on drug prices under a ban on off-label prescriptions, the ban would further increase the treatment cost without leading to an improvement in health outcomes.

Keywords: Physician Behavior, Prescription Drugs, Drug Efficacy, Off-Label Drugs, Bargaining.

JEL Codes: I10, D12, C25

*HEC Montreal, email: tuba.tuncel@gmail.com

†I would like to thank Pierre Dubois for his support, advice and for the data access to the CEGEDIM and IMS Health databases. I also thank Gautam Gowrisankaran, Bruno Jullien, Thierry Magnac, Philippe Choné, Mathias Reynaert, James Hammitt, Michael Dickstein, Isis Durrmeyer, Elaine Kelly as well as participants in many seminars and conferences. I also thank the NBER-IFS International Network on the Value of Health Research for pilot funding and HEC Montreal for financial support.

1 Introduction

While marketing authorizations are given based on evidence of safety and efficacy for the specific requested indications, hundreds of millions of dollars are spent every year on off-label prescriptions of drugs in the US, Canada and most European countries. The use of prescription drugs for an indication (e.g., a disease or a symptom) other than the approved indications is called off-label use[1].The famous example of Avastin, a drug approved for cancer treatment but used widely for the treatment of age-related macular degeneration (AMD), shows the importance of off-label use and its regulation. Even though Avastin is not officially approved in Europe for AMD, there is scientific evidence that it is effective for the treatment of AMD. As the average price of Avastin in Europe is 40 euros per injection, whereas the approved drug for AMD treatment (Lucentis) is approximately 900 euros per injection, there are large incentives for off-label use of Avastin.

Another example is the off-label use of acetylsalicylic acid, famously known as aspirin, to prevent secondary myocardial infarctions (heart attacks). Preliminary evidence that aspirin could lower the risk of a second heart attack began to emerge in the 1960s and 1970s. However, aspirin was already a cheap generic drug back then, as its patent had expired in 1917 and no company had an incentive to conduct clinical trials to prove that aspirin prevents heart attacks. Thanks to government financing of such studies, aspirin was approved for secondary myocardial infarction only in 1985 (Weisman and Healy (1987)). Until approval, aspirin was used off-label as a heart attack preventive for decades[2]. More recent examples for off-label use are the drugs that are being used in treatment of coronavirus all around the world[3].

In the pharmaceutical market which accounts for 12% of total health spending in the U.S. (roughly 2.1% of U.S. GDP) and 14% of total health spending in France (1.6% of French GDP) (OECD (2017)), off-label use is very common. Using nationally representative data, Radley et al. (2006) finds that, among the 160 most prescribed drugs in the US, off-label prescriptions account for approximately 21% of overall use. Papers on off-label use in Europe typically study this practice within narrowly defined clinical populations such as pediatric patients (Chalumeau et al. (2000), Bücheler et al. (2002)) or inpatients in a single hospital on a single day (Martin-Latry et al. (2007)). While there is extensive literature on the determinants of physicians' treatment decisions, studies of physicians' choices between approved and off-label alternatives

[1]The term "off-label use" can also apply to the use of a drug in a patient population (e.g., pediatric), or in a dosage form that has not been approved. However, in this study, off-label use refers to the use of a drug to treat an indication for which the drug has not received approval. To avoid conflicts with patient population-based off-label use, we focus on the sample of adult patients.

[2]Source of this example: Article titled "Drug safety: Off-label drug use" on http://consumerhealthchoices.org/report/off-label-drug-use/.

[3]As of June 5, 2020, which is the date this note is written, there are no drugs approved for the treatment of coronavirus. The FDA issued only an emergency use authorization for Remdesivir.

are sparse. Bradford et al. (2018) provide an important first step documenting off-label prescription rates and identifying some determinants of such decisions. Shapiro (2018) investigates the impact of detailing on physicians' prescriptions of on-label versus off-label use.

Our paper is the first study that identifies the monetary costs and treatment outcomes of off-label prescriptions relative to on-label alternatives. We build a structural model of demand and supply to answer the policy relevant question of whether we should prevent off-label drug prescriptions. On the demand side, we model physicians' prescription decisions across on-label and off-label alternatives and study the impact of prescription choice on health outcomes. One important feature of the demand model is that we jointly estimate the prescription choice and health outcome taking into account selection into treatment based on unobserved patient health state which can bias the estimates for treatment impact of drugs if not controlled for. On the supply side, we use a Nash-in-Nash bargaining model between the pharmaceutical companies and the government in which they negotiate to determine drug prices. In the supply model we allow the government to care for both the consumer surplus and the cost of treatment. In health care systems with generous coverage like the one in France, it is possible that demand is price insensitive. The objective to contain total cost of treatment implies that the government still wishes to lower prices in negotiation with pharmaceutical companies, even if demand is price insensitive. Using the estimates of the structural model, we, then, conduct counterfactual analysis to investigate how banning off-label prescriptions would impact the equilibrium of drug prices, the monetary cost of treatment and the health outcomes.

In a different context, some recent literature has studied the effects of restricting supply in prescription drug plan choices in the U.S., as done by some insurance companies. Lucarelli et al. (2012) estimate an equilibrium model of the Medicare Part D market to study the welfare impacts of reducing consumers' choices because of public concerns that there are "too many" plans to choose from. Decarolis et al. (2020) study the interaction between insurer behavior and public subsidies within Medicare Part D Prescription Drug Plan markets using a similar structural model of supply and demand. In those cases, supply-side models are slightly different than government bargaining over the prices of drugs, which occurs in France.

Regulation of off-label use differs across countries; some countries permit off-label use, some partially restrict it, and still others restrict it through the limited reimbursement of off-label use. Regulating off-label use is a difficult task mainly because of the unknown costs and benefits involved. Off-label use may provide important benefits for some patients if their individual needs require the use of an off-label drug, as in cases where approved treatments have failed (Stafford (2008)). In addition, off-label use may provide financial

benefits, as in the use of Avastin to treat AMD. However, if off-label use is not effective in treatment, it would lead to wasteful spending. Additionally, even if off label drugs are as effective as indicated ones for the treatment of a disease, allowing wider choice of treatments can affect expenses because prices may be different.

We use data from France during the years 2000 to 2008, a period in which there was no restriction on off-label prescriptions. Physicians were free to choose among drugs, regardless of their label status and they do not have monetary incentives to prescribe one drug vs. another. The data period and the institutional setting is ideal to study physician behavior in terms of off-label use. We analyze the share of off-label prescriptions conditional on the diagnosis, that is, the propensity of the physician to prescribe an off-label treatment conditional on the diagnosed disease, namely, depression. The existing medical literature on off-label use provides information on which drugs have the highest number of off-label prescriptions. Our analysis is based on depression treatment for which off-label use is indeed prevalent among French general practitioners and we find that 21% of the drugs prescribed for depression treatment are off-label drugs.

This unique French dataset provides longitudinal information on a representative sample of physicians and all their patients for a period of nine years. The longitudinal dimension of the data on office visits allows us to follow patients' visits to their physicians and thus observe their treatment outcomes. Hence, we can investigate whether treatment outcomes are different for patients treated with approved drugs than for those treated with off-label drugs.

We uniquely contribute to identifying the treatment costs and benefits of off-label use relative to on-label alternatives and shed light on how off-label prescriptions versus approved alternatives affect patients' health. Using a model of prescription behavior in which patient-level unobserved heterogeneity is allowed to be correlated with treatment outcomes, we identify the demand for on-label and off-label drugs in depression treatment. We separately identify the impact of treatment choice and the impact of patients' unobserved health state on treatment outcomes using detailing (advertising) expenditures that affect physicians' prescription decisions and that can credibly be excluded from the health outcome equation, conditional on the treatment chosen. The results show that patients' unobserved health state impacts both the treatment choice and the treatment outcome and that off-label drugs are not worse than the on-label alternatives in terms of the probability of recovery. Consistent with the medical literature, our results show that treatment impact of drugs is heterogenous across patients after controlling for both observables and unobserved health state.

Current regulations question policies that would restrict or fully prevent off-label prescriptions. In France, the current system in effect since 2011 aims at strictly regulating off-label prescriptions with "Temporary Recommendations for Use (TRU)", while in the U.S., the formulary drug lists of health insurance companies contribute to limiting off-label prescriptions. To shed light on the best regulations of off-label use, we evaluate the counterfactual effects of a ban on off-label prescriptions.

Using the structural model parameter estimates, we first evaluate the impact of this ban on choice probabilities, treatment outcomes, and the costs of prescription drugs, keeping the prices of drugs fixed. The counterfactual simulations show that banning off-label prescriptions would lead to an increase in the cost of treatment because of substitution to more expensive products, whereas it would not lead to an improvement in terms of health outcomes.

We, then, investigate how the ban on off-label prescriptions would impact the price negotiations between the pharmaceutical companies and the government. Using a Nash-in-Nash bargaining model (as in Crawford and Yurukoglu (2012), Grennan (2013), Gowrisankaran et al. (2015), Ho and Lee (2017)) between the government and pharmaceutical companies allows the partial identification of marginal costs and bargaining parameters thanks to the observed price equilibrium. In this bargaining model, we assume that firms' objective functions are their profits, while the government cares about a weighted sum of consumer surplus and the cost of treatment. Prices are assumed to be determined by a Nash equilibrium of bilateral Nash bargaining problems (Horn and Wolinsky (1988)) between each pair of firm and the government. We show how we can set-identify the bargaining parameters and the weights the government puts on consumer surplus versus the cost of treatment thanks to marginal cost restrictions. We restrict the marginal cost of drugs with today's prices which are smaller than drug prices during the sample period. This identification strategy narrows the interval set of the bargaining parameters which are between 0.58 and 0.79 for the firms. We can then simulate the new equilibrium outcome that would prevail had bargaining on the prices of drugs happened under a strict ban on off-label prescriptions.

The Nash bargaining equilibrium implies that the price cost margin of a drug is a function of the drug's additional value added in consumer surplus, its marginal impact on total cost of treatment and of the price elasticity of demand. The ban can impact drug prices through all these channels, through its impact on total cost, demand and consumer surplus elasticities.

A drug approved for depression is used not only in its on-label market (depression) but also for indications for which it is not approved: off-label markets. Similarly, a drug approved for depression faces competition

from drugs approved for other indications but used off-label in depression treatment. With the ban on off-label use, the drug's market power in the on-label market will increase because the drug will stop facing competition from off-label drugs in this market. However, the drug will lose all the sales in the off-label markets in which it is used when there is no ban. Therefore, whether the aggregate demand for the drug will increase or decrease with the ban relative to without the ban is ambiguous. Hence, the impact of the ban on the negotiated price of a drug can go in any direction depending on the prevalence on different indications and on the price elasticity of demand for the drug on the on-label indication market when off-label drugs are present and when they are absent and also on its elasticity in the off-label indication markets. We thus also estimate the demand for approved drugs for depression in the off-label markets for these drugs[4].

Similarly, the ban can impact the drug prices through the channel of the price elasticity of total spending and consumer surplus of drugs. The value added of a drug in consumer surplus is the marginal surplus provided by the drug relative to the other drugs in physicians' choice set. As physicians' choice set shrinks with the ban, the marginal surplus of a drug increases and becomes less elastic. Therefore, the ban will increase the price through the channel of consumer surplus elasticity. Overall, the effect of the ban on the price of a drug will depend on whether its impact through all these channels go in the same direction or on which one dominates if they vary in opposite directions.

We estimate the impact of the ban on prices given the values of the bargaining parameters and weights in the identified set. The results show that prices of approved drugs for depression treatment increase when off-label prescriptions are banned. The price increase depends on the combination of the bargaining parameters and weights and ranges from 0.4% to 34% across drugs. Therefore, the ban on off-label drug prescriptions would increase treatment expenditures even more under the counterfactual bargaining equilibrium prices. Keeping prices constant, physicians substitute towards on-label drugs, and because on-label drugs are more expensive on average, this substitution effect leads to an 3.8% increase in prescription expenses when off-label prescriptions are banned. In equilibrium when prices are negotiated under the ban, the prescription expenses increase by 24.3% due to both the price and substitution effects. As a result, banning off-label prescriptions would increase the cost of treatment, not only if drug prices are assumed to be the same as those under the no-ban benchmark but also if prices are negotiated under the ban. The counterfactual results also show that the ban does not provide any improvement in patients' health outcomes.

[4]The list of the off-label markets for approved drugs for depression treatment is provided later on in the text.

The paper proceeds as follows: Section 2 summarizes the relevant information on the institutions and regulations in the French health care system, and Section 3 describes the data and provides descriptive statistics. Section 4 presents the structural models of demand and supply. Section 5 describes the econometric identification and empirical estimates. Section 6 presents the counterfactual analysis, and section 7 concludes.

2 Institutional Background

2.1 Drug Approval Process and Regulations in France

Both the French drug-regulatory agency and the European Medicines Agency (EMA) can issue marketing authorization for drugs in France. For authorization, they require substantial evidence of efficacy and safety determined by clinical trials for a given indication. In France, approximately 20% of drugs are approved by the EMA (Cohen et al. (2007)). The reimbursement of drugs depends on the evaluation of the transparency commission, which assesses each approved drug's safety, efficacy, and ease of use. The commission also takes into consideration the severity of the disease targeted by each specific drug and the availability of alternative therapies. The commission does not evaluate a drug's cost effectiveness. It ranks each drug according to a measure of actual medical benefit. The commission also compares the new drugs with the existing drugs and assigns a score for the improvement in medical benefits.

Each drug has a reimbursement level according to its rate: 100% reimbursement for irreplaceable drugs, 35% reimbursement for drugs treating disorders that are not considered serious, and 65% for all other drugs (Cohen et al. (2007)). It is important to note that, once the reimbursement level is determined and the drug is on the market, the determined reimbursement level then applies to all the prescriptions of the drug regardless of the indication the drug is prescribed for, i.e., regardless of whether the drug is prescribed for an approved or off-label indication. Patients are reimbursed through the national healthcare system according to the determined reimbursement rate. Almost all the drugs used in depression treatment have a reimbursement rate of 65%.

After the commission determines the reimbursement level of a new drug, the economic committee on health products negotiates the drug's price with manufacturers. Pricing decisions depend on the medical benefit improvement rating, prices of therapeutic alternatives, the size of the target patient population, expected sales volume, and associated budget impact (Cohen et al. (2007)). It is important to note that the medical benefit improvement score of a drug is based on the main indication(s) of the drug. Potential off-label uses of

the drug do not play a role in determining this score. Unlike in the U.S., drug samples and direct-to-consumer promotion of prescription drugs are tightly restricted in France (Gallini et al. (2013)).

Figure 2.1 shows an example of three drugs that are approved for different indications but can be used off-label for non-approved indications. In this example, only drug A is approved for depression, but drugs B and C are also used, off-label, to treat depression, while drug B is approved for alcoholism and drug C for epilepsy. Finally, drug A is also used off-label to treat alcoholism. As we will see later, banning off-label prescriptions would mean that off-label prescriptions displayed with dashed arrows in this example would not be allowed (see figure A.1 in appendix A.3).

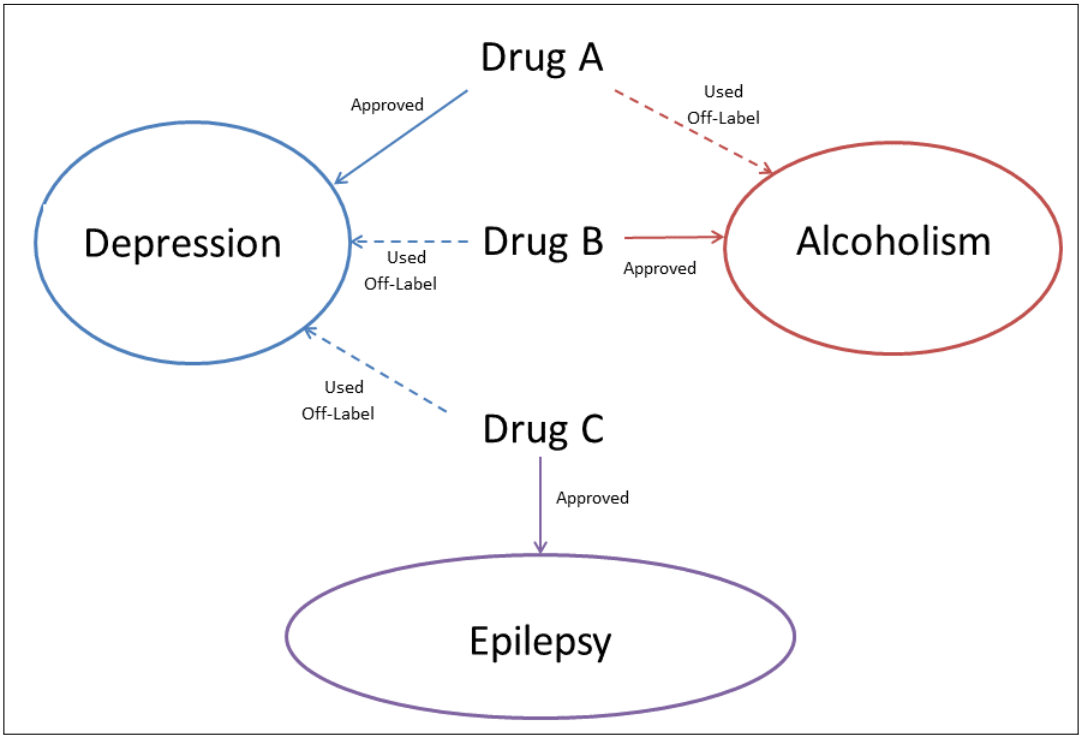


Figure 2.1: *Example of On- and Off-Label Use of Drugs*

2.2 Health Care System

Health insurance is mandatory in France, and all residents are automatically enrolled in the national insurance system with different categories depending on their occupational status. The public national health insurance reimbursement rate is almost always 65% for the drugs prescribed for depression treatment. Even though national health insurance includes different health insurance plans for different occupational groups, they are all regulated under the same statutory framework (Rodwin (2003)).

As in the case of the Italian market, discussed by Crawford and Shum (2005), this feature of the French market attenuates the agency problem, which may come into play in the case of a market with heterogeneous third-party payers, as doctors in France face a uniform incentive scheme. For instance, the heterogeneous constraints on doctors' choices induced by drug formularies in the U.S. do not come into play in the French market. Physicians do not have monetary incentives to prescribe one drug vs. another; however, detailing towards doctors by pharmaceutical companies may play a role in their prescription choices which we control for in our analysis.

3 Data and Descriptive Statistics

3.1 Data

The main data are proprietary to CEGEDIM, a global technology and services company specializing in health care, and contain the exhaustive prescriptions written by 386 general practitioners to all of their patients in France between 2000 and 2008, with patient-level anonymous IDs that allow us to follow individual patients over time. At the physician level, the dataset includes the physicians' age, gender and region of operation. At the patient level, it includes socio-demographic information (age, gender, and employment status) and information on health (chronic disease conditions and body mass index).

For each patient visit, physicians record all diagnoses and indicate the drug therapy specifically used to treat each diagnosis; they also record the exam results transmitted to them. Thus, we observe all the diagnosis-drug prescription pairs for each visit, including details such as the dosage and renewal of treatments. Note that to study off-label prescriptions, it is essential to observe the diagnosis and the drug prescribed for each diagnosis. The drug-level information is the Anatomical Therapeutic Chemical (ATC) code at the finest level, the reimbursement level, and whether the drug is generic or branded. For each visit, the patient and physician identification numbers allow us to identify unique physician-patient pairs.

Data about on-label indications of drugs were self-collected from the websites of the ministry of health and from the French regulatory agency[5]. Diagnoses are considered as on-label if they can be matched to the therapeutic indications approved by the French drug regulatory agency or the European Medicines Agency. Any diagnosis that cannot be matched to a labeled indication is considered off-label. We also use the DRUGDEX system (Thomson Micromedex, Greenwood Village, Colo), a nationally recognized pharmaceutical

[5]http://base-donnees-publique.medicaments.gouv.fr (official website of Ministry of Health) and http://www.theriaque.org (approved by French regulatory agency, HAS)

compendium in the U.S. that describes the efficacy and scientific documentation for prescription drugs. It contains readily available summaries of evidence-based information on the indications (both on-label and off-label) of drugs. As in Radley et al. (2006), each drug indication is considered to have scientific support if, according to DRUGDEX, its effectiveness has been shown in controlled trials or observed in clinical settings[6]. Combining these data allows us to identify whether a drug is on-label or off-label for a particular indication in France and also if an off-label drug has been shown to have efficacy for a particular indication in the medical literature.

Drug prices and drug formats (number of pills and mg per pill in a drug package) are obtained from the drug database of the French Social Security Health Care System ("Base de Médicaments et Informations Tarifaires" on http://www.codage.ext.cnamts.fr). The website reports package formats, reimbursement level, and drug prices from the 1990s to today at the CIP level (code that identifies the presentation of a drug format, i.e., package). "Defined Daily Dose (DDD)", which is the assumed average maintenance dose per day for a drug used for its main indication in adults is obtained from ATC/DDD Index of the World Health Organization (https://www.whocc.no/atc_ddd_index/).

Finally, we also use drug detailing data from the IMS Health Global Promotional Track for France (IMS Health - Base Global Promo Track - [2001 - 2008]), which reports monthly detailing expenditures for each drug to general practitioners in France for the period from May 2001 to December 2008.

3.2 Descriptive Statistics

Table 3.1 provides the summary statistics of the shares and the average prices of off-label and approved drugs used in depression treatment. The statistics are based on the prescriptions given at the visit during which the disease is diagnosed for the first time. In total, 21% of the drugs prescribed for depression are off-label drugs; for only 14% of them, the medical literature provides evidence that they have efficacy in depression treatment. For the remaining 86% of off-label drugs, there is no research in the medical literature investigating their efficacy in depression treatment, and hence, they are neither shown to be effective nor ineffective in depression treatment[7]. It is important to note that these are not off-label drugs "without efficacy", these are off-label

[6]The information on indications comes from sources in different languages. For the correspondence of diseases in French and English, we use the International Classification of Diseases (ICD) published by the World Health Organization in both languages.

[7]All phyisicians prescribe off-label drugs and a substantial share of them prescribe an off-label drug around 20% of the time they prescribe a drug for depression. Therefore, the average number of off-label prescriptions across all prescriptions, 21%, is not caused by some physicians very agressively prescribing off-label drugs and others prescribing very little. They all prescribe off-label drugs and very large share of them prescribe an off-label drug between 15-25% of the time they prescribe a drug for depression treatment (for distribution of share of off-label prescriptions across physicians, see the online appendix B.6).

drugs without any published research on their efficacy in depression treatment. They are called as "Other Off-Label Drugs" throughout the paper.

The average price of on-label drugs, 0.83€, is more than four times as large as the average price of off-label prescriptions, 0.17€. Table 3.1 also reports the total number of depressed patients over the sample period.

Table 3.1: *On-Label versus Off-label Prescriptions in Depression Treatment*

	Share	Average Price
On-Label Drugs	79%	0.83€
Off-Label Drugs with Efficacy	3%	0.49€
Other Off-Label Drugs	18%	0.17€
Number of Patients	37,510	

Notes: The second column shows the share of prescriptions at the first time the patient is diagnosed with depression. Price is the price for one-day treatment calculated using "Defined Daily Dose (DDD)" assigned by the World Health Organization. For each active ingredient, it is the price per mg times mg per day according to the DDD. The third column shows the average price for one-day treatment for the drug group in the first column of the corresponding row.

The treatment outcome considered in the analysis is whether the patient recovers from the disease after a certain treatment period. We allow for a certain treatment period; then, we check whether the patient is diagnosed with depression again during an observation period after the treatment period. The recovery is defined by the patient not being diagnosed with depression during this observation period after the treatment period.

Tables 3.2 and 3.3 show the percentage of patients recovering from the disease for combinations of two treatment periods, six-month and one-year, and two observation periods, six-month and one-year. Note that recovery from the disease is one of the main health outcomes the medical literature considers in depression treatment (see Kilbourne et al. (2018), Licht-Strunk et al. (2009)). Table 3.2 shows what percentage of the patients recover from the disease six months after the first time they are diagnosed with depression. The table reports treatment outcomes across all patients and also across patients depending on their prescriptions: approved drugs, off-label drugs, off-label drugs with efficacy, and other off-label drugs. There are three columns showing three different time periods. 'One-year Period' represents the one-year observation period starting six months after the first diagnosis. 'Six-month Period' represents the six-month observation period starting six months after the first diagnosis. 'Anytime' represents the entire period starting six months after the first diagnosis until the end of the sample period.

The share of patients who recover from the disease is higher among those who are prescribed off-label drugs: patients who are prescribed off-label drugs are approximately 13 percentage points more likely to

recover than those who are prescribed approved drugs. The recovery rate is slightly higher among patients who are prescribed off-label drugs with efficacy than among patients who are prescribed other off-label drugs.

Table 3.2: *Recovery Rates - Six Months after the First Diagnosis*

	No Depression Diagnosis in Six-month Period	No Depression Diagnosis in One-year Period	No Depression Diagnosis Anytime
All Patients	66%	60%	49%
Among Patients who are Prescribed:			
On-Label Drugs	63%	57%	46%
Off-Label Drugs	76%	70%	59%
Off-Label Drugs with Efficacy	76%	71%	59%
Other Off-Label Drugs	76%	70%	59%

Notes: For the 'one-year' and 'anytime' observation periods, second and third columns, the recovery rates are lower because in these cases, some of the patients have another cycle of depression (relapse cases).

Table 3.3 reports statistics on recovery after a one-year treatment period after the first diagnosis. Recovery rates after one year are approximately 5 percentage points higher than recovery rates after six months, whereas the recovery patterns among patients prescribed off-label drugs and patients prescribed on-label drugs are the same as in Table 3.2.

Table 3.3: *Recovery Rates - One Year after the First Diagnosis*

	No Depression Diagnosis in Six-month Period	No Depression Diagnosis in One-year Period	No Depression Diagnosis Anytime
All Patients	72%	65%	55%
Among Patients who are Prescribed:			
On-Label Drugs	69%	63%	52%
Off-Label Drugs	79%	74%	64%
Off-Label Drugs with Efficacy	79%	73%	64%
Other Off-Label Drugs	79%	74%	64%

Notes: See the notes in Table 3.2.

These descriptive statistics show that there are clear correlations between prescriptions and health status, which could be a combination of the treatment effects of drugs and selection into treatments by physicians. We will now develop a model allowing us to disentangle causality effects from correlated effects due to heterogeneity. This approach will allow us to identify the counterfactual effects of choices and health outcomes when off-label alternatives become unavailable to the prescriber.

4 Structural Model of Demand and Supply Including the Off-Label Use of Drugs

4.1 A Joint Model of Prescription Choice and Health Outcome

We assume that the drug prescription choice is based on a random utility model in which physician i prescribes drug d to patient j at time t to maximize some payoff function U_{ijdt} specified as follows:

$$U_{ijdt} = \alpha_d(z_i, z_j) - \beta p_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d I_j + \varepsilon_{ijdt}$$

where $\alpha_d(z_i, z_j)$ is a drug-specific effect depending on z_i, and z_j, some observed characteristics of physicians and patients such as age and gender, p_{dt} is the price, x_{dt} is the drug-specific detailing expenditures stock, I_j is an unobserved patient state affecting the propensity to recover (health outcome), which can affect drug preferences differently with coefficient λ_d. ε_{ijdt} is a deviation from the mean utility of d at t assumed to be independent of all other variables. $l(d)$ is a dummy variable indicating whether drug d is on-label or off-label so that $\gamma_{l(d)}$ allows detailing to differently impact the on-label and the off-label use of drugs.

To take into account the long-lasting effects of advertising, we build an advertising stock for each month t using all past detailing expenditures for each drug d such that[8] $x_{dt} = \sum_{\tau=-\infty}^{t} (0.75)^{t-\tau}$ (detailing expenditures)$_{d\tau}$. As health insurance coverage is quite large, it is possible that the coefficient β is small or zero but we allow the out of pocket payment of patients to possibly affect drug choices. Moreover, it could be that physicians are not completely indifferent to price because of the warnings they receive from the government to keep the cost of treatment low and/or additional payments by the health insurance system which depends on their rate of generic prescriptions. With this utility specification we consider the physician-patient pair as the decision maker which is reasonable given the features of the French healthcare market (for details see section 2.2).

The unobserved state I_j is assumed to be discrete with two types of patients: high- and low-types. A patient is of high-type such that $I_j = \overline{I}$ with probability q and of low-type such that $I_j = \underline{I}$ with probability $(1-q)$. The drug choice is denoted by $y_{ijt} \in \{1, .., D\}$ and the alternative 0, corresponding to the group of other off-label drugs, has a normalized utility $U_{ij0t} = \gamma x_{0t} + \varepsilon_{ij0t}$ (details on choice alternatives are provided in Section 5.1).

[8] We also performed some robustness checks by varying the time discounting value.

Under the assumption that ε_{ijdt} is independently and identically distributed according to a Gumbel (type I extreme value) distribution, the choice probability of drug d by physician i for patient j, conditional on the unobserved patient state, is

$$P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt},I_j\right)=\frac{\exp\left(\alpha_d\left(z_i,z_j\right)-\beta p_{dt}+\gamma_{l(d)}x_{dt}+\lambda_d I_j\right)}{\sum_{d'=0}^{D}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\gamma_{l(d')}x_{d't}+\lambda_{d'}I_j\right)} \tag{4.1}$$

The choice probability of drug d by physician i for patient j unconditional on the unobserved patient state is then

$$\begin{aligned}P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt}\right) &= q\,\frac{\exp\left(\alpha_d\left(z_i,z_j\right)-\beta p_{dt}+\gamma_{l(d)}x_{dt}+\lambda_d\overline{I}\right)}{\sum_{d'=0}^{D}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\gamma_{l(d')}x_{d't}+\lambda_{d'}\overline{I}\right)}\\ &\quad+(1-q)\,\frac{\exp\left(\alpha_d\left(z_i,z_j\right)-\beta p_{dt}+\gamma_{l(d)}x_{dt}+\lambda_d\underline{I}\right)}{\sum_{d'=0}^{D}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\gamma_{l(d')}x_{d't}+\lambda_{d'}\underline{I}\right)}\end{aligned}$$

The recovery status of patient j who has visited physician i is denoted r_{ijt} and equal to one or zero depending on whether the patient has recovered. We assume that it depends on an unobserved propensity to recover from the disease, r^*_{ijt}, such that $r_{ijt}=1_{\{r^*_{ijt}\geq 0\}}$ as

$$r^*_{ijt}=z_j'\theta+\sum_{d=1}^{D}\delta_{ijd}1_{\{y_{ijt}=d\}}+I_j+\eta_{ijt}$$

where z_j is the patient covariates and δ_{ijd} is the treatment effect of drug d, relative to the reference drug, for individual i treated by physician j. This specification allows the unobserved patient state I_j to be correlated with each patient's propensity to recover: *High(low)-type* patients are more (less) likely to recover from the disease. For example, we can consider the *high-type* patients as the mild cases and *low-type* patients as the more severe cases of depression. λ_d allows the unobserved patient state to affect drug preferences differently across drugs. Therefore, the patient state, which is observed by the physician but unobserved by the econometrician, is allowed to impact both prescription probability and recovery probability.

Assuming that η_{ijt} is independent of other variables and standard normal, the probability of recovering conditional on prescription choice y_{ijt} and the unobserved patient-state I_j is

$$P(r_{ijt}=1|z_j,I_j,y_{ijt}=d,\delta_{ijd})=\varphi\left(z_j'\theta+\sum_{d'=1}^{D}\delta_{ijd'}\mathbf{1}_{\{y_{ijt}=d\}}+I_j\right)$$

where $\varphi(.)$ is the normal cumulative distribution function.

The recovery probability unconditional on the unobserved patient-state is then

$$P(r_{ijt}=1|z_j,y_{ijt}=d,\delta_{ijd})=q\ \varphi\left(z_j'\theta+\sum_{d'=1}^{D}\delta_{ijd'}1_{\{y_{ijt}=d\}}+\overline{I}\right)+(1-q)\ \varphi\left(z_j'\theta+\sum_{d'=1}^{D}\delta_{ijd'}\mathbf{1}_{\{y_{ijt}=d\}}+\underline{I}\right)$$

The joint probability of drug choice and treatment outcome can then be written as

$$P(y_{ijt}=d,r_{ijt}=1|z_i,z_j,p_{dt},x_{dt},I_j,\delta_{ijd})=P(r_{ijt}=1|z_j,I_j,y_{ijt}=d,\delta_{ijd})\ P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt},I_j\right)$$

where

$$\delta_{ijd}=\delta_d+\sigma_d\delta_{ij}^d$$

The treatment impact of drugs, δ_{ijd}, is allowed to be heterogeneous across patients with the same observable characteristics and the unobservable patient-state. Assuming that $\delta_{ij}^d\sim N(0,1)$ for all d, we have

$$P(y_{ijt}=d,r_{ijt}=1|z_i,z_j,p_{dt},x_{dt},I_j)=\int P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt},I_j\right)\ P(r_{ijt}=1|z_j,I_j,y_{ijt}=d,\delta_{ijd})\ d\varphi\left(\delta_{ij}^d\right)$$

$$=\int\frac{\exp\left(\alpha_d\left(z_i,z_j\right)-\beta p_{dt}+\gamma_{l(d)}x_{dt}+\lambda_d I_j\right)}{\sum_{d'=0}^{D}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\gamma_{l(d')}x_{d't}+\lambda_{d'}I_j\right)}\varphi\left(z_j'\theta+\sum_{d=1}^{D}\left(\delta_d+\sigma_d\delta_{ij}^d\right)\mathbf{1}_{\{y_{ijt}=d\}}+I_j\right)d\varphi\left(\delta_{ij}^d\right)$$

We can then use simulated maximum likelihood to estimate the model parameters.

The log-likelihood of the sample of physician choices and patient recovery status for all physician-patient pairs will then be $\sum_{(i,j,t)}\ln l(y_{ijt}=d,r_{ijt}=1)$ where

$$l(y_{ijt}=d,r_{ijt}=1)=\frac{1}{S}\sum_{s=1}^{S}\left\{qF\left(\delta_{ij}^{ds},\overline{I}\right)+(1-q)F\left(\delta_{ij}^{ds},\underline{I}\right)\right\}$$

and

$$F\left(\delta_{ij}^{ds},I_j\right)=\frac{\exp\left(\alpha_d\left(z_i,z_j\right)-\beta p_{dt}+\gamma_{l(d)}x_{dt}+\lambda_d I_j\right)}{\sum_{d'=0}^{D}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\gamma_{l(d')}x_{d't}+\lambda_{d'}I_j\right)}\varphi\left(z_j'\theta+\sum_{d=1}^{D}\left(\delta_d+\sigma_d\delta_{ij}^{ds}\right)\mathbf{1}_{\{y_{ijt}=d\}}+I_j\right)$$

and S is the total number of simulation draws of the random variable δ_{ij}^d.

Note that this model of prescription choice leads to an aggregate demand for drug d that can be written as

$$q_{dt} = \sum_{j \in J} \left\{ \frac{q \exp\left(\alpha_d\left(z_i, z_j\right) - \beta p_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d \overline{I}\right)}{\sum_{d'=0}^{D} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} \overline{I}\right)} + \frac{(1-q) \exp\left(\alpha_d\left(z_i, z_j\right) - \beta p_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d \underline{I}\right)}{\sum_{d'=0}^{D} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} \underline{I}\right)} \right\}$$

where the set J denotes the set of all patients diagnosed with depression.

Once the model is estimated, we can obtain the unconditional recovery probability for each patient as the sum, over all the alternatives, of the conditional recovery probability with each alternative times the prescription probability of that alternative as

$$\begin{aligned} E\left[r_{ijt}\right] &= P\left(r_{ijt}=1\right) = \sum_{d=0}^{D} P(r_{ijt}=1|y_{ijt}=d)\ P(y_{ijt}=d) \\ &= \sum_{d=0}^{D} \int P\left(r_{ijt}=1|z_i, I_j, y_{ijt}=d, \delta_{ijd}\right)\ P\left(y_{ijt}=d|z_i, z_j, p_{dt}, x_{dt}, I_j\right)\ d\varphi\left(\delta_{ij}^{d}\right) \end{aligned}$$

We use an exclusion restriction on drug advertising that is allowed to affect treatment choice but not treatment outcome conditional on the treatment choice[9]. If patients treated with drugs that are advertised more obtain better health outcomes, the model attributes this effect to the drug choice[10]. Because the direct-to-consumer promotion of prescription drugs is forbidden in France (Gallini et al. (2013)), patients are not exposed to drug advertisement.

4.2 A Supply-Side Model of Price Setting with Off-Label Drugs

We now show how we can use a price negotiation model *à la* Crawford and Yurukoglu (2012) between the pharmaceutical companies and the regulator to infer the possible changes in the outcomes of these price negotiations if off-label prescriptions are banned.

Let us assume that the price negotiations between the French regulator and pharmaceutical companies can be represented by a bargaining process for the pricing of each active ingredient in which the firm cares about its profit and the government cares about a weighted average of the consumer surplus and the cost of

[9]The model is identified even if the exogenous covariates in the drug choice equation are the same as the exogenous covariates in the treatment outcome equation. However, it is preferable that some variables in the drug choice equation are excluded from the treatment outcome equation so that our identification is not only driven by parametric assumptions but also comes from exclusion restrictions.

[10]The identification assumption is that advertisement is uncorrelated with unobserved heterogeneity. As a response to the argument that "if there is seasonality in both unobserved heterogeneity and drug advertisement, they would be correlated", we do not observe seasonality in drug advertisement (for details on no-seasonality in drug advertisement see the online appendix B.4).

treatment. The fact that the government does not simply account for consumer surplus, which is obtained by the demand shape, can be justified by the fact that public health insurance coverage is large and thus the demand elasticity (which may even be zero) may not represent the health insurance budget opportunity cost.

Bargaining models have been used to represent negotiations between insurers and hospitals in the U.S. (Gowrisankaran et al. (2015)) and between hospitals and medical device providers (Grennan (2013)). Dubois and Lasio (2018) show how to identify price cost margins in the French regulatory environment and present conditions under which a model with price caps set by the regulators can be strategically equivalent to a Nash bargaining model. We choose to model the effects of price regulation using bargaining because in the counterfactual analysis, we do not intend to change the price-setting behavior of the regulator but only the ability of physicians in off-label prescriptions.

The profit of the firm for active ingredient d at period t is

$$\Pi_{dt}(\mathbf{p}_t) = [p_{dt} - c_{dt}]\, q_{dt}^{tot}\left(\mathbf{p}_t^{on}, \mathbf{p}_t^{off}\right)$$

where $q_{dt}^{tot}\left(\mathbf{p}_t^{on}, \mathbf{p}_t^{off}\right)$ is the total sales of drug d in markets for both on-label and off-label indications, at the vector of prices $\mathbf{p}_t = \left(\mathbf{p}_t^{on}, \mathbf{p}_t^{off}\right)$ where $\mathbf{p}_t^{on}$ is the vector of prices of all drugs in the relevant on-label market and $\mathbf{p}_t^{off}$ is the vector of prices of all drugs in the relevant off-label market[11].

We assume that ε_{ijdt} is i.i.d. with type I extreme value, and hence, the consumer surplus on the on-label market of drug d has the standard closed form (Small and Rosen (1981)):

$$CS_t(\mathbf{p}_t^{on}) = \frac{1}{\beta}\sum_{j\in J_{on}} \ln\left(\sum_{d'\in D^{on}} q\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't} + \lambda_{d'}\overline{I}\right) + (1-q)\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't} + \lambda_{d'}\underline{I}\right)\right)$$

where D^{on} is the set of drugs prescribed for the on-label indication of drug d, which contains both on-label drugs and off-label drugs, and the set J_{on} denotes the set of all patients diagnosed with the approved indication of drug d (on-label indication, namely depression)[12].

Let us now assume that the price negotiation can be represented by Nash bargaining with a bargaining parameter of μ for the firm. The Nash in Nash equilibrium concepts involve that all contracts remain the same if another negotiation fails (see, e.g., Crawford and Yurukoglu (2012), Gowrisankaran et al. (2015), Dubois and Sæthre (2020)). Assuming a Nash in Nash equilibrium (Horn and Wolinsky (1988)), whose microfoundation

[11] The price and the reimbursement level of a drug do not change depending on the indication for which it is prescribed. However, since the set of drugs used for an on-label indication of drug d is different than the set of drugs used for an off-label indication of drug d, the vectors of prices p_t^{on} and p_t^{off} are different.

[12] Drug advertisement is not included in consumer surplus.

has been recently clarified (Collard-Wexler et al. (2019)), active ingredient by active ingredient amounts to maximizing the Nash product given the other prices:

$$\max_{p_{dt}} \left[\Pi_{dt}\left(\mathbf{p}_t^{on}, \mathbf{p}_t^{off}\right)\right]^{\mu} \left[w\Delta_d CS_t\left(\mathbf{p}_t^{on}\right) - (1-w)\Delta_d TC_{dt}\left(\mathbf{p}_t^{on}\right)\right]^{1-\mu}$$

where Π_{dt} is the profit of the firm from sales of active ingredient d. Note that the difference in firm profits with or without drug d is equal to the profit from drug d because we assume firms behave to maximize profit active ingredient by active ingredient. $\Delta_d CS_t(\mathbf{p}_t^{on}) \equiv CS_t(\mathbf{p}_t^{on}) - CS_{t,-d}(\mathbf{p}_t^{on})$ is drug d's value added in consumer surplus for the on-label indication. $\Delta_d TC_{dt}(\mathbf{p}_t^{on})$ is the difference between the total cost of treatment of the on-label indication of drug d when drug d is on the market and when it is not, such that

$$\Delta_d TC_{dt}(\mathbf{p}_t^{on}) = \sum_{\tilde{d}\in D^{on}} p_{\tilde{d}t} q_{\tilde{d}t}(\mathbf{p}_t^{on}) \quad - \quad \sum_{\tilde{d}\in D^{on}\setminus\{d\}} p_{\tilde{d}t} q_{\tilde{d}t}^{-d}(\mathbf{p}_t^{on})$$

where $q_{\tilde{d}t}(\mathbf{p}_t^{on})$ is the demand of drug $\tilde{d}$ in the on-label market and $q_{\tilde{d}t}^{-d}(\mathbf{p}_t^{on})$ is the demand of drug $\tilde{d}$ when drug d is absent. w is the weight the government puts on consumer surplus, and hence, $(1-w)$ is the weight the government puts on the change in cost of treatment drug d brings. Note that the consumer surplus absent drug d is

$$CS_{t,-d}(\mathbf{p}_t^{on}) = \frac{1}{\beta} \sum_{j\in J_{on}} \ln\left(\sum_{d'\in D^{on}\setminus\{d\}} q\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't} + \lambda_{d'}\overline{I}\right) + (1-q)\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't} + \lambda_{d'}\underline{I}\right)\right)$$

The first-order conditions of this Nash equilibrium are

$$\frac{\mu}{1-\mu}\frac{\partial \ln \Pi_{dt}\left(\mathbf{p}_t^{on}, \mathbf{p}_t^{off}\right)}{\partial p_{dt}} + \frac{\partial \ln\left[w\Delta_d CS_t\left(\mathbf{p}_t^{on}\right) - (1-w)\Delta_d TC_{dt}\left(\mathbf{p}_t^{on}\right)\right]}{\partial p_{dt}} = \mathbf{0} \tag{4.2}$$

Thus, the firm's marginal cost is

$$c_{dt} = p_{dt} + \frac{1}{\left(\frac{\partial \ln q_{dt}^{tot}\left(\mathbf{p}_t^{on}, \mathbf{p}_t^{off}\right)}{\partial p_{dt}}\right) + \frac{1-\mu}{\mu}\left(\frac{\partial \ln[w\Delta_d CS_t(\mathbf{p}_t^{on}) - (1-w)\Delta_d TC_{dt}(\mathbf{p}_t^{on})]}{\partial p_{dt}}\right)} \tag{4.3}$$

where $\frac{\partial \ln q_{dt}^{tot}\left(\mathbf{p}_t^{on}, \mathbf{p}_t^{off}\right)}{\partial p_{dt}}$ is the price semi-elasticity of aggregate demand in both the on-label and off-label markets.

Note that once we know the demand shape, we also have

$$\begin{aligned}\frac{\partial \Delta_d CS_t\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}} &= \frac{\partial CS_{td}\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}} - \frac{\partial CS_{t,-d}\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}} \\ &= \frac{1}{\beta}\sum_{j\in J_{on}} -\beta\left\{ q\frac{\exp\left(\alpha_d(z_i,z_j)-\beta p_{dt}+\lambda_d\overline{I}\right)}{\sum\limits_{d'\in D^{on}}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\lambda_{d'}\overline{I}\right)}\right. \\ &\quad \left. +(1-q)\frac{\exp\left(\alpha_d(z_i,z_j)-\beta p_{dt}+\lambda_d\underline{I}\right)}{\sum\limits_{d'\in D^{on}}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\lambda_{d'}\underline{I}\right)}\right\}\end{aligned}$$

and

$$\frac{\partial \Delta_d TC_{dt}\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}} = \sum_{\tilde{d}\in D^{on}} \frac{\partial p_{\tilde{d}t} q_{\tilde{d}t}\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}} - \sum_{\tilde{d}\in D^{on}\setminus\{d\}} \frac{\partial p_{\tilde{d}t} q_{\tilde{d}t}^{-d}\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}} = q_{dt}\left(\mathbf{p}_t^{on}\right) + \sum_{\tilde{d}\in D^{on}} p_{\tilde{d}t}\frac{\partial q_{\tilde{d}t}\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}}$$

Moreover, the profit of the firm in both markets is

$$\Pi_{dt}\left(\mathbf{p}_t^{on},\mathbf{p}_t^{off}\right) = \left[p_{dt}-c_{dt}\right] q_{dt}^{tot}\left(\mathbf{p}_t^{on},\mathbf{p}_t^{off}\right) = \left[p_{dt}-c_{dt}\right]\left[q_{dt}^{on}\left(\mathbf{p}_t^{on}\right) + q_{dt}^{off}\left(\mathbf{p}_t^{off}\right)\right]$$

implying that

$$\frac{\partial \Pi_{dt}\left(\mathbf{p}_t^{on},\mathbf{p}_t^{off}\right)}{\partial p_{dt}} = q_{dt}^{on}\left(\mathbf{p}_t^{on}\right) + q_{dt}^{off}\left(\mathbf{p}_t^{off}\right) + \left[p_{dt}-c_{dt}\right]\left[\frac{\partial q_{dt}^{on}\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}} + \frac{\partial q_{dt}^{off}\left(\mathbf{p}_t^{off}\right)}{\partial p_{dt}}\right]$$

where

$$\frac{\partial q_{dt}^{on}\left(\mathbf{p}_t^{on}\right)}{\partial p_{dt}} = -\beta\sum_{j\in J_{on}}\left\{P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt},I_j,on\right)\left(1-P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt},I_j,on\right)\right)\right\}$$

and

$$\frac{\partial q_{dt}^{off}\left(\mathbf{p}_t^{off}\right)}{\partial p_{dt}} = -\beta\sum_{j\in J_{ojj}}\left\{P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt},I_j,off\right)\left(1-P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt},I_j,off\right)\right)\right\}$$

which are also known from demand estimates. The set J_{off} denotes the set of all patients diagnosed with the indications drug d is not approved for but used off-label (off-label indications of drug d). Note that antidepressants are used off-label in the treatment of the following diseases: alcoholism, anguish, anxiety, asthenia, bipolar, dementia, headaches and migraines, high blood pressure, insomnia, other psychotic disorders, pain, smoking, and schizophrenia. In estimating the profits of firms, the market shares of each of the active

ingredients in each of these markets is taken into account, and hence, the profits coming from these markets are included in the total profits of the firms.

In the case where demand is price insensitive and thus $\beta = 0$, equation (4.3) simplifies to

$$c_{dt} = p_{dt} + \frac{\Delta_d TC_{dt}\left(\mathbf{p}_t^{on}\right) - \frac{w}{1-w}\ \Delta_d CS_t\left(\mathbf{p}_t^{on}\right)}{\frac{1-\mu}{\mu}\ q_{dt}\left(\mathbf{p}_t^{on}\right)}$$

5 Econometric Identification and Empirical Estimates

5.1 Demand Model Identification and Estimates

The most commonly used on-label drugs in depression treatment are Selective Serotonin Reuptake Inhibitors (SSRIs), which include the active ingredients Citalopram, Escitalopram, Fluoxetine, Paroxetine, and Sertraline. In estimating choice probabilities, they are considered distinct alternatives in the physicians' choice set. All the remaining approved active ingredients, each of which has a smaller market share, are classified under the choice called 'Other On-Label Drugs'. Off-label drugs are classified under two distinct categories: 'Off-Label Drugs with Efficacy', which are the off-label active ingredients for which scientific evidence from the medical literature shows their efficacy in the treatment of depression, and 'Other Off-Label Drugs', which include all the other off-label active ingredients and is the reference alternative in the estimation (details on the drugs aggregated under alternatives 'Other On-Label Drugs', 'Off-Label Drugs with Efficacy' and 'Other Off-Label Drugs' can be found in appendix A.1). In total, there are 8 exclusive alternatives in the physicians' choice set (because co-prescriptions are less than 3% of the cases, they are excluded from the analysis).

Patients that are diagnosed with depression but did not receive any drug prescription are excluded from the analysis as their share is only 0.4% of all the depression patients. To avoid dealing with the impact of learning about patients' response to prescription choice, as in Crawford and Shum (2005) and Dickstein (2018), we consider the prescription choice on the first visit at which the patient is diagnosed with depression[13].

The treatment outcome considered in the analysis is whether the patient recovers from the disease after six months. We allow for a six-month treatment period; then, we check whether the patient was diagnosed with depression again in the one-year period after the six-month treatment period[14]. The recovery is defined by

[13]The correlation between the prescription at the first visit of diagnosis and the prescriptions during the entire treatment period is provided in the online appendix. The estimation of the treatment outcome equation conditional on the drug that is prescribed the most often during the treatment period is also provided in the online appendix.

[14]For robustness, the same analysis was performed by allowing for combinations of six-month or one-year treatment periods with six-month or one-year observation periods after the treatment period. The recovery rates are higher after one year of treatment; however, the impact of counterfactual policies are the same.

the patient not being diagnosed with depression during this one-year observation period after the six-month treatment period. Given that the sample period ends in December 2008, only the patients who are diagnosed with depression until the end of June 2007 are included in the analysis[15]. The patients who stopped visiting their family physician during or after the six-month treatment period, who are no longer in the sample, are not included in the analysis. The share of patients who stopped visiting their family physician is the same for patients treated with on-label drugs and for those treated with off-label drugs (for details, see appendix A.4), and these dropouts are likely to be due to moving both within the city and out of the city.

As a baseline for the analysis, we first estimate the model assuming no correlation between the unobserved patient state affecting prescription choices and treatment outcomes, which means imposing $\lambda_d = 0$ for all d. The treatment choice model is then a logit and the recovery model is a binary probit model. Table 5.1 reports the parameters of the logit estimation of treatment choice in this case, and Table 5.2 provides the parameters of the binary probit estimation of the treatment outcome. In the estimations, the reference category is 'Other Off-Label Drugs'; therefore, coefficients are interpreted relative to 'Other Off-Label Drugs'. For instance, the positively significant coefficient for patients' age for the on-label drug 'Citalopram' means that older patients are more likely to obtain a 'Citalopram' prescription relative to an 'Other Off-Label Drug'. The drug-specific constant terms control for the time-invariant variation across drugs. Physician and patient characteristics (age and gender) are also introduced.

Table 5.1 shows that demand is price sensitive with a significantly negative coefficient for price[16]. Additionally, detailing expenditures positively affect the demand in the on-label market. However, its impact on the demand in the off-label market is limited, similar to the findings in Shapiro (2018). This result is coherent with the fact that pharmaceutical firms can advertise their drugs for on-label use but are not allowed to advertise them for off-label use. Though it is coherent with the regulation, we still cannot identify whether the very small impact of drug advertisement on off-label use of drugs is because firms are not advertising drugs for off-label use or because, even though they do, physicians' response to these advertisements is limited.

Table 5.2 presents the estimation outcome of the binary probit model for the treatment outcome. The control group is 'Other Off-Label Drugs', and the drug-specific coefficients are the δ_ds in the model, the treatment effect of drug d relative to the control group. As a baseline for the analysis, we do not allow

[15]To visit a psychiatrist, patients need a referral from their family physician. We can observe whether a patient is referred to a specialist. The patients who are referred to a psychiatrist on the first visit of depression diagnosis or within the six-month treatment period are not included in the analysis. The patients who are referred to a specialist after the six-month period are considered as 'still-depressed'; hence, the treatment outcome of recovery is zero for them.

[16]In the case of depression treatment, there is almost no variation in reimbursement rates across drugs, as they are almost all reimbursed at 65%. Therefore, we do not take into account the reimbursement rate in the estimations.

for heterogeneity in treatment impact across patients; hence, $\sigma_d = 0$. δ_ds are significantly negative for all the approved drugs, which means that patients treated with approved drugs have lower recovery rates than patients treated with 'Other Off-Label Drugs'. The recovery rates with alternative 'Off-Label Drugs with Efficacy' are not different from the recovery rates with 'Other Off-Label Drugs'.

Table 5.1: *Logit Estimation of Treatment Choice*

	Alternative Specific Parameters				
	Patients'		Physicians'		
	Age	Sex	Age	Sex	Constant
On-Label Drugs:					
Citalopram	0.015	-0.072	-0.014	-0.092	-2.332
	(0.001)	(0.044)	(0.003)	(0.052)	(0.424)
Sertraline	0.006	-0.124	0.001	-0.185	-2.918
	(0.001)	(0.046)	(0.003)	(0.056)	(0.425)
Paroxetine	0.012	-0.149	0.003	-0.218	-2.595
	(0.001)	(0.037)	(0.002)	(0.045)	(0.412)
Fluoxetine	0.012	0.040	0.001	-0.153	-3.575
	(0.001)	(0.042)	(0.003)	(0.050)	(0.407)
Escitalopram	0.010	-0.115	-0.002	-0.417	-3.800
	(0.001)	(0.056)	(0.004)	(0.073)	(0.434)
Other	0.022	-0.175	0.006	-0.022	-4.104
	(0.001)	(0.035)	(0.002)	(0.042)	(0.407)
Off-Label Drugs:					
with Efficacy	-0.011	-0.265	-0.001	-0.127	-0.843
	(0.002)	(0.067)	(0.004)	(0.084)	(0.022)
Parameters Common Across Alternatives					
Price ($-\beta$)	Advertising*On-Label			Advertising*Off-Label	
-1.042	0.227			0.004	
(0.278)	(0.014)			(0.023)	
Observations	37,510				

Notes: Advertising is the natural logarithm of the stock of advertising. The "Sex" variable is 1 for females and 0 for males. Standard errors are in parentheses.

Table 5.3 provides the results of the joint estimation of treatment choice and treatment outcomes, taking into account unobserved patient heterogeneity, which can potentially affect both the treatment choice and the treatment outcome. In the estimation we normalize the higher value of I_j to 1 and lower value of I_j to -1 and we estimate the shares[17], q.

Detailing expenditures to general practitioners at the drug-month level are excluded from the health outcome equation. Although the model is identified when the exogenous covariates in the drug choice equation are the same as the exogenous covariates in the treatment outcome equation, it is preferable that some variables in the drug choice equation are excluded from the treatment outcome equation so that our

[17]For robustness, we jointly estimate the demand and health outcome equation normalizing $\overline{I}$ and $\underline{I}$ to a few different values. The estimation and counterfactual results in all the cases are consistent with the current results and are reported in the online appendix B.8.

identification is not only driven by parametric assumptions but also comes from exclusion restrictions. We use drug advertising as an excluded variable in the health outcome equation. Advertising affects drug choices but, conditional on the drug choice, it should not impact the health outcomes. The model attributes the change in health outcomes correlated to a different drug choice steered by advertising as coming from the drug choice.

Table 5.2: *Binary Probit Estimation of Treatment Outcome*

	Parameter	Std. Error	Marginal Effect	Std. Error
Patients' Age	-0.013	(0.000)	-0.005	(0.000)
Patients' Sex (female=1)	-0.079	(0.014)	-0.030	(0.005)
On-Label Drugs				
Citalopram	-0.295	(0.026)	-0.111	(0.010)
Sertraline	-0.289	(0.028)	-0.108	(0.010)
Paroxetine	-0.285	(0.022)	-0.107	(0.008)
Fluoxetine	-0.290	(0.025)	-0.108	(0.009)
Escitalopram	-0.316	(0.033)	-0.118	(0.012)
Other	-0.342	(0.021)	-0.128	(0.008)
Off-Label Drugs				
with Efficacy	-0.015	(0.043)	-0.005	(0.016)
Constant	1.163	(0.027)		
Number of Observations	37,510			

Notes: Standard errors are in parentheses.

The first part of Table 5.3 reports the parameters of the estimation of treatment choice, including the impact of unobserved patient-state on the prescription choice of each alternative through λ_d. The second part reports the parameters of the estimation of treatment outcomes. The drug-specific coefficients in part 2 of Table 5.3 are the δ_ds in the model, the treatment effect of drug d relative to the control group. The simulated estimation is based on 400 normalized Halton draws for δ_{ij}^d for each patient[18].

A statistically significant parameter of unobserved heterogeneity, λ_d, means that there is unobserved selection into treatment, which impacts both the prescription probability and the recovery probability. The δ_ds show the average treatment impact of each drug relative to 'Other Off-Label Drugs', after controlling for the unobserved patient-level heterogeneity. The estimate of q shows that 90% of the patients are high types; hence, 10% are low types. The negatively significant estimates of λ_d for the on-label alternatives show that these active ingredients are less (more) likely to be prescribed to high (low) types, which are the patients who are more (less) likely to recover from the disease due to their unobserved state. The severity of the disease

[18]Halton sequences are preferred to pseudo-random draws thanks to two desirable properties. First, they give more even coverage over the domain of the mixing distribution. Because the draws for each observation are more evenly spread, the simulated probabilities vary less over observations relative to the probabilities calculated by random draws. Second, with Halton draws, the simulated probabilities are negatively correlated over observations, and this negative correlation decreases the variance in the simulated likelihood function (Deb and Trivedi (2006)).

can be an example of this type of heterogeneity, i.e., relatively more severe cases (low-types) are more likely to be prescribed the on-label alternatives, relative to 'Other Off-Label Drugs', and because they are more severe cases, they are less likely to recover from the disease.

Once we control for the unobserved heterogeneity, the relative average treatment impact parameter for most of the on-label alternatives is either positive or not statistically different from zero. The significantly positive estimate for δ_d for the alternative called 'Sertraline' shows that, on average, the treatment impact of 'Sertraline' relative to 'Other Off-Label Drugs' is positive. The estimate for the δ_d parameter is not statistically different than zero for the alternatives 'Citalopram', 'Paroxetine' and 'Escitalopram', which means that, on average, the treatment impact of these alternatives relative to the reference group is not statistically different from zero, in contrast to the results presented in Table 5.2, which does not take into account unobserved heterogeneity. The average relative treatment impact of the alternatives 'Fluoxetine' and 'Other On-Label' drugs is negative despite the selection into treatment based on unobservables. Overall, the results in Table 5.3 show that there is unobserved patient-level heterogeneity that impacts both drug choice and recovery status.

Statistically significant estimates of σ_d show that the treatment effect of each drug, relative to the reference group, is heterogeneous across patients, even after controlling for observables and patient-level unobserved state. This means, for some patients, treatment with 'Citalopram', for instance, leads to lower recovery rates than treatment with 'Other Off-Label Drugs'. However, for some other group of patients, it leads to higher recovery rates, even though the average relative impact, δ_d, is not statistically different from zero. It is important to not misinterpret the estimates of σ_d. σ_d estimates show the heterogeneity in relative treatment effect of each drug, relative to the reference group.

A statistically significant estimate of σ_d means for some group of patients drug d leads to higher recovery rates than 'Other Off-Label Drugs' whereas for some other group of patients, it leads to lower recovery rates than 'Other Off-Label Drugs'. It does not, at all, mean that drug d is detrimental for some share of the population, it just means it is worse than 'Other Off-Label Drugs'. Heterogeneity in treatment impact of drugs across patients in depression treatment is a well-documented result in the medical literature (see, e.g., Simon and Perlis (2010), Perlis (2014), Uher et al. (2012)). The coefficients of the observed patient characteristics in Table 5.3 show that older patients and female patients are less likely to recover from depression.

Table 5.3: *Joint Estimation of Treatment Choice and Treatment Outcome*

Part 1: Treatment Choice Equation					
Alternative Specific Parameters					
	Patients'		Physicians'		λ_d
	Age	Sex	Age	Sex	
On-Label Drugs					
Citalopram	0.014	-0.087	-0.017	-0.122	-2.124
	(0.001)	(0.044)	(0.003)	(0.052)	(0.470)
Sertraline	0.004	-0.162	0.001	-0.220	-3.336
	(0.002)	(0.058)	(0.003)	(0.072)	(0.577)
Paroxetine	0.012	-0.165	0.003	-0.235	-2.683
	(0.001)	(0.040)	(0.003)	(0.049)	(0.500)
Fluoxetine	0.012	0.039	-0.001	-0.164	-0.284
	(0.001)	(0.042)	(0.003)	(0.050)	(0.000)
Escitalopram	0.010	-0.117	-0.003	-0.421	-0.526
	(0.002)	(0.057)	(0.004)	(0.074)	(0.000)
Other	0.022	-0.177	0.005	-0.033	-0.841
	(0.001)	(0.035)	(0.002)	(0.042)	(0.000)
Off-Label Drugs					
with Efficacy	-0.011	-0.271	-0.001	-0.133	2.252
	(0.002)	(0.068)	(0.004)	(0.084)	(0.210)

Parameters Common Across Alternatives			
	Advertising		Share of
Price ($-\beta$)	On-Label	Off-Label	*High Types (q)*
-1.105	0.250	0.043	0.899
(0.307)	(0.020)	(0.024)	(0.180)

Part 2: Treatment Outcome Equation				
	δ_d	Std. Err.	σ_d	Std. Err.
On-Label Drugs				
Citalopram	-0.026	(0.079)	1.128	(0.253)
Sertraline	0.927	(0.325)	0.855	(0.373)
Paroxetine	0.348	(0.213)	1.422	(0.300)
Fluoxetine	-0.211	(0.040)	0.883	(0.199)
Escitalopram	-0.147	(0.082)	1.590	(0.410)
Other	-0.242	(0.030)	1.192	(0.169)
Off-Label Drugs with Efficacy	0.131	(0.148)	0.840	(0.400)
	Coef.	Std. Err.		
Patients' Age	-0.019	(0.001)		
Patients' Sex	-0.101	(0.021)		
Constant	0.441	(0.051)		
Observations	37.510			

Notes: Standard errors are in parentheses. Advertising is the natural logarithm of the stock of advertising. The "Sex" variable is 1 for females and 0 for males. The estimation includes a dummy that is 1 for the period after the warning in December 2004 about the increase of suicidal thinking in children and adolescents treated with SSRI type antidepressants (for details, see Dubois and Tunçel (2020)).

Table 5.4 reports the quantiles of the relative treatment effects of the drugs relative to the reference alternative 'Other Off-Label Drugs'. For every on-label alternative, there are some group of patients for whom the on-label alternative is better than 'Other Off-Label Drugs' and some other group of patients for whom 'Other Off-Label Drugs' is better than the on-label alternative. Hence, for all the on-label drugs, there are some patients for whom the treatment effect relative to 'Other Off-Label Drugs' is positive, with the lowest being 40% for the 'Fluoxetine' treatment.

Table 5.4: *Quantiles on the Marginal Effect of Treatment*

	Quantiles				Percentage with Positive Relative Treatment Effect
	25%	50%	75%	95%	
On-Label					
Citalopram	-9.0	-0.4	16.8	43.9	49%
Sertraline	7.4	19.9	32.8	63.2	86%
Paroxetine	-6.2	7.5	26.7	67.7	60%
Fluoxetine	-10.8	-2.4	8.7	29.9	40%
Escitalopram	-11.7	-1.6	20.0	58.7	46%
Other On-Label	-12.2	-2.5	13.0	40.2	42%
Off-Label					
with Efficacy	-5.6	2.4	16.2	37.2	56%

Notes: The quantiles are in terms of percentage points showing the treatment impact of the drug in the first column relative to the reference alternative 'Other Off-Label Drugs'. For instance, the first cell shows that for 25% of the patients recovery rates when treated with 'Citalopram' are at least 9 percentage points lower than recovery rates when treated with 'Other Off-Label Drugs'.

Table 5.5 reports the recovery probabilities across patients if all the patients are treated with the same drug separately for *low-* and *high-type* patients. For instance, if all patients are treated with 'Citalopram', the average recovery rate is 0.16 for *low-type* patients and 0.62 for the *high types*. The average recovery rate is highest when all patients are treated with 'Sertraline' both for *high-* and *low-type* patients.

Table 5.5: *Quantiles on the Recovery Probabilities*

	Quantiles				
	25%	50%	75%	95%	Mean
If all *low-type* patients are treated with:					
On-Label					
Citalopram	.01	.06	.23	.66	.16
Sertraline	.12	.28	.51	.82	.33
Paroxetine	.02	.12	.43	.90	.26
Fluoxetine	.01	.04	.14	.43	.11
Escitalopram	.00	.05	.29	.85	.19
Other On-Label	.00	.04	.18	.61	.14
Off-Label					
with Efficacy	.02	.08	.22	.54	.15
Other	.04	.07	.10	.15	.07
If all *high-type* patients are treated with:					
On-Label					
Citalopram	.37	.68	.90	.99	.62
Sertraline	.79	.92	.98	1.00	.85
Paroxetine	.45	.80	.97	1.00	.68
Fluoxetine	.36	.61	.82	.97	.58
Escitalopram	.23	.64	.93	1.00	.57
Other On-Label	.28	.60	.86	.99	.56
Off-Label					
with Efficacy	.51	.73	.89	.98	.68
Other	.62	.70	.77	.83	.68

5.2 Supply-Side Identification and Estimates

We now present our identification strategy on parameters of the supply-side model and estimates of these parameters

The first-order conditions (4.2) depend on demand, parameters μ, w and marginal costs c_{dt}. We can set identify μ, and w using some simple and robust cost restrictions imposing that marginal costs of drugs are positive and smaller than the lowest observed prices for these drugs. Using (4.3), the firm's marginal cost depends on the demand parameter vector denoted by $\mathbf{\Lambda}$ and the supply-side parameters μ and w as follows

$$c_{dt}\left(\mathbf{\Lambda},\mu,w,\mathbf{k}_t\right)=p_{dt}+\frac{1}{\left(\frac{\partial \ln q_{dt}^{tot}\left(\mathbf{p}_t^{on},\mathbf{p}_t^{off}\right)}{\partial p_{dt}}\right)+\frac{1-\mu}{\mu}\left(\frac{\partial \ln[w\Delta_d CS_t(\mathbf{p}_t^{on})-(1-w)\Delta_d TC_{dt}(\mathbf{p}_t^{on})]}{\partial p_{dt}}\right)}$$

where $\mathbf{k}_t$ is the vector of observed variables used to obtain marginal costs (prices and characteristics).

We then assume that true marginal costs c_d^0 are time invariant and that ζ_{dt} is mean independent of true cost c_d^0 such that

$$c_{dt}\left(\mathbf{\Lambda},\mu,w,\mathbf{k}_t\right)=c_d^0+\zeta_{dt}$$

These imply the moment condition

$$E\left(c_{dt}\left(\mathbf{\Lambda},\mu,w,\mathbf{k}_t\right)\right)=E\left(c_d^0+\zeta_{dt}\right)=c_d^0$$

Using the set of natural economic inequalities that cost should be positive and lower than price at any time,

$$0\leq c_d^0\leq \underline{p_d}\equiv\min_t p_{dt}\qquad\forall d=1,..,D$$

we obtain the following moment inequalities:

$$0\leq E\left(c_{dt}\left(\mathbf{\Lambda},\mu,w,\mathbf{k}_t\right)\right)\leq \underline{p_d}\quad\forall d=1,..,D$$

Note that we could slightly generalize by allowing marginal costs c_d^0 to vary over time and then use time-specific upper bounds on the cost. In practice, the minimum price $\underline{p_d}$ will be the minimum price of the corresponding drug d over an observation period that goes beyond the period of estimates of costs[19], until 2018.

Then, for a given vector of demand parameters $\mathbf{\Lambda}$, the identified set of supply-side parameters is

$$S_{(\mu,w)}\left(\mathbf{\Lambda}\right)=\left\{(\mu,w)\,|0\leq E\left(c_{dt}\left(\mathbf{\Lambda},\mu,w,\mathbf{k}_t\right)\right)\leq \underline{p_d},\,\forall d=1,..,D\right\}$$

and can be empirically estimated using

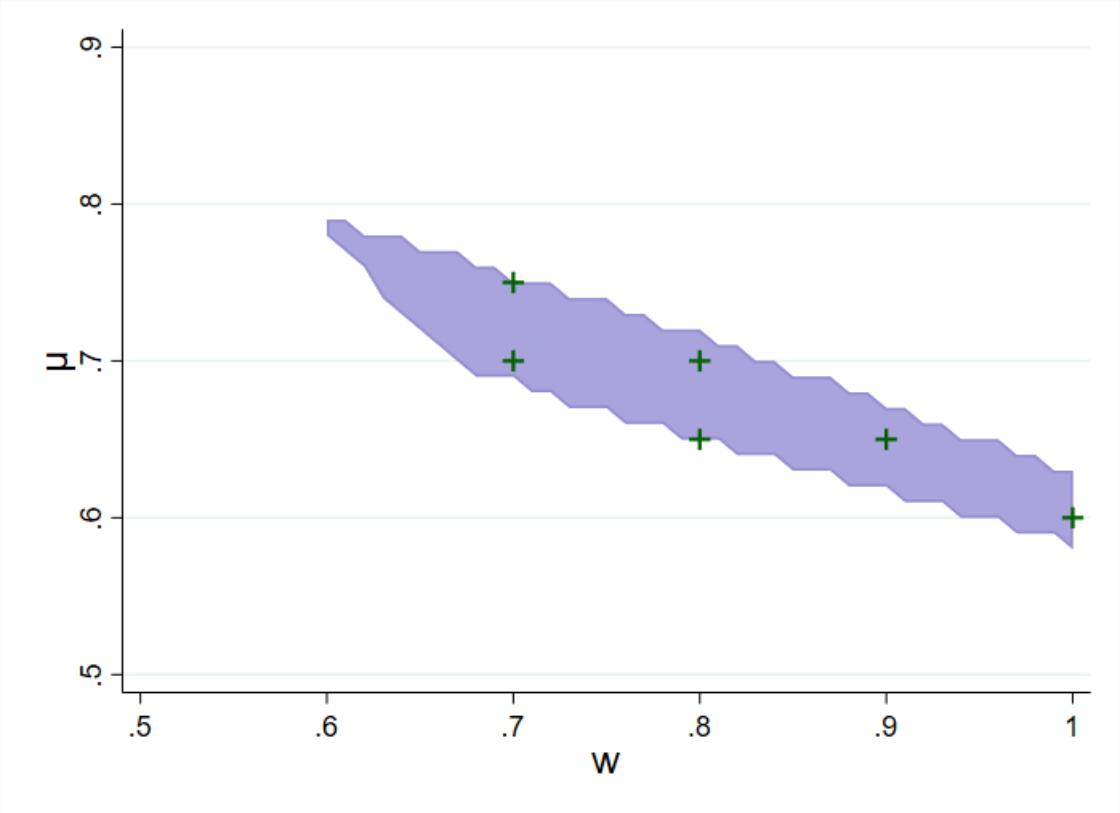
$$\hat{S}_{(\mu,w)}\left(\mathbf{\Lambda}\right)=\left\{(\mu,w)|0\leq\frac{1}{T}\sum\nolimits_{t=1}^{T}c_{dt}\left(\mathbf{\Lambda},\mu,w,\mathbf{k}_t\right)\leq \underline{p_d},\,\forall d=1,..,D\right\}$$

Remark that demand parameters are unknown and, while they are independently estimated, we should also account for the uncertainty in estimates of $\mathbf{\Lambda}$ and obtain random sets $\hat{S}_{(\mu,w)}\left(\mathbf{\Lambda}\right)$. The asymptotic theory of these identified sets can be studied using methods proposed by Chernozhukov et al. (2007) or Andrews and Barwick (2012) for a fixed $\mathbf{\Lambda}$ but would require some extension when $\mathbf{\Lambda}$ is unknown and estimated from an auxiliary model. For simplicity, we use the estimated parameter $\widehat{\mathbf{\Lambda}}$ to compute the identified set $\hat{S}_{(\mu,w)}(\widehat{\mathbf{\Lambda}})$ and simply make some robustness checks with respect to $\widehat{\mathbf{\Lambda}}$.

[19]Drug prices decrease over time in France mostly due to generic entry and/or regulatory changes.

In the identified set[20], the bargaining power of the firm, μ, takes values between 0.58 and 0.79 and the weight the government puts on consumer surplus, w, is between 0.60 and 1. The shaded are in figure 5.1 shows the whole identified set for μ and w. For instance, when $\mu = 0.60$, w takes values between 0.94 and 1, whereas when $\mu = 0.78$, w takes values on a smaller range, between 0.60 and 0.61. Because it would not be feasible to do the counterfactuals for all the combinations of μ and w, we will provide the marginal cost estimates and counterfactual results for some combinations of μ and w in the identified set, which are marked by cross signs in figure 5.1. Table 5.6 provides the average prices during the sample period, prices in 2018 and the marginal cost of on-label alternatives for these combinations of μ and w.

Figure 5.1: *Identified Set for the Bargaining Parameter and the Weight*



Notes: The cross signs represent the combinations of μ and w for which we provide the marginal cost estimates and counterfactual results.

[20]In estimation of the profit function, quantities include all the visits, not only the first visit. Market shares of choice alternatives across the first visit prescriptions and across all the visits are very similar. The impact of advertisement and price on demand across all the visits is also similar to their impact on demand on the first visit prescriptions (for details on market shares and demand estimation including all visits, see online appendix B.5).

We use product by product bargaining as most firms have a single product on this market except one. We assume, for simplicity, that the only firm (Lundbeck) that has two products (Citalopram and Escitalopram) on this market is bargaining at the active ingredient level. Also for simplicity, we assume that the aggregate choice alternative "Other On-Label Drugs" behave similarly in price setting as an additional firm. Firms that market the main molecules on this market do not have very large share in this aggregate alternative.

Table 5.6: *Marginal Cost Estimates of On-Label Drugs*

	Average Price During Sample Period (€)	Prices in 2018 (€)						
Bargaining Parameter (μ)			0.60	0.65	0.65	0.70	0.70	0.75
Weight on CS (w)			1	0.80	0.90	0.70	0.80	0.70
					Marginal Cost (€/DDD)			
Citalopram	.836	.271	.252	.266	.234	.251	.216	.197
Sertraline	.804	.286	.172	.173	.149	.153	.126	.103
Paroxetine	.776	.203	.161	.171	.142	.154	.122	.101
Fluoxetine	.744	.267	.038	.096	.036	.088	.029	.015
Escitalopram	.685	.209	.043	.137	.064	.140	.073	.070
Other On-Label Drugs	.928	.526	.201	.388	.255	.462	.312	.383

Notes: Price is the price for one-day treatment calculated using the DDD assigned by the World Health Organization. For each active ingredient, it is the price per mg times mg per day according to the DDD.

6 Counterfactual Simulation of a Ban on Off-Label Prescriptions

Using the structural choice model, we perform counterfactual simulations of prescription choices when off-label drugs are not in physicians' choice set. In France, the current regulation under "Temporary Recommendations for Use (TRU)" aims at strictly regulating off-label prescriptions. In the U.S., the formulary drug lists of health insurance companies contribute to limiting off-label prescriptions. Before 2011 in France, physicians were perfectly free to prescribe off-label drugs if they wanted to do so. We thus predict counterfactual prescriptions in the case of an off-label ban by simulating demand if off-label drugs were removed from the physicians' choice set.

After simulating such prescriptions, we can simulate the associated expected cost at the observed prices had the prices been identical in the case of an off-label ban. In the second part of the counterfactual analysis, using the supply-side bargaining model, we will investigate how prices would have been different had negotiations took place under a ban on off-label prescriptions and simulate the associated counterfactual expected demand and cost using the counterfactual prices.

6.1 Theory

6.1.1 Keeping Prices Fixed

Assuming prices would be the same in the case of a ban on off-label use, we can simulate not only the counterfactual prescription choices but also their expected costs and the expected recovery rate of patients simply by removing off-label alternatives from the physicians' choice set.

Removing off-label drugs from the choice set, the counterfactual choice probability that physician i will prescribe an approved drug for which $l(d) = 1$ to patient j is now equal to

$$\begin{aligned} P_c\left(y_{ijt} = d|z_i, z_j, p_{dt}, x_{dt}, I_j\right) &= q \frac{\exp\left(\alpha_d\left(z_i, z_j\right) - \beta p_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d \overline{I}\right)}{\sum_{d' \in D_{on}^{on}} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} \overline{I}\right)} \\ &+ (1-q) \frac{\exp\left(\alpha_d\left(z_i, z_j\right) - \beta p_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d \underline{I}\right)}{\sum_{d' \in D_{on}^{on}} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} \underline{I}\right)} \end{aligned}$$

where $D_{on}^{on} = D^{on} \backslash \{d'' \in D^{on} | l(d'') = 0\}$ is the set of all the on-label drugs for the indication for which d is on-label ($l(d) = 1$ if drug d is an approved drug for depression and 0 otherwise). Note that because the reference group is 'Other Off-Label Drugs', with a ban on off-label drugs, the choice probability of the reference group is 0.

The ex ante expected cost of treatment for patients diagnosed with the on-label indication is

$$\sum\nolimits_{j \in J_{on}} \sum\nolimits_{d \in D^{on}} P\left(y_{ijt} = d|z_i, z_j, p_{dt}, x_{dt}, I_j\right) p_{dt} \tag{6.1}$$

where, as defined above, D^{on} is the set of drugs prescribed for the on-label indication of drug d, which contains both on-label drugs and off-label drugs. The treatment cost ex post (after removing off-label drugs from the choice set) on this on-label indication market is

$$\sum\nolimits_{j \in J_{on}} \sum\nolimits_{d \in D_{on}^{on}} P_c\left(y_{ijt} = d|z_i, z_j, p_{dt}, x_{dt}, I_j\right) p_{dt} \tag{6.2}$$

Therefore, the change in the expected cost is the difference between (6.1) and (6.2). Since drugs approved for depression treatment are, on average, more expensive than off-label drugs used in depression treatment, assuming drug prices will be the same under the ban, we would expect cost of treatment would increase with the ban on off-label prescriptions. When off-label drugs are no longer in the choice set, physicians will necessarily substitute approved drugs for off-label drugs. Because approved drugs are more expensive alternatives, expected prescription expenses will increase due to this substitution effect.

The ex ante expected recovery rate for patient j using the ex ante prescription choice probabilities is

$$E\left[r_{ijt}\right] = \sum_{d \in D^{on}} P(r_{ijt} = 1|z_i, I_j, y_{ijt} = d) P(y_{ijt} = d|z_i, z_j, p_{dt}, x_{dt}, I_j) \tag{6.3}$$

while the expected counterfactual recovery rate r^c_{ijt} using the new prescription choice probabilities is

$$E\left[r^c_{ijt}\right] = \sum_{d \in D^{on}_{on}} P(r_{ijt} = 1 | z_i, I_j, y_{ijt} = d) P_c(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j) \tag{6.4}$$

6.1.2 Price Negotiations Under a Ban on Off-Label Use

If the government bans off-label prescriptions, the negotiated drug price equilibrium is likely to be different. With the off-label ban, the choice probability of drug d by physician i for patient j who is diagnosed with an approved indication of drug d is

$$\begin{aligned} P_c\left(y_{ijt} = d | z_i, z_j, p^{ban}_{dt}, x_{dt}, I_j\right) &= q \, \frac{\exp\left(\alpha_d\left(z_i, z_j\right) - \beta p^{ban}_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d \overline{I}\right)}{\sum_{d' \in D^{on}_{on}} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p^{ban}_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} \overline{I}\right)} \\ &+ (1-q) \, \frac{\exp\left(\alpha_d\left(z_i, z_j\right) - \beta p^{ban}_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d \underline{I}\right)}{\sum_{d' \in D^{on}_{on}} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p^{ban}_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} \underline{I}\right)} \end{aligned} \tag{6.5}$$

where, as defined above, D^{on}_{on} is the set of all the on-label drugs for the on-label indication considered (depression) and p^{ban}_{dt} is the price of drug d negotiated with the regulator under the ban on off-label use. Note that because the reference group is 'Other Off-Label Drugs', with a ban on off-label drugs, the choice probability of the reference group is 0.

Then, we can simulate the new price equilibrium in the case of a ban on off-label prescriptions, assuming that the same bargaining model would apply. For any drug $d \in D^{on}_{on}$, its new price, p^{ban}_{dt}, would be the solution of

$$\max_{p^{ban}_{dt}} \left[\Pi^{ban}_{dt}\left(\mathbf{p}^{ban}_t\right)\right]^{\mu} \left[w \Delta_d CS^{ban}_t\left(\mathbf{p}^{ban}_t\right) - (1-w) \Delta_d TC^{ban}_{dt}\left(\mathbf{p}^{ban}_t\right)\right]^{1-\mu}$$

where $\Pi^{ban}_{dt}\left(\mathbf{p}^{ban}_t\right)$ is the profit of the firm for active ingredient d when off-label drugs competing with d are banned, and $\mathbf{p}^{ban}_t$ is the vector of prices of all the on-label drugs when off-label drugs are banned. $\Delta_d CS^{ban}_t\left(\mathbf{p}^{ban}_t\right) \equiv CS^{ban}_t\left(\mathbf{p}^{ban}_t\right) - CS^{ban}_{t,-d}\left(\mathbf{p}^{ban}_t\right)$ is the consumer surplus provided by drug d for the on-label indication where

$$CS^{ban}_t\left(\mathbf{p}^{ban}_t\right) = \frac{1}{\beta} \sum_{j \in J_{on}} \ln\left(\sum_{d' \in D^{on}_{on}} q \exp\left(\alpha_{d'}(z_i, z_j) - \beta p^{ban}_{d't} + \lambda_{d'} \overline{I}\right) + (1-q) \exp\left(\alpha_{d'}(z_i, z_j) - \beta p^{ban}_{d't} + \lambda_{d'} \underline{I}\right)\right)$$

and

$$CS_{t,-d}^{ban}\left(\mathbf{p}_t^{ban}\right)=\frac{1}{\beta}\sum_{j\in J_{on}}\ln\left(\sum_{d'\in D_{on}^{on}\setminus\{d\}}q\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}^{ban}+\lambda_{d'}\overline{I}\right)+(1-q)\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}^{ban}+\lambda_{d'}\underline{I}\right)\right)$$

The firm's profits are now

$$\Pi_{dt}^{ban}\left(\mathbf{p}_t^{ban}\right)=\left[p_{dt}^{ban}-c_{dt}\right]q_{dt}^{ban}\left(\mathbf{p}_t^{ban}\right)$$

where $q_{dt}^{ban}\left(p_t^{ban}\right)$ is the aggregate demand of drug d in the case of an off-label ban and marginal costs are estimated according to (4.3). Note that in the case of a ban, the market share of drug d is zero by definition in the off-label indication markets. Therefore, with the ban on off-label use, the aggregate demand of drug d can increase or decrease relative to the benchmark case where off-label use is allowed. The market share of drug d in the on-label market increases with the ban as there is no competition from off-label drugs in the on-label market for drug d (in the depression market). However, drug d will lose all sales in the off-label markets. List of off-label markets for antidepressants is provided in section 4.2.

The marginal impact of the drug d on the total cost of treatment is now

$$\Delta_d TC_{dt}^{ban}\left(\mathbf{p}_t^{ban}\right)=\sum_{\tilde{d}\in D_{on}^{on}}p_{\tilde{d}t}^{ban}q_{\tilde{d}t}^{ban}\left(\mathbf{p}_t^{ban}\right)-\sum_{\tilde{d}\in D_{on}^{on}\setminus\{d\}}p_{\tilde{d}t}^{ban}q_{\tilde{d}t}^{ban\,-d}\left(\mathbf{p}_t^{ban}\right)$$

which is the difference between the total cost of treatment when drug d is in the choice set and the total cost when drug d is absent. $q_{\tilde{d}t}^{ban\,-d}\left(\mathbf{p}_t^{ban}\right)$ is the demand for drug $\tilde{d}$, at equilibrium prices under the ban, when drug d is absent.

The first-order condition can be written as

$$\frac{\mu}{1-\mu}\frac{\partial\ln\Pi_{dt}^{ban}\left(\mathbf{p}_t^{ban}\right)}{\partial p_{dt}}+\frac{\partial\ln\left[w\Delta_d CS_t^{ban}\left(\mathbf{p}_t^{ban}\right)-(1-w)\Delta_d TC_{dt}^{ban}\left(\mathbf{p}_t^{ban}\right)\right]}{\partial p_{dt}}=\mathbf{0} \tag{6.6}$$

Using this system of first-order conditions, we can find the new prices as solutions of this system given the estimated marginal costs:

$$p_{dt}^{ban}=c_{dt}+\frac{1}{\left(\frac{-\partial\ln q_{dt}^{on}\left(\mathbf{p}_t^{ban}\right)}{\partial p_{dt}}\right)+\frac{1-\mu}{\mu}\left(\frac{-\partial\ln\left[w\Delta_d CS_t^{ban}\left(\mathbf{p}_t^{ban}\right)-(1-w)\Delta_d TC_{dt}^{ban}\left(\mathbf{p}_t^{ban}\right)\right]}{\partial p_{dt}}\right)}$$

where $\frac{\partial \ln q_{dt}^{on}\left(\mathbf{p}_t^{ban}\right)}{\partial p_{dt}}$ is the price semi-elasticity of demand at the new price equilibrium in the on-label market and $\frac{\partial \ln \Delta_d CS_t^{ban}\left(\mathbf{p}_t^{ban}\right)}{\partial p_{dt}}$ is the price semi-elasticity of the marginal consumer surplus by drug d under the ban, when off-label drugs are not in the choice set.

This estimation allows us to compute the counterfactual prescription probabilities of on-label drugs for each patient and the market shares given the new prices. We can then predict the associated counterfactual treatment expenses and recovery probabilities.

Note that price cost margin is a function of price elasticity of demand and price elasticity of drug d's additional value added to the consumer surplus at the new price equilibrium. It is also a function of price elasticity of drug d's marginal impact on total spending at the new prices. The ban can impact drug prices through its impact on these elasticities. As mentioned before, whether the aggregate demand for drug d will increase or decrease with the ban vs. without the ban is ambiguous because $q_{dt}^{tot}\left(p_t^{on}, p_t^{off}\right) = q_{dt}^{on}\left(p_t^{on}\right) + q_{dt}^{off}\left(p_t^{off}\right)$ is the sum of the sales of drug d on the on-label market and all the off-label markets for which drug d is used, whereas $q_{dt}^{on}\left(p_t^{ban}\right)$ is the on-label sales of drug d when no off-label competing drug is present to treat the on-label indication of drug d. Therefore, whether $\frac{\partial \ln q_{dt}^{on}\left(\mathbf{p}_t^{ban}\right)}{\partial p_{dt}}$ is larger or smaller than $\frac{\partial \ln q_{dt}^{tot}\left(\mathbf{p}_t^{on}, \mathbf{p}_t^{off}\right)}{\partial p_{dt}}$ is ambiguous. Semi-elasticities can be ranked in any direction depending on the price at which they are evaluated and depending on the price elasticity of drug d on the on-label indication market when off-label drugs are present and when they are absent and its elasticity in the off-label indication markets.

Similarly, the ban can impact the drug prices through the channel of the price elasticity of the consumer surplus value added of drugs. The consumer surplus value added of drug d is the marginal surplus provided by drug d relative to the other drugs in physicians' choice set. Physicians' choice set shrinks with the ban, and hence, the marginal surplus of drug d relative to the drugs in the new choice set increases; hence, $\Delta_d CS_t^{ban}\left(p_t^{ban}\right) > \Delta_d CS_t\left(p_t^{on}\right) > 0$. The order of derivatives of the consumer surplus value added of drug d with respect to price is such that $\frac{\partial \Delta_d CS_t^{ban}\left(\mathbf{p}_t^{ban}\right)}{\partial p_{dt}} < \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\partial p_{dt}} < 0$. Then, the semi-elasticity of $\Delta_d CS$ is larger with fewer drugs than with more drugs in the choice set, $\frac{\partial \ln \Delta_d CS_t(\mathbf{p}_t^{on})}{\partial p_{dt}} < \frac{\partial \ln \Delta_d CS_t^{ban}\left(\mathbf{p}_t^{ban}\right)}{\partial p_{dt}} < 0$ (proof is provided in the online appendix B.1). Therefore, the ban has an increasing impact on the price through the channel of consumer surplus elasticity.

In short, the overall impact of the ban on drug prices through these channels is ambiguous. It can increase drug prices if the effects go in the same direction, but it can also decrease the prices if the effect through demand elasticity is in the opposite direction of and dominates the effect through consumer surplus elasticity.

6.2 Empirical Results

6.2.1 Banning Off-Label Prescriptions while Keeping the Prices of On-Label Drugs Constant

In this section, we simulate counterfactual prescription probabilities, their expected costs and expected recovery of patients under a ban on off-label prescriptions, assuming that prices will be the same as in the benchmark scenario of no ban. Two cases are worth considering in the counterfactual analysis. In one of them, prescriptions of all off-label drugs are banned, and in the other case, prescriptions of 'Off-Label Drugs with Efficacy' are allowed, whereas prescriptions of off-label drugs for which there is no information in the medical literature on efficacy in depression treatment, 'Other Off-Label Drugs', are banned.

Table 6.1 reports the statistics on the expected cost of a one-day treatment for the cases when off-label prescriptions are allowed and for the two counterfactual scenarios. In both of the counterfactual scenarios, the cost increases, and the increase is larger when we ban all off-label drugs. The expected cost increases, from a benchmark of 76 euro cents per day, by 3.8% when we ban all off-label drugs, and by 1.7% when we ban only 'Other Off-Label Drugs' and still allow for prescriptions of 'Off-Label Drugs with Efficacy'.

Table 6.2 shows the expected recovery rate after a six-month treatment period after the first diagnosis. First, the probability of recovery after six months, conditional on the drug treatment, is calculated using the estimated parameters of the structural model. Then, given the estimated prescription probabilities, the unconditional recovery probability is calculated according to (6.3) in the benchmark and according to (6.4) in the counterfactuals, as the sum over all the alternatives of the recovery probability conditional on each alternative multiplied by the prescription probability of that alternative. Table 6.2 provides statistics on the unconditional recovery probability; the first column is the case when all drugs are in the choice set of the physicians; the second column shows the case when all approved drugs and off-label drugs with efficacy are in the choice set of the physicians; the last column shows the expected recovery probability when only approved drugs are in the choice set. The results show that removing the off-label drugs from the choice set does not have a very large impact on the recovery probability; recovery probability decreases by 1.4%.

These results show that keeping prices identical, banning off-label prescriptions has little impact on recovery rates for depression after a six-month treatment period but raises expenditures on drugs as physicians substitute off-label drugs with more expensive on-label alternatives.

Table 6.1: *Expected Prescription Expense of Drug Treatment (€)*

Choice Set		When All Drugs are in the Choice Set (Benchmark)	When 'Other Off-Label Drugs' are Removed		When All Off-Label Drugs are Removed	
Patients		Expenses	Expenses	Change	Expenses	Change
All	Mean	.760 [.754,.761]	.773 [.767,.776]	.013 (1.7%)	.789 [.787,.790]	.029 (3.8%)
	Min.	.684 [.681,.684]	.702 [.692,.712]	.018 (2.6%)	.754 [.753,.756]	.070 (10.2%)
	Max.	.814 [.807,.815]	.822 [.818,.823]	.008 (1.0%)	.830 [.827,.830]	.016 (2.0%)
Female	Mean	.759 [.752,.760]	.772 [.766,.775]	.013 (1.7%)	.788 [.786,.788]	.029 (3.8%)
Male	Mean	.763 [.757,.764]	.774 [.770,.777]	.011 (1.4%)	.793 [.791,.795]	.030 (3.9%)
Old	Mean	.779 [.772,.780]	.788 [.784,.790]	.009 (1.2%)	.799 [.797,.799]	.020 (2.6%)
Young	Mean	.750 [.743,.751]	.764 [.758,.768]	.014 (1.9%)	.784 [.782,.785]	.034 (4.5%)

Notes: Expected expense is the cost of one-day treatment, which is the sum across all the active ingredients, of price per day times the prescription probability for each active ingredient. The 'price per day' is the price per mg times mg per day according to the DDD for each active ingredient. Old patients are patients older than 60. Confidence intervals at the 90% confidence level constructed by 500 bootstrap draws drawn from the estimated distribution of parameters are in square brackets. Percentage change is reported in parentheses.

Table 6.2: *Expected Recovery Rate After Six Months*

Choice Set		When All Drugs are in the Choice Set (Benchmark)	When 'Other Off-Label Drugs' are Removed		When All Off-Label Drugs are Removed	
Patients		Recovery Probability	Recovery Probability	Change	Recovery Probability	Change
All	Mean	.502 [.418,.591]	.500 [.415,.590]	-.002 (-0.4%)	.495 [.414,.576]	-.007 (-1.4%)
	Min.	.275 [.190,.363]	.275 [.191,.363]	.000 (0.0%)	.274 [.190,.361]	-.001 (-0.4%)
	Max.	.776 [.716,.836]	.771 [.706,.835]	-.005 (-0.6%)	.751 [.706,.802]	-.025 (-3.2%)
Female	Mean	.493 [.407,.584]	.491 [.404,.584]	-.002 (-0.4%)	.486 [.402,.571]	-.007 (-1.4%)
Male	Mean	.523 [.445,.605]	.521 [.442,.604]	-.002 (-0.4%)	.515 [.441,.590]	-.008 (-1.5%)
Old	Mean	.419 [.329,.512]	.418 [.327,.511]	-.001 (-0.2%)	.415 [.326,.503]	-.004 (-1.0%)
Young	Mean	.549 [.470,.636]	.547 [.465,.635]	-.002 (-0.4%)	.540 [.464,.618]	-.009 (-1.6%)

Notes: Old patients are patients older than 60. Confidence intervals at the 90% significance level constructed by 500 bootstrap draws drawn from the estimated distribution of parameters are in square brackets. Percentage change is reported in parentheses.

6.2.2 Banning Off-Label Prescriptions and the New Price Equilibrium

Now, we investigate what prices of on-label drugs would be if the negotiations were to take place under a strict ban on off-label prescriptions. Then, we will identify the change in the expected cost of treatment and the change in the expected recovery rate, both of which also depend on new prices through the counterfactual choice probabilities. When off-label drugs are banned, the market share of on-label drugs in the depression market will increase. However, the on-label drugs for depression treatment are used off-label for the treatment of other diseases. Therefore, with an off-label ban, these drugs' market shares in the treatment of other diseases will decrease to zero even though their shares in the depression market will increase. The impact of an off-label ban on the aggregate sales of the drugs approved for depression is thus ambiguous. We therefore also estimate the sales of these drugs in the markets for the off-label indications they are used for. The list of off-label indications for antidepressants is provided in section 4.2.

After removing off-label drugs from the choice set, the counterfactual choice probability that physician i will prescribe an approved drug for which $l(d) = 1$, to patient j is given by (6.5). Then, the treatment cost for this on-label indication is

$$\sum\nolimits_{j \in J_{on}} \sum\nolimits_{d \in D_{on}^{on}} P_c \left(y_{ijt} = d | z_i, z_j, p_{dt}^{ban}, x_{dt}, I_j \right) p_{dt}^{ban}$$

We can now identify the change in the expected cost of treatment and, finally, the change in expected recovery rate that also depends on new prices through the counterfactual choice probabilities of alternatives.

Table 6.3 presents the differences between the prices in the benchmark scenario of no-ban and the prices in the counterfactual case of a ban estimated using the system of first-order conditions in (6.6) for different values of μ and w in the identified set. When off-label drugs are banned, we observe an increase in the prices of on-label alternatives in all six combinations of μ and w, with the exception of 'Citalopram' and 'Paroxetine' when $\mu = 0.6$ and $w = 1$ in which cases the price is not different than the benchmark scenario of no-ban. The increase in price depends on the drug and the values of μ and w and ranges from 0.4% to 34%.

Table 6.3: *Counterfactual Price Changes in the Case of a Ban on Off-Label Prescriptions*

	(1)	(2)	(3)	(4)	(5)	(6)
μ	0.60	0.65	0.65	0.70	0.70	0.75
w	1	0.80	0.90	0.70	0.80	0.70
Drugs						
Citalopram	-.001 (-0.1%)	.014 (1.7%)	.005 (0.6%)	.023 (2.7%)	.013 (1.6%)	.020 (2.4%)
	[-.173,.461]	[-.174,.508]	[-.182,.503]	[-.181,.551]	[-.190,.546]	[-.198,.587]
Sertraline	.003 (0.4%)	.004 (0.5%)	.005 (0.6%)	.006 (0.8%)	.006 (0.8%)	.007 (0.9%)
	[-.171,.456]	[-.181,.491]	[-.181,.494]	[-.193,.528]	[-.194,.530]	[-.205,.568]
Paroxetine	-.001 (-0.1%)	.008 (1.0%)	.005 (0.7%)	.014 (1.8%)	.010 (1.3%)	.015 (2.0%)
	[-.168,.461]	[-.173,.500]	[-.177,.502]	[-.181,.542]	[-.186,.542]	[-.194,.583]
Fluoxetine	.064 (9.6%)	.124 (18.0%)	.096 (14.1%)	.152 (21.9%)	.125 (18.2%)	.152 (22.0%)
	[-.145,.602]	[-.088,.670]	[-.125,.664]	[-.070,.724]	[-.106,.721]	[-.091,.776]
Escitalopram	.048 (7.1%)	.133 (19.8%)	.099 (14.8%)	.165 (24.5%)	.137 (20.4%)	.169 (25.1%)
	[-.178,.634]	[-.086,.694]	[-.135,.702]	[-.060,.743]	[-.105,.759]	[-.083,.813]
Other On-Label	.112 (12.4%)	.234 (25.9%)	.181 (20.0%)	.296 (32.8%)	.246 (27.2%)	.310 (34.4%)
	[-.150,.773]	[-.010,.852]	[-.087,.856]	[.049,.924]	[-.027,.935]	[.032,1.014]

Notes: μ is the bargaining parameter of the firm and w is the weight on CS. Confidence intervals at the 90% significance level constructed by 500 bootstrap draws drawn from the estimated distribution of parameters are in square brackets. Percentage change is reported in parentheses.

Table 6.4 provides a comparison of the expected prescription expenditure on drugs (expected cost of treatment), for given values of $\mu = 0.65$ and $w = 0.80$, in the three cases: when all drugs are in the choice set; and for the two counterfactual scenarios: when off-label drugs are banned and drug prices are assumed to be constant and when off-label drugs are banned and drug prices are negotiated under the ban. The expected

prescription expense of drug treatment increases by 24% in the counterfactual scenario when drug prices are allowed to be different under the ban, which is higher than the increase in the case when drug prices are kept constant, 3.7%. When off-label drugs are banned, physicians substitute towards on-label drugs and because, on average, on-label drugs are more expensive, this substitution effect leads to an increase in prescription expenses. When off-label drugs are banned and prices are negotiated under the ban, the prices of on-label drugs, on average, increase, and hence, the prescription expenses increase due to both price and substitution effects. Therefore, in the second column of Table 6.4 the increase in prescription expenses is due to the substitution effect whereas the increase in the third column is due to both substitution and price effects.

Table 6.4: *Expected Prescription Expense of Drug Treatment (€) (when $\mu = 0.65, w = 0.80$)*

Choice Set		When All Drugs are in the Choice Set (Benchmark)	When All Off-Label Drugs are Removed and Prices are Constant		When All Off-Label Drugs are Removed (with Counterfactual Prices)	
Patients		Expenses	Expenses	Change	Expenses	Change
All	Mean	.760 [.754,.761]	.789 [.787,.790]	.029 (3.8%)	.945 [.712,1.527]	.185 (24.3%)
	Min.	.684 [.681,.684]	.754 [.753,.756]	.070 (10.2%)	.892 [.666,1.461]	.208 (30.4%)
	Max.	.814 [.807,.815]	.830 [.827,.830]	.016 (2.0%)	1.009 [.768,1.603]	.195 (24.0%)
Female	Mean	.759 [.752,.760]	.788 [.786,.788]	.029 (3.8%)	.943 [.711,1.524]	.184 (24.2%)
Male	Mean	.763 [.757,.764]	.793 [.791,.795]	.030 (3.9%)	.951 [.717,1.536]	.188 (24.6%)
Old	Mean	.779 [.772,.780]	.799 [.797,.799]	.020 (2.6%)	.960 [.726,1.545]	.181 (23.2%)
Young	Mean	.750 [.743,.751]	.784 [.782,.785]	.034 (4.5%)	.937 [.705,1.517]	.187 (24.9%)

Notes: See the notes in Table 6.1.

Similarly, Table 6.5 reports a comparison of the probability of recovery six months after the first diagnosis in the same three cases. The decrease in the average recovery rate in the counterfactual scenario when drug prices are allowed to be different, 0.2%, is smaller than the decrease when drug prices are kept constant. Table 6.4 and 6.5 report the counterfactual results for given values of $\mu = 0.65$ and $w = 0.80$, the counterfactual results for the other values of μ and w are reported in the online appendix B.3. For all the values μ and w the expected cost of treatment increases with the ban on off-label drugs and for none of the values the ban leads to an improvement in the health outcome.

7 Conclusion

Using a unique dataset that provides longitudinal information over nine years on a sample of physicians, their office visits and all their patients, we develop a structural model of demand and supply on off-label drug use. On the demand side, we estimate a model of prescription behavior with potential unobserved patient-level heterogeneity, which is allowed to be correlated with treatment choices and treatment outcomes and estimate

Table 6.5: *Expected Recovery Rate After Six Months*

Choice Set		When All Drugs are in the Choice Set (Benchmark)	When All Off-Label Drugs are Removed and Prices are Constant		When All Off-Label Drugs are Removed (with Counterfactual Prices)	
Patients		Recovery Probability	Recovery Probability	Change	Recovery Probability	Change
All	Mean	.502 [.418,.591]	.495 [.414,.576]	-.007 (-1.4%)	.500 [.414,.580]	-.002 (-0.4%)
	Min.	.275 [.190,.363]	.274 [.190,.361]	-.001 (-0.4%)	.281 [.200,.366]	.006 (2.2%)
	Max.	.776 [.716,.836]	.751 [.706,.802]	-.025 (-3.2%)	.747 [.695,.799]	-.029 (-3.7%)
Female	Mean	.493 [.407,.584]	.486 [.402,.571]	-.007 (-1.4%)	.491 [.403,.575]	-.002 (-0.4%)
Male	Mean	.523 [.445,.605]	.515 [.441,.590]	-.008 (-1.5%)	.519 [.440,.593]	-.004 (-0.8%)
Old	Mean	.419 [.329,.512]	.415 [.326,.503]	-.004 (-1.0%)	.421 [.331,.508]	.002 (0.5%)
Young	Mean	.549 [.470,.636]	.540 [.464,.618]	-.009 (-1.6%)	.544 [.462,.621]	-.005 (-0.9%)

Notes: See the notes in Table 6.2.

the demand for on-label and off-label drugs for depression treatment. The results show that there is significant patient-level heterogeneity impacting both the treatment choice and the treatment outcome. We find that, on average, treatment outcomes with off-label drugs are not worse than the treatment outcomes with on-label alternatives. On the supply side, we develop a Nash-in-Nash bargaining model between the government and pharmaceutical firms at which prices are determined by Nash equilibrium of bilateral Nash bargaining problems.

We perform counterfactual analysis by simulating the demand, supply and the associated treatment outcomes in the case of a ban on off-label use. We consider two cases of counterfactual scenarios. In the first one, we assume that drug prices are the same in the case of a ban on off-label drug prescriptions. In the second counterfactual scenario, we allow drug prices to be different under a ban by allowing the ban to impact the bargaining outcome. We then simulate the associated counterfactual expected demand, treatment cost, and treatment outcome using the counterfactual equilibrium prices. The results suggest that banning off-label prescriptions would lead to an increase in the cost of prescription drugs due to a substitution effect in the first counterfactual scenario and due to both a substitution effect and to higher prices in equilibrium in the second counterfactual scenario. In both cases, banning off-label prescriptions does not lead to significant changes in terms of health outcomes.

This analysis shows that regulatory decisions concerning physicians' ability to be flexible in their prescription behavior regarding the label status of drugs may be an important factor in determining welfare and health expenditures. We have shown, in the context of France, that banning off-label prescriptions would not negatively affect health outcomes on average, but would significantly increase drug expenditures. The argument that a strict enforcement of label-status-based prescriptions can improve health outcomes by more

strictly controlling physicians' behavior does not seem to be valid in this context. Moreover, it would lead to higher expenses. However, we do not analyze the behavior of pharmaceutical companies in terms of demands for approval of different indications. Their behavior would possibly be different in the counterfactual environment in which they would anticipate the impossibility of physicians to use off-label prescriptions. Analysing this in the counterfactual environment would require modeling the investment decisions in clinical trials and regulatory applications for approval as a determinant of profit maximization. We leave this interesting question for future research.

References

Ameli (2019): Drug database of the French Social Security Health Care System (Base de Médicaments et Informations Tarifaires), Available at http://www.codage.ext.cnamts.fr.

Andrews, D. W. K. and P. J. Barwick (2012): "Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure," *Econometrica*, 80, 2805–2826.

Bradford, D. W., J. L. Turner, and J. W. Williams (2018): "Off-Label Use Of Pharmaceuticals: A Detection Controlled Estimation Approach," *The Journal of Industrial Economics*, 66, 866–903.

Bücheler, R., M. Schwab, K. Mörike, B. Kalchthaler, H. Mohr, H. Schröder, P. Schwoerer, and C. H. Gleiter (2002): "Off-Label Prescribing to Children in Primary Care in Germany: Retrospective Cohort Study," *BMJ*, 324, 1311–1312.

Chalumeau, M., J. Tréluyer, B. Salanave, R. Assathiany, G. Chéron, N. Crocheton, C. Rougeron, M. Mares, G. Bréart, and G. Pons (2000): "Off-Label and Unlicensed Drug Use Among French Office-Based Pediatricians," *Archives of Disease in Childhood*, 83, 502–505.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243–1284.

Cohen, J., L. Faden, S. Predaris, and B. Young (2007): "Patient Access to Pharmaceuticals: An International Comparison," *The European Journal of Health Economics : HEPAC : Health Economics in Prevention and Care*, 8, 253—266.

Collard-Wexler, A., G. Gowrisankaran, and R. S. Lee (2019): "Nash-in-Nash Bargaining: A Microfoundation for Applied Work," *Journal of Political Economy*, 127, 163–195.

Crawford, G. S. and M. Shum (2005): "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73, 1137–1173.

Crawford, G. S. and A. Yurukoglu (2012): "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, 102, 643–685.

Deb, P. and P. Trivedi (2006): "Specification and Simulated Likelihood Estimation of a Non-normal Treatment-Outcome Model with Selection: Application to Health Care Utilization," *Econometrics Journal*, 9, 307–331.

Decarolis, F., M. Polyakova, and S. Ryan (2020): "Subsidy Design in Privately-Provided Social Insurance: Lessons from Medicare Part D," *Journal of Political Economy*, 128, 1712–1752.

Dickstein, M. (2018): "Efficient Provision of Experience Goods: Evidence from Antidepressant Choice," *Working Paper*.

Dubois, P. and L. Lasio (2018): "Identifying Industry Margins with Price Constraints: Structural Estimation on Pharmaceuticals," *American Economic Review*, 108, 3685–3724.

Dubois, P. and M. Sæthre (2020): "On the Effect of Parallel Trade on Manufacturers' and Retailers' Profits in the Pharmaceutical Sector," *Econometrica*, 88, 2503–2545.

Dubois, P. and T. Tunçel (2020): "Identifying the Effects of Scientific Information and Recommendations on Physicians Prescribing Behavior," *TSE Working Paper*.

Eudralex (2015): "The Rules Governing Medicinal Products in the European Union," Vol. 2A: Procedures for Marketing Authorization, Chapter 1: Marketing Authorization. European Commission Health and Food Safety Directorate-General.

Gallini, A., J. M. Donohue, and H. A. Huskamp (2013): "Diffusion of Antipsychotics in the U.S. and French Markets, 1998–2008," *Psychiatric Services*, 64, 680–687.

Gowrisankaran, G., A. Nevo, and R. Town (2015): "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry," *American Economic Review*, 105, 172–203.

Grennan, M. (2013): "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *American Economic Review*, 103, 145–177.

Ho, K. and R. S. Lee (2017): "Insurer Competition in Health Care Markets," *Econometrica*, 85, 379–417.

Horn, H. and A. Wolinsky (1988): "Bilateral Monopolies and Incentives for Merger," *The RAND Journal of Economics*, 19, 408–419.

IMS (2001-2008): Health Press Releases.

Kilbourne, A. M., K. Beck, B. Spaeth-Rublee, P. Ramanuj, R. W. O'Brien, N. Tomoyasu, and H. A. Pincus (2018): "Measuring and Improving the Quality of Mental Health Care: A Global Perspective," *World Psychiatry*, 17, 30–38.

Licht-Strunk, E., H. Van Marwijk, T. Hoekstra, J. Twisk, M. De Haan, and A. Beekman (2009): "Outcome of Depression in Later Life in Primary Care: Longitudinal Cohort Study with Three Years' Follow-up," *BMJ*, 338, a3079.

Lucarelli, C., J. Prince, and K. Simon (2012): "The Welfare Impact of Reducing Choice in Medicare Part D: A Comparison of Two Regulation Strategies," *International Economic Review*, 53, 1155–1177.

Martin-Latry, K., C. Ricard, and H. Verdoux (2007): "A One-Day Survey of Characteristics of Off-Label Hospital Prescription of Psychotropic Drugs," *Pharmacopsychiatry*, 40, 116–120.

MICROMEDEX (2016): Healthcare Series, Greenwood Village, Colo: MICROMEDEX.

OECD (2017): "Pharmaceutical Spending (Indicator) and Health Spending (Indicator)," Available at data.oecd.org.

Perlis, R. (2014): "Pharmacogenomic Testing and Personalized Treatment of Depression," *Clinical Chemistry*, 60, 53–59.

Radley, D. C., S. Finkelstein, and R. Stafford (2006): "Off-Label Prescribing Among Office-Based Physicians," *Archives of Internal Medicine*, 166, 1021–1026.

Rodwin, V. (2003): "The Health Care System Under French National Health Insurance: Lessons for Health Reform in the United States," *American Journal of Public Health*, 93, 31–37.

Shapiro, B. (2018): "Informational Shocks, Off-Label Prescribing and the Effects of Physician Detailing," *Management Science*, 64, 5925–5945.

Simon, G. and R. Perlis (2010): "Personalized Medicine for Depression: Can We Match Patients with Treatments?" *American Journal of Psychiatry*, 167, 1445–1455.

Small, K. and H. Rosen (1981): "Applied Welfare Economics of Discrete Choice Models," *Econometrica*, 49, 105–130.

Stafford, R. (2008): "Regulating Off-Label Drug Use-Rethinking the Role of the FDA," *New England Journal of Medicine*, 358, 1427–1429.

Uher, R., K. Tansey, K. Malki, and R. Perlis (2012): "Biomarkers Predicting Treatment Outcome in Depression: What Is Clinically Significant?" *Pharmacogenomics*, 13, 233–240.

Weisman, H. and B. Healy (1987): "Myocardial Infarct Expansion, Infarct Extension, and Reinfarction: Pathophysiologic Concepts," *Progress in Cardiovascular Disease*, XXX, 73–110.

A Appendix

A.1 Details on Aggregated Choice Alternatives

Table A.1 provides details on the share of drugs aggregated under the choice alternatives 'Other Off-Label Drugs', 'Off-Label Drugs with Efficacy', and 'Other On-Label Drugs'. For instance, among 'Other Off-Label Drugs', 81% are nervous system drugs. Among them, psycholeptics have the highest share, 91%, and among psycholeptics, anxiolytics have the highest share, 56%. Among all the off-label drugs, considering both *'Other Off-Label Drugs'* and *'Off-Label Drugs with Efficacy',* 84% are nervous system drugs.

Table A.1: *Drug Classification of Off-Label and 'Other On-Label' Drugs*

	Percentage		
Drugs in Alternative 'Other Off-Label Drugs'			
Nervous System (N)	81%		
Psycholeptics (N05)		91%	
Antipsychotics (N05A)			6%
Anxiolytics (N05B)			56%
Hypnotics and Sedatives (N05C)			38%
Analgesics (N02)		5%	
Antiepileptics (N03)		2%	
Other Nervous System		2%	
Alimentary Tract and Metabolism (A)	9%		
Other	10%		
Drugs in Alternative 'Off-Label Drugs with Efficacy'			
Alprazolam (N05BA12)	87%		
Buspirone (N05BE01)	7%		
Olanzapine (N05AH03)	5%		
Other	1%		
Drugs in Alternative 'Other On-Label Drugs'			
Other Antidepressants (N06AX)	82%		
Non-Selective Monoamine Reuptake Inhibitors (N06AA)	12%		
Other	6%		

Note: ATC codes are in parentheses.

A.2 Clarification on the Definition of Off-Label Prescriptions

As mentioned before, in this study, off-label use refers to the use of a drug for an indication for which the drug has not received approval. In the data, we observe which diagnoses are made at each office visit for a given patient. We also observe which drug is prescribed for each diagnosis. Using the three cases below as examples of possible diagnosis-prescription pairs at a given office visit, we explain how off-label prescriptions are determined in this study and how we were conservative in determining them.

In case 1, the patient is diagnosed with depression and alcoholism. A drug that is approved for depression is prescribed for the depression diagnosis (on-label drug for depression), and another drug is prescribed for the alcoholism diagnosis. In the analysis, the patient in case 1 is considered a depression patient who is being treated with an on-label drug. In case 2, the patient is diagnosed only with depression (the patient does not have an alcoholism diagnosis). A drug that is not approved for depression but is approved for alcoholism is prescribed for the depression diagnosis (off-label drug for depression). In the analysis, the patient in case 2 is considered a depression patient who is being treated with an off-label drug. In case 3, the patient is diagnosed with depression and alcoholism. A drug that is not approved for depression but approved for alcoholism is prescribed for the depression diagnosis, and another drug is prescribed for the alcoholism diagnosis. In this case, there are two possibilities: it could be that the physician prescribes an off-label drug to treat depression. The second possibility is that the physician believes that depression is a secondary condition to alcoholism, and he prescribes two drugs to treat alcoholism and does not prescribe any drug for depression and considers that depression will go away once alcoholism is treated. To be conservative in determining off-label prescriptions, in the analysis we exclude cases such as case three. These cases correspond to only 0.98% of off-label prescriptions and 0.21% of all the prescriptions. Determining these cases requires that for every drug not approved but prescribed for depression, we check for which indications the drug is approved and whether the patient is diagnosed with any of these indications.

- Case 1:
  - Diagnosis: Depression → Prescription: A drug approved for depression (on-label drug for depression)
  - Diagnosis: Alcoholism → Prescription: A drug for alcoholism
- Case 2:
  - Diagnosis: Depression → Prescription: A drug not approved for depression (off-label depression drug), which is approved for alcoholism (on-label alcoholism drug)
  - No Alcoholism Diagnosis

- Case 3:
  - Diagnosis: Depression → Prescription: A drug not approved for depression (off-label depression drug), which is approved for alcoholism (on-label alcoholism drug)
  - Diagnosis: Alcoholism → Prescription: A drug for alcoholism

A.3 Graphical Example of Banning Off-Label Drug Prescriptions

Figure A.1 shows an example of three drugs that are approved for different indications but can be used off-label for non-approved indications and what happens when off-label prescriptions are banned. In this example, only drug A is approved for depression, B is approved for alcoholism and C is for epilepsy. Drugs B and C are used off-label in the treatment of depression. Drug A is also used off-label to treat alcoholism. When off-label prescriptions are banned, drugs B and C can no longer be used in depression treatment; hence, the market share of drug A in the depression market will increase with the ban. However, drug A can no longer be used in alcoholism treatment; hence, all the sales of drug A in the alcoholism market will disappear with the ban. How aggregate demand for drug A changes with the ban depends on how much its demand will increase in the depression market and how much it loses in the off-label markets, i.e., the alcoholism market in Figure A.1.

A.4 Dropouts

Table A.2 shows the percentage of patients who stopped visiting their physicians among all the patients and among patients prescribed approved vs. off-label drugs. We call these patients 'dropouts'; however, one has to keep in mind that the objective of the data is to follow physicians, not patients, over time. Therefore, when a patient changes her generalist (for instance, because of moving to another place), she is no longer in the dataset. It would be worrying if dropout rates were different among patients receiving approved drugs than among those receiving off-label drugs. The share of patients who stopped visiting their physician among patients who were prescribed approved drugs is not much different than the share among patients who were prescribed off-label drugs.

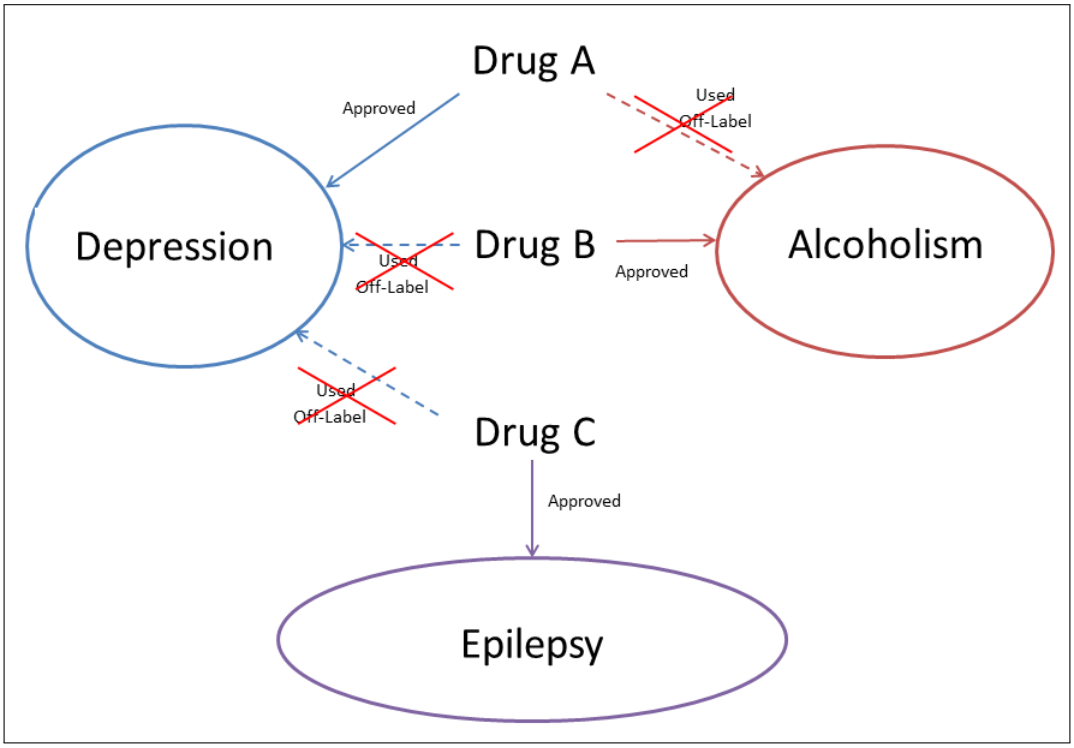


Figure A.1: *Example of Banning Off-Label Drug Prescriptions*

Table A.2: *Dropout Rates among Patients*

Among All Patients	24.57%
Among Patients that are Prescribed:	
On-Label Drugs	24.18%
Off-Label Drugs	26.03%
Off-Label Drugs with Efficacy	26.30%
Other Off-Label Drugs	25.98%

B Appendix for Online Publication

B.1 Proof of Change in the Semi-Elasticity of CS When Off-Label Drugs are Banned

The semi-elasticity of CS before the ban is

$$\frac{\partial \ln[\Delta_d CS_t(\mathbf{p}_t^{on})]}{\partial p_{dt}} = \frac{1}{\Delta_d CS_t(\mathbf{p}_t^{on})} \frac{\partial \Delta_d CS_t(\mathbf{p}_t^{on})}{\partial p_{dt}}$$

where $\Delta_d CS_t(\mathbf{p}_t^{on}) \equiv CS_t(\mathbf{p}_t^{on}) - CS_{t,-d}(\mathbf{p}_t^{on})$.

The semi-elasticity of CS after the ban is

$$\frac{\partial \ln[\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})]}{\partial p_{dt}} = \frac{1}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})} \frac{\partial \Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}{\partial p_{dt}}$$

The CS of all drugs in the choice set before the ban is

$$CS_t(\mathbf{p}_t^{on}) = \frac{1}{\beta} \sum_{j \in J_{on}} \ln \left(\sum_{d' \in D^{on}} q \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \lambda_{d'} \overline{I}\right) + (1-q) \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \lambda_{d'} \underline{I}\right) \right)$$

and the CS of all drugs excluding drug d before the ban is

$$CS_{t,-d}(\mathbf{p}_t^{on}) = \frac{1}{\beta} \sum_{j \in J_{on}} \ln \left(\sum_{d' \in D^{on} \setminus \{d\}} q \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \lambda_{d'} \overline{I}\right) + (1-q) \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \lambda_{d'} \underline{I}\right) \right)$$

We will first show how the marginal surplus of drug d changes with the ban. Then, we will show how its derivative changes with the ban.

Let us denote the mean utility of drug d by A such that

$$A \equiv q \exp\left(\alpha_d(z_i, z_j) - \beta p_{dt} + \lambda_d \overline{I}\right) + (1-q) \exp\left(\alpha_d(z_i, z_j) - \beta p_{dt} + \lambda_d \underline{I}\right)$$

and the sum of mean utilities of all the drugs before the ban other than drug d by B such that

$$B \equiv \sum_{d' \in D^{on} \setminus \{d\}} \left(q \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \lambda_{d'} \overline{I}\right) + (1-q) \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \lambda_{d'} \underline{I}\right) \right)$$

Then,

$$\begin{aligned}\Delta_d CS_t(\mathbf{p}_t^{on}) &\equiv CS_t(\mathbf{p}_t^{on}) - CS_{t,-d}(\mathbf{p}_t^{on}) \\ &= \frac{1}{\beta}\left[\ln(A+B) - \ln(B)\right]\end{aligned}$$

The marginal surplus of drug d after the ban is

$$\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban}) \equiv CS_t^{ban}(\mathbf{p}_t^{ban}) - CS_{t,-d}^{ban}(\mathbf{p}_t^{ban})$$

where

$$CS_t^{ban}(\mathbf{p}_t^{ban}) = \frac{1}{\beta}\sum_{j\in J_{on}} \ln\left(\sum_{d'\in D_{on}^{on}} q\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't}^{ban} + \lambda_{d'}\overline{I}\right) + (1-q)\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't}^{ban} + \lambda_{d'}\underline{I}\right)\right)$$

and

$$CS_{t,-d}^{ban}(\mathbf{p}_t^{ban}) = \frac{1}{\beta}\sum_{j\in J_{on}} \ln\left(\sum_{d'\in D_{on}^{on}\setminus\{d\}} q\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't}^{ban} + \lambda_{d'}\overline{I}\right) + (1-q)\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't}^{ban} + \lambda_{d'}\underline{I}\right)\right)$$

Let us denote the sum of the mean utilities of all on-label drugs other than drug d, in the on-label market for drug d, which are the drugs in set $D_{on}^{on}\setminus\{d\}$, by C such that

$$C \equiv \sum_{d'\in D_{on}^{on}\setminus\{d\}} \left(q\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't} + \lambda_{d'}\overline{I}\right) + (1-q)\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't} + \lambda_{d'}\underline{I}\right)\right)$$

Then,

$$\begin{aligned}\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban}) &\equiv CS_t^{ban}(\mathbf{p}_t^{ban}) - CS_{t,-d}^{ban}(\mathbf{p}_t^{ban}) \\ &= \frac{1}{\beta}\left[\ln(A+C) - \ln(C)\right]\end{aligned}$$

Note that $B > C$.

Then,

$$\begin{aligned}\Delta_d CS_t(\mathbf{p}_t^{on}) - \Delta_d CS_t^{ban}(\mathbf{p}_t^{ban}) &= \frac{1}{\beta}\left[\ln(A+B) - \ln(B)\right] - \frac{1}{\beta}\left[\ln(A+C) - \ln(C)\right] \\ &= \frac{1}{\beta}\left[\ln(A+B) - \ln(B) - \ln(A+C) + \ln(C)\right] \leq 0\end{aligned}$$

This comes from $|\ln(A+B)-\ln(A+C)| \leq |\ln(C)-\ln(B)|$ since the log is a concave function and $B > C$.

Thus, $\Delta_d CS_t(\mathbf{p}_t^{on}) \leq \Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})$, meaning that the marginal contribution of drug d to a larger set is smaller than the one to a smaller set, with everything else equal. ΔCS_d increases when going from a larger to a smaller set, that is, when off-label drugs are banned.

Now, let us look at the derivative of the marginal surplus of drug d with respect to p_d before the ban. We have

$$\begin{aligned}\frac{\partial \Delta_d CS_t(\mathbf{p}_t^{on})}{\partial p_{dt}} &= \frac{\partial}{\partial p_d}\left[\frac{1}{\beta}\left[\ln(A+B)-\ln(B)\right]\right] \\ &= \frac{\partial}{\partial p_d}\left[\frac{1}{\beta}\ln(A+B)\right] = \frac{1}{\beta}\frac{-\beta A}{A+B} = \frac{-A}{A+B}\end{aligned}$$

The derivative of the marginal surplus of drug d with respect to p_d after the ban is

$$\begin{aligned}\frac{\partial \Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}{\partial p_{dt}} &= \frac{\partial}{\partial p_d}\left[\frac{1}{\beta}\left[\ln(A+C)-\ln(C)\right]\right] \\ &= \frac{\partial}{\partial p_d}\left[\frac{1}{\beta}\ln(A+C)\right] = \frac{1}{\beta}\frac{-\beta A}{A+C} = \frac{-A}{A+C}\end{aligned}$$

We thus have

$$\begin{aligned}&\frac{\partial \ln[\Delta_d CS_t(\mathbf{p}_t^{on})]}{\partial p_{dt}} - \frac{\partial \ln[\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})]}{\partial p_{dt}} \\ =\ & \underbrace{\frac{1}{\Delta_d CS_t(\mathbf{p}_t^{on})}}_{>0}\underbrace{\frac{-A}{A+B}}_{<0} - \underbrace{\frac{1}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}}_{>0}\underbrace{\frac{-A}{A+C}}_{<0} \\ =\ & \frac{1}{\Delta_d CS_t(\mathbf{p}_t^{on})}\left[\frac{-A}{A+B} - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}\frac{-A}{A+C}\right] \\ =\ & \frac{-A}{\Delta_d CS_t(\mathbf{p}_t^{on})}\left[\frac{1}{A+B} - \frac{\frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}}{A+C}\right] \\ =\ & \frac{-A}{\Delta_d CS_t(\mathbf{p}_t^{on})}\left[\frac{(A+C)-\frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}(A+B)}{(A+B)(A+C)}\right] \\ =\ & \underbrace{\frac{-A}{\Delta_d CS_t(\mathbf{p}_t^{on})(A+B)(A+C)}}_{<0}\left[\left(1-\frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}\right)(A+C) - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}(B-C)\right] \\ =\ & \underbrace{\frac{-A}{\Delta_d CS_t(\mathbf{p}_t^{on})(A+B)(A+C)}}_{<0}\left[(A+C) - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}A - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}B\right]\end{aligned}$$

because $\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban}) > \Delta_d CS_t(\mathbf{p}_t^{on})$ and $A > 0$, $B > C > 0$.

We cannot determine the sign of $\left[(A+C) - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}A - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}B\right]$ by knowing simply

$$0 < \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})} < 1$$

However, $\frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}$ is related to A, B and C such that

$$\frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})} = \frac{\ln(A+B) - \ln(B)}{\ln(A+C) - \ln(C)}$$

Then,

$$\left[(A+C) - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}A - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}B\right] = \left[(A+C) - \frac{\ln(A+B) - \ln(B)}{\ln(A+C) - \ln(C)}(A+B)\right]$$

We can check that $\frac{(A+C)}{(A+B)} > \frac{\ln(A+B)-\ln(B)}{\ln(A+C)-\ln(C)}$ for all $A > 0$, $B > C > 0$.

Hence

$$\begin{aligned} & \frac{\partial \ln[\Delta_d CS_t(\mathbf{p}_t^{on})]}{\partial p_{dt}} - \frac{\partial \ln[\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})]}{\partial p_{dt}} \\ = & \underbrace{\frac{-A}{\Delta_d CS_t(\mathbf{p}_t^{on})(A+B)(A+C)}}_{<0} \left[\underbrace{(A+C) - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}A - \frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}B}_{>0}\right] < 0 \end{aligned}$$

Thus, the semi-elasticity of $\Delta_d CS$ is smaller without the ban (with more drugs in the choice set) than with the ban (with fewer drugs in the choice set). Because both values are negative, we have

$$\frac{\partial \ln[\Delta_d CS_t(\mathbf{p}_t^{on})]}{\partial p_{dt}} < \frac{\partial \ln[\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})]}{\partial p_{dt}} < 0$$

Because the marginal CS of drug d becomes less elastic, the ban increases the drug prices through this channel.

B.2 Additional Descriptive Statistics and Robustness Checks for Health Outcomes

Table B.1 reports the relationship between the prescription on the first visit the patient is diagnosed with depression and the majority of the prescriptions the patient receives in the six-month treatment period. In

total, 96% (87%) of the patients who received an on-label (off-label) drug on the first visit are also prescribed an on-label (off-label) drug the majority of the time during the six-month period.

Table B.1: *Prescription at First Diagnosis vs. Majority of Prescriptions (Six-Month Period)*

	Majority of Prescriptions		Equal
First Visit	On-Label	Off-Label	Share
On-Label	96%	1%	3%
Off-Label	7%	87%	6%

Notes: The element in the first row of the first column is the share of patients who received an on-label drug for the majority of the time during the treatment period among the patients who received an on-label drug on the day of the first-diagnosis.

Table B.2 reports the same statistics as in Table B.1 for a one-year treatment period.

Table B.2: *Prescription at First Diagnosis vs. Majority of Prescriptions (One-Year Period)*

	Majority of Prescriptions		Equal
First Visit	On-Label	Off-Label	Share
On-Label	95%	2%	3%
Off-Label	10%	83%	7%

Notes: See the notes in Table B.1.

Table B.3 provides the relationship between the prescription on the first visit the patient is diagnosed with depression and the majority of the prescriptions the patient receives in the six-month treatment period at the level of choice alternatives considered in the estimations. For instance, 93.5% of the patients who received Citalopram on the first visit also received Citalopram the majority of the time they are prescribed a drug in the six-month period, whereas 1.2% of them received Sertraline the majority of the time they are prescribed a drug.

Table B.3: *Prescription at First Diagnosis vs. Majority of Prescriptions (Six-Month Period)*

	Majority of Prescriptions							
First Visit	Citalop.	Sertral.	Paroxet.	Fluoxet.	Escitalop.	Other On-Lab.	Off-Lab. w. Eff.	Other Off-Lab.
Citalopram	93.5%	1.2%	0.9%	0.6%	0.5%	2.0%	0.1%	1.0%
Sertraline	0.6%	94.2%	0.9%	0.7%	0.4%	1.8%	0.1%	1.3%
Paroxetine	0.9%	0.7%	93.1%	1.1%	0.1%	2.4%	0.2%	1.3%
Fluoxetine	0.9%	0.7%	0.7%	94.1%	0.4%	1.9%	0.1%	1.1%
Escitalopram	0.3%	0.5%	0.6%	0.6%	94.4%	2.1%	0.2%	1.3%
Other On-Label	1.0%	0.8%	1.6%	1.2%	0.7%	93.4%	0.2%	1.1%
Off-Label w. Efficacy	2.4%	1.7%	4.7%	2.6%	1.4%	5.8%	79.4%	1.9%
Other Off-Label	1.3%	1.4%	3.1%	1.9%	0.9%	4.0%	0.5%	86.9%

Notes: The element in the first row of the first column is the share of patients who received a Citalopram prescription for the majority of the time during the treatment period among the patients who received a Citalopram prescription on the day of the first-diagnosis. Similarly, the element in the first row of the second column is the share of patients who received a Sertraline prescription for the majority of the time during the treatment period among the patients who received a Citalopram prescription on the day of the first-diagnosis.

Table B.4 reports the same statistics as Table B.3 for a one-year treatment period.

Table B.4: *Prescription at First Diagnosis vs. Majority of Prescriptions (One-Year Period)*

	Majority of Prescriptions							
First Visit	Citalop.	Sertral.	Paroxet.	Fluoxet.	Escitalop.	Other On-Lab.	Off-Lab. w. Eff.	Other Off-Lab.
Citalopram	89.8%	1.7%	1.3%	1.0%	1.0%	3.0%	0.3%	1.8%
Sertraline	0.7%	91.2%	1.3%	1.1%	0.4%	2.9%	0.2%	2.1%
Paroxetine	1.3%	0.9%	90.1%	1.8%	0.6%	3.1%	0.2%	2.0%
Fluoxetine	1.3%	1.3%	1.3%	90.6%	0.5%	2.8%	0.3%	1.9%
Escitalopram	0.4%	0.8%	0.7%	0.8%	91.5%	3.2%	0.4%	2.1%
Other On-Label	1.7%	1.1%	2.2%	1.5%	1.0%	90.2%	0.3%	2.1%
Off-Label w. Efficacy	3.4%	2.2%	5.7%	2.9%	1.5%	6.5%	74.8%	2.9%
Other Off-Label	1.8%	1.7%	3.7%	2.5%	1.2%	5.2%	0.6%	83.3%

Notes: See the notes in Table B.3.

Table B.5 reports the binary probit estimation of the treatment outcome, recovery after six months, taking into account the 'majority of prescriptions' during the treatment period. The results are not different from the recovery estimation when we focus on the prescription on the first visit of the depression diagnosis.

Table B.5: *Binary Probit Estimation of Treatment Outcomes*

	Parameter	Std. Error	Marginal Effect	Std. Error
Patients' Age	-0.01	(0.00)	-0.00	(0.00)
Patients' Sex (female=1)	-0.06	(0.02)	-0.02	(0.00)
Constant	1.64	(0.03)		
If majority of prescriptions during the treatment period is:				
On-Label Drugs				
Citalopram	-0.29	(0.03)	-0.10	(0.01)
Sertraline	-0.31	(0.03)	-0.10	(0.01)
Paroxetine	-0.29	(0.02)	-0.10	(0.01)
Fluoxetine	-0.30	(0.03)	-0.10	(0.01)
Escitalopram	-0.32	(0.03)	-0.11	(0.01)
Other	-0.36	(0.02)	-0.12	(0.01)
Off-Label Drugs				
with Efficacy	-0.01	(0.05)	-0.00	(0.02)
Number of Observations			37,521	

Notes: Standard errors are in parentheses.

One may be concerned that the lower recovery rates with on-label drugs are due to the aggregation of off-label drugs under one choice alternative. To check if this is the case, we disaggregate the choice alternative 'Other Off-Label Drugs' and estimate the binary probit treatment outcome model with more alternatives for off-label drugs. The off-label active ingredients we disaggregate from the alternative 'Other Off-Label Drugs' are the ones that have the highest share among 'Other Off-Label Drugs': 'Bromazepam', 'Etifoxine', 'Valeriane',

'Zolpidem', 'Prazepam', and 'Zopiclone'. Table B.6 reports the estimation in which the reference category is all the off-label drugs excluding 'Off-Label with Efficacy' and 'Bromazepam', 'Etifoxine', 'Valeriane', 'Zolpidem', 'Prazepam', and 'Zopiclone'. The results show that on-label active ingredients lead to lower recovery rates than the reference category and also lower than 'Off-Label with Efficacy' and all the other off-label active ingredients. We observe that recovery rates are higher with 'Bromazepam', 'Etifoxine', and 'Valeriane' than with 'Other Off-Label Drugs', and recovery rates with 'Zolpidem', 'Prazepam', and 'Zopiclone' are the same as recovery rates with 'Other Off-Label Drugs'. The results show that lower recovery rates with on-label drugs are not due to aggregation of off-label drugs under one choice alternative. Note that in this estimation, the unobserved patient state is not taken into account; hence, the results in Table B.6 should be considered as a robustness check for the results in Table 5.2.

Table B.7 presents the percentage of patients who are no longer diagnosed with depression six months after they are diagnosed with depression for the first time, separately for each choice alternative considered in the estimations. We observe that recovery rates are higher with off-label drugs.

Table B.6: *Binary Probit Estimation of Treatment Outcomes*

	Parameter	Std. Error	Marginal Effect	Std. Error
Patients' Age	-0.01	(0.00)	0.00	(0.00)
Patients' Sex	-0.08	(0.01)	-0.03	(0.01)
Constant	1.08	(0.03)		
On-Label Drugs				
Citalopram	-0.22	(0.03)	-0.09	(0.01)
Sertraline	-0.21	(0.03)	-0.08	(0.01)
Paroxetine	-0.21	(0.03)	-0.08	(0.01)
Fluoxetine	-0.21	(0.03)	-0.08	(0.01)
Escitalopram	-0.24	(0.04)	-0.09	(0.02)
Other	-0.26	(0.03)	-0.10	(0.01)
Off-Label Drugs				
with Efficacy	0.06	(0.04)	0.02	(0.02)
Other Off-Label Drugs				
Bromazepam	0.08	(0.05)	0.03	(0.02)
Etifoxine	0.37	(0.06)	0.13	(0.02)
Valeriane	0.38	(0.05)	0.14	(0.02)
Zolpidem	-0.03	(0.07)	-0.01	(0.03)
Prazepam	-0.00	(0.08)	-0.00	(0.03)
Zopiclone	-0.10	(0.08)	-0.04	(0.03)
Number of Observations		37,510		

Notes: Standard errors are in parentheses.

Table B.8 presents the percentage of patients who are no longer diagnosed with depression one year after they are diagnosed with depression for the first time.

Table B.7: *Recovery Rates - Six Months after the First Diagnosis*

	No Depression Diagnosis in Six-month Period	No Depression Diagnosis in One-year Period	No Depression Diagnosis Anytime
Among Patients who are Prescribed:			
Citalopram	63%	57%	46%
Sertraline	64%	59%	47%
Paroxetine	64%	58%	47%
Fluoxetine	64%	58%	46%
Escitalopram	64%	57%	51%
Other On-Label	61%	54%	44%
Off-Label Drugs with Efficacy	76%	71%	59%
Other Off-Label Drugs	76%	70%	59%

Notes: For the 'one-year' and 'anytime' observation periods, second and third columns, the recovery rates are lower because in these cases, some of the patients have another cycle of depression (relapse cases).

Table B.8: *Recovery Rates - One Year after the First Diagnosis*

	No Depression Diagnosis in Six-month Period	No Depression Diagnosis in One-year Period	No Depression Diagnosis Anytime
Among Patients who are Prescribed:			
Citalopram	71%	64%	52%
Sertraline	72%	66%	54%
Paroxetine	71%	64%	53%
Fluoxetine	71%	64%	52%
Escitalopram	68%	62%	59%
Other On-Label	67%	61%	51%
Off-Label Drugs with Efficacy	79%	73%	64%
Other Off-Label Drugs	79%	74%	64%

Notes: Notes: See the notes in Table B.7.

B.3 Counterfactual Results for Different Values of Bargaining Power and Weight

Table B.9 provides the change in the expected prescription expenditure on drugs (expected cost of treatment) for the counterfactual scenario when off-label drugs are banned and drug prices are negotiated under the ban, for all the values of μ and w other than $\mu = 0.65$ and $w = 0.80$ which is provided in the text. Similarly, Table B.10 shows the change in the probability of recovery six months after the first diagnosis for the same counterfactual scenario. For all the values of μ and w, the expected cost of treatment increases with the ban on off-label drugs and the ban does not lead to an improvement in the average recovery rates.

Table B.9: *Change in Expected Prescription Expense of Drug Treatment (€)*
(with Counterfactual Prices)

	μ	0.60	0.65	0.70	0.70	0.75
	w	1	0.90	0.70	0.80	0.70
Patients						
All	Mean	.100 (13.2%)	.149 (19.7%)	.223 (29.3%)	.190 (25.0%)	.228 (30.0%)
	Min.	.130 (19.1%)	.174 (25.6%)	.242 (35.3%)	.212 (31.0%)	.246 (35.9%)
	Max.	.099 (12.2%)	.154 (19.0%)	.240 (29.5%)	.202 (24.9%)	.247 (30.4%)
Female	Mean	.099 (13.1%)	.147 (19.4%)	.222 (29.2%)	.189 (24.9%)	.227 (29.9%)
Male	Mean	.102 (13.4%)	.151 (19.9%)	.228 (29.8%)	.193 (25.3%)	.233 (30.5%)
Old	Mean	.095 (12.3%)	.144 (18.6%)	.222 (28.5%)	.188 (24.2%)	.227 (29.1%)
Young	Mean	.104 (14.0%)	.151 (20.2%)	.224 (29.8%)	.192 (25.6%)	.230 (30.6%)

Notes: Expected expense is the cost of one-day treatment, which is the sum across all the active ingredients, of the price per day times the prescription probability for each active ingredient. The 'price per day' is the price per mg times mg per day according to the DDD for each active ingredient. Percentage change is reported in parentheses. Old patients are patients older than 60.

Table B.10: *Change in Expected Recovery Rate After Six Months*
(with Counterfactual Prices)

	μ	0.60	0.65	0.70	0.70	0.75
	w	1	0.90	0.70	0.80	0.70
Patients						
All	Mean	-.003 (-0.6%)	-.003 (-0.6%)	-.002 (-0.4%)	-.002 (-0.4%)	-.002 (-0.4%)
	Min.	.006 (2.2%)	.006 (2.2%)	.007 (2.5%)	.006 (2.2%)	.007 (2.5%)
	Max.	-.029 (-3.7%)	-.029 (-3.7%)	-.029 (-3.7%)	-.029 (-3.7%)	-.028 (-3.6%)
Female	Mean	-.002 (-0.4%)	-.002 (-0.4%)	-.001 (-0.2%)	-.001 (-0.2%)	-.001 (-0.2%)
Male	Mean	-.005 (-1.0%)	-.004 (-0.8%)	-.003 (-0.6%)	-.004 (-0.8%)	-.003 (-0.6%)
Old	Mean	.002 (0.5%)	.002 (0.5%)	.003 (0.7%)	.003 (0.7%)	.003 (0.7%)
Young	Mean	-.006 (-1.1%)	-.005 (-0.9%)	-.004 (-0.7%)	-.005 (-0.9%)	-.004 (-0.7%)

Notes: Percentage change is reported in parentheses. Old patients are patients older than 60.

B.4 Test for Seasonality in Drug Advertisement

Table B.11 reports regression results when we regress the stock of advertisement for each drug choice across the set of monthly dummies and a linear annual time trend. The omitted month is January. According to the joint test of significance of the coefficients of month dummies, we cannot reject the null hypothesis that they are jointly equal to zero. The t-statistics for coefficients of months support the same result that the stock of advertisement is not statistically different from each other across different months. We would still like to mention that the drug advertisement seems to be lower in August, though not statistically significant for most drugs. This is likely to be related to the nationwide vacation period in August in France rather than being correlated with unobserved patient heterogeneity.

Table B.11: *Test for Seasonality in Drug Advertisement*

	Dependent Variable: Stock of Advertisement for:							
	Citalop.	Sertraline	Paroxetine	Fluoxetine	Escitalop.	Other On-Lab.	Off-Lab. w. Eff.	Other Off-Lab.
Feb.	.01 (0.01)	.09 (0.15)	-.10 (-0.09)	-.01 (-0.01)	.01 (0.03)	.03 (0.25)	-.09 (-0.13)	.03 (0.24)
Mar.	.03 (0.07)	.01 (0.01)	-.19 (-0.17)	.01 (0.02)	.12 (0.52)	.08 (0.69)	.54 (0.77)	.09 (0.68)
Apr.	-.12 (-0.29)	-.14 (-0.25)	-.33 (-0.29)	-.03 (-0.10)	.14 (0.62)	.07 (0.57)	.56 (0.80)	.07 (0.54)
May	-.29 (-0.69)	-.23 (-0.39)	-.47 (-0.42)	-.09 (-0.28)	.15 (0.64)	.06 (0.56)	.47 (0.68)	.03 (0.21)
June	-.42 (-0.97)	-.13 (-0.23)	-.53 (-0.47)	-.15 (-0.48)	.22 (0.96)	.09 (0.76)	.29 (0.42)	.03 (0.26)
July	-.62 (-1.45)	-.33 (-0.54)	-.73 (-0.65)	-.28 (-0.88)	.15 (0.64)	.01 (0.04)	.13 (0.18)	-.09 (-0.65)
Aug.	-.89 (-2.07)	-.54 (-0.90)	-1.82 (-1.63)	-.54 (-1.70)	-.09 (-0.39)	-.21 (-1.86)	-.06 (-0.09)	-.32 (-2.40)
Sep.	-.53 (-1.22)	-.31 (-0.52)	-1.79 (-1.60)	-.52 (-1.64)	.06 (0.25)	-.11 (-0.96)	-.16 (-0.23)	-.14 (-1.01)
Oct.	-.56 (-1.29)	-.35 (-0.58)	-.45 (-0.41)	-.49 (-1.54)	.13 (0.58)	-.04 (-0.33)	-.17 (-0.24)	-.02 (-0.14)
Nov.	-.71 (-1.66)	-.49 (-0.82)	-.56 (-0.50)	-.48 (-1.53)	.25 (1.15)	-.03 (-0.23)	-.35 (-0.50)	.01 (0.10)
Dec.	-.74 (-1.73)	-.66 (-1.10)	-.55 (-0.50)	-.63 (-1.99)	.24 (1.10)	-.07 (-0.60)	-.54 (-0.78)	-.03 (-0.24)
Joint Test of Significance of Month Dummies								
F-stat	1.12	0.30	0.55	1.23	0.43	1.19	0.52	1.40
Prob¿F	0.36	0.98	0.86	0.28	0.92	0.31	0.88	0.19

Notes: t-statistics are in parentheses. Specifications include linear annual time trend and a constant term.

B.5 Market Shares and Demand Estimation for Prescriptions Across All Visits

Table B.12 reports the market shares of eight choice alternatives among the prescriptions on the first visit and among the prescriptions across all the visits. Table B.13 reports the demand estimation including all the visits and shows similar impact of price and advertisement as in the case of demand estimation of prescriptions only on the first visit.

Table B.12: *Market Shares Across First Visit vs. All the Visits*

	Market Shares Across:	
	First Visit	All Visits
Citalopram	10.5%	8.8%
Sertraline	8.6%	7.1%
Paroxetine	18.7%	16.0%
Fluoxetine	12.6%	13.3%
Escitalopram	5.1%	3.6%
Other On-Lab.	23.8%	29.3%
Off-Lab. w. Eff.	3.0%	2.7%
Other Off-Lab.	17.7%	19.3%

Table B.13: *Logit Estimation of Treatment Choice Including all the Visits*

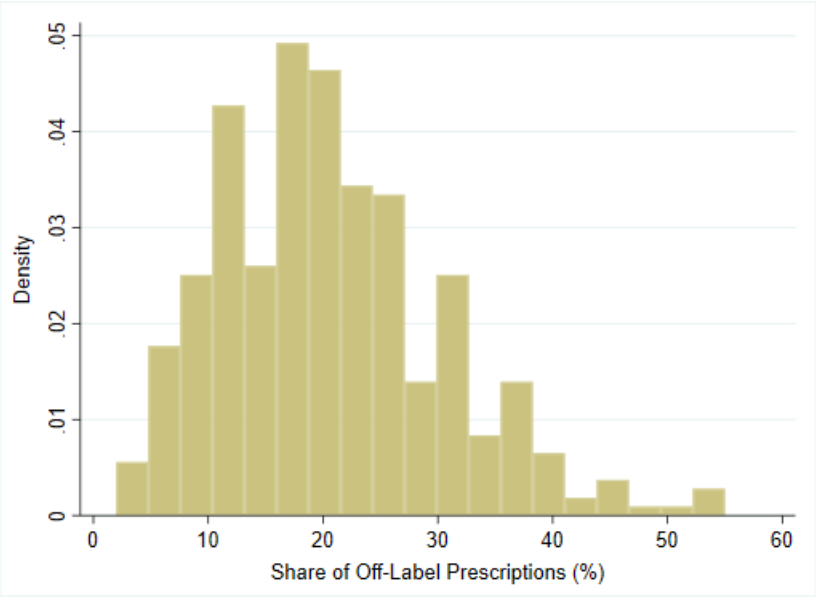
	Alternative Specific Parameters				
	Patients'		Physicians'		
	Age	Sex	Age	Sex	Constant
On-Label Drugs:					
Citalopram	-0.009	0.131	-0.015	0.041	-1.002
	(0.001)	(0.014)	(0.001)	(0.016)	(0.131)
Sertraline	-0.018	0.002	-0.010	-0.231	-0.910
	(0.001)	(0.014)	(0.001)	(0.019)	(0.131)
Paroxetine	-0.009	0.091	-0.001	-0.085	-1.175
	(0.001)	(0.011)	(0.001)	(0.014)	(0.126)
Fluoxetine	-0.010	0.301	-0.003	-0.080	-1.649
	(0.001)	(0.013)	(0.001)	(0.015)	(0.125)
Escitalopram	-0.024	0.070	0.001	-0.267	-1.915
	(0.001)	(0.021)	(0.001)	(0.027)	(0.139)
Other	0.001	-0.061	-0.003	-0.015	-1.534
	(0.001)	(0.009)	(0.001)	(0.012)	(0.124)
Off-Label Drugs:					
with Efficacy	-0.021	-0.067	-0.004	0.006	-0.516
	(0.001)	(0.022)	(0.001)	(0.027)	(0.082)
Parameters Common Across Alternatives					
Price ($-\beta$)	Advertising*On-Label			Advertising*Off-Label	
-0.511	0.143			-0.001	
(0.088)	(0.004)			(0.007)	
Observations	415,380				

Notes: Advertising is the natural logarithm of the stock of advertising. The "Sex" variable is 1 for females and 0 for males. Standard errors are in parentheses.

B.6 Heterogeneity in Physicians' Off-Label Prescriptions

Figure B.1 shows the distribution of physicians in terms of their share of off-label prescriptions. For each physician, the share of off-label prescriptions across all their prescriptions for depression treatment is calculated and the figure reports these shares across all the physicians. Majority of the physicians prescribe an off-label drug between 15-25% of the time they prescribe a drug for depression. Therefore, the average number of off-label prescriptions across all prescriptions, 21%, is not caused by some physicians very aggressively prescribing off-label drugs and others prescribing very little. All physicians prescribe off-label drugs and a mass of them prescribe an off-label drug around 20% of the time they prescribe a drug for depression.

Figure B.1: *Distribution of Within Physician Share of Off-Label Prescriptions*



B.7 Elasticities

B.7.1 Price Elasticity of Demand

The market share of drug d at price $\mathbf{p}_t$ on the on-label market for drug d, $s_{dt}(\mathbf{p}_t)$, which is the average choice probability of drug d across all the patients diagnosed with the on-label indication, is

$$s_{dt}(\mathbf{p}_t) = \frac{1}{J} \sum_{j \in J_{on}} P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j\right)$$

The formula for demand elasticities are traditional for logit. The own and cross-price elasticities of the market share s_{dt} on the on-label market for drug d are

$$\frac{\partial \ln s_{dt}}{\partial \ln p_{kt}} = \begin{cases} -\beta \frac{p_{kt}}{s_{dt}} \frac{1}{J} \sum_{j \in J_{on}} \left\{ qP\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \overline{I}\right) \left[1 - P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \overline{I}\right)\right] \right. \\ \left. +(1-q) P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \underline{I}\right) \left[1 - P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \underline{I}\right)\right] \right\} \end{cases} \text{if } d = k$$

$$\frac{\partial \ln s_{dt}}{\partial \ln p_{kt}} = \begin{cases} \beta \frac{p_{kt}}{s_{dt}} \frac{1}{J} \sum_{j \in J_{on}} \left\{ qP\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \overline{I}\right) P\left(y_{ijt} = k | z_i, z_j, p_{dt}, x_{dt}, \overline{I}\right) \right. \\ \left. +(1-q) P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \underline{I}\right) P\left(y_{ijt} = k | z_i, z_j, p_{dt}, x_{dt}, \underline{I}\right) \right\} \end{cases} \text{if } d \neq k$$

Table B.14 reports the own- and cross-price elasticities of demand for all the choice alternatives.

B.7.2 Advertisement Elasticity of Demand

The own- and cross-advertisement elasticities of the market share s_{dt} on the on-label market for drug d are

$$\frac{\partial \ln s_{dt}}{\partial x_{kt}} = \begin{cases} \gamma_{l(d)} \frac{1}{s_{dt}} \frac{1}{J} \sum_{j \in J_{on}} \{ qP\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \overline{I}\right) \left[1 - P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \overline{I}\right)\right] \\ \quad + (1-q) P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \underline{I}\right) \left[1 - P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \underline{I}\right)\right] \} \end{cases} \quad \text{if } d = k$$

$$\frac{\partial \ln s_{dt}}{\partial x_{kt}} = \begin{cases} -\gamma_{l(d)} \frac{1}{s_{dt}} \frac{1}{J} \sum_{j \in J_{on}} \{ qP\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \overline{I}\right) P\left(y_{ijt} = k | z_i, z_j, p_{dt}, x_{dt}, \overline{I}\right) \\ \quad + (1-q) P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, \underline{I}\right) P\left(y_{ijt} = k | z_i, z_j, p_{dt}, x_{dt}, \underline{I}\right) \} \end{cases} \quad \text{if } d \neq k$$

where x_{kt} is the natural logarithm of stock of detailing expenditures.

Note that in the estimation, $\gamma_{l(d)}$ differs across on-label and off-label drugs; hence, γ is γ_{on} if $l(k) = 1$ and γ_{off} if $l(k) = 0$.

Table B.15 reports the own- and cross-advertisement elasticities of demand for all the choice alternatives.

Table B.14: *Own- and Cross-Price Elasticities*

	Citalopram	Sertraline	Paroxetine	Fluoxetine	Escitalopram	Other On-Lab.	Off-Lab. w. Eff.	Other Off-Lab.
Citalopram	-.7780	.0430	.0360	.0069	.0074	.0086	.0059	.0065
Sertraline	.0898	-.6012	.1376	.0062	.0088	.0140	.0023	.0046
Paroxetine	.1059	.1948	-.5429	.0136	.0162	.0221	.0090	.0115
Fluoxetine	.0931	.0408	.0628	-.4721	.1452	.1403	.1508	.1497
Escitalopram	.1168	.0671	.0868	.1689	-.5595	.1621	.1756	.1715
Other On-Lab.	.3364	.2624	.2925	.4028	.4003	-.5473	.3923	.3978
Off-Lab. w. Eff.	.0136	.0025	.0070	.0254	.0254	.0230	-.4704	.0276
Other Off-Lab.	.0020	.0007	.0012	.0034	.0034	.0032	.0038	-.1640

Note: Each cell is the price elasticity of demand for the drug in the row with respect to the price of the drug in the column.

Table B.15: *Own- and Cross-Advertisement Elasticities*

	Citalopram	Sertraline	Paroxetine	Fluoxetine	Escitalopram	Other On-Lab.	Off-Lab. w. Eff.	Other Off-Lab.
Citalopram	.2413	-.0134	-.0112	-.0021	-.0023	-.0027	-.0018	-.0020
Sertraline	-.0288	.1925	-.0441	-.0020	-.0028	-.0045	-.0007	-.0015
Paroxetine	-.0374	-.0689	.1919	-.0048	-.0057	-.0078	-.0032	-.0041
Fluoxetine	-.0374	-.0164	-.0253	.1898	-.0584	-.0564	-.0606	-.0602
Escitalopram	-.0400	-.023	-.0297	-.0578	.1915	-.0555	-.0601	-.0587
Other On-Lab.	-.0879	-.0686	-.0764	-.1053	-.1046	.1430	-.1025	-.1039
Off-Lab. w. Eff.	-.0012	-.0002	-.0006	-.0022	-.0022	-.0020	.0408	-.0024
Other Off-Lab.	-.0005	-.0002	-.0003	-.0009	-.0009	-.0008	-.0010	.0424

Note: Each cell is the advertisement elasticity of demand for the drug in the row with respect to the advertisement of the drug in the column.

B.8 Demand Estimation and Counterfactual Results for Different Values of Types

This section reports the results of counterfactual analysis using the estimates of the joint estimation of treatment choice and treatment outcome with different values of unobserved health state I_j. When $\overline{I}$ and $\underline{I}$ are normalized to different values the estimate of share of high type, q, is also different. However, the results concerning the selection into treatment based on unobservables, relative treatment impact of drugs and counterfactual simulations are robust across different values of types. Tables B.16 and B.18 report the statistics on the expected cost of a one-day treatment, given the values of I_j reported on the tables, for the cases when off-label prescriptions are allowed and for the two counterfactual scenarios: when prescriptions of all off-label drugs are banned, and when prescriptions of 'Off-Label Drugs with Efficacy' are allowed, whereas prescriptions of 'Other Off-Label Drugs' are banned. Similar to the results in Table 6.1, in both of the counterfactual scenarios, the cost of treatment increases, and the increase is larger when we ban all off-label drugs. Tables B.17 and B.19 show the expected recovery rate after a six-month treatment period after the first diagnosis, for the benchmark case and for the two counterfactual scenarios, for the values of I_j reported on the tables. Similar to the results in Table 6.2, removing the off-label drugs from the choice set has a small decreasing impact on the recovery outcomes.

Table B.16: *Expected Prescription Expense of Drug Treatment (€) (When $\overline{I} = 0.5$, $\underline{I} = -0.5$)*

Choice Set		When All Drugs are in the Choice Set (Benchmark)	When 'Other Off-Label Drugs' are Removed		When All Off-Label Drugs are Removed	
Patients		Expenses	Expenses	Change	Expenses	Change
All	Mean	.764	.771	.007 (0.9%)	.787	.023 (3.0%)
	Min.	.693	.704	.011 (1.6%)	.753	.060 (8.7%)
	Max.	.814	.818	.004 (0.5%)	.826	.012 (1.5%)
Female	Mean	.763	.770	.007 (0.9%)	.785	.022 (2.9%)
Male	Mean	.766	.772	.006 (0.8%)	.790	.024 (3.1%)
Old	Mean	.780	.786	.006 (0.8%)	.796	.016 (2.1%)
Young	Mean	.754	.763	.009 (1.2%)	.782	.028 (3.7%)

Notes: See the notes in Table B.9.

Table B.17: *Expected Recovery Rate After Six Months (When $\overline{I} = 0.5$, $\underline{I} = -0.5$)*

Choice Set		When All Drugs are in the Choice Set (Benchmark)	When 'Other Off-Label Drugs' are Removed		When All Off-Label Drugs are Removed	
Patients		Recovery Probability	Recovery Probability	Change	Recovery Probability	Change
All	Mean	.513	.512	-.001 (-0.2%)	.506	-.007 (-1.4%)
	Min.	.280	.281	.001 (0.4%)	.280	.000 (0.0%)
	Max.	.783	.781	-.002 (-0.3%)	.762	-.021 (-2.7%)
Female	Mean	.504	.503	-.001 (-0.2%)	.498	-.006 (-1.2%)
Male	Mean	.534	.533	-.001 (-0.2%)	.527	-.007 (-1.3%)
Old	Mean	.429	.428	-.001 (-0.2%)	.426	-.003 (-0.7%)
Young	Mean	.561	.559	-.002 (-0.4%)	.553	-.008 (-1.4%)

Notes: See the notes in Table B.10.

Table B.18: *Expected Prescription Expense of Drug Treatment (€) (When $\overline{I} = 1.5$, $\underline{I} = -1.5$)*

Choice Set		When All Drugs are in the Choice Set (Benchmark)	When 'Other Off-Label Drugs' are Removed		When All Off-Label Drugs are Removed	
Patients		Expenses	Expenses	Change	Expenses	Change
All	Mean	.763	.771	.008 (1.0%)	.789	.026 (3.4%)
	Min.	.684	.696	.012 (1.8%)	.751	.067 (9.8%)
	Max.	.817	.822	.005 (0.6%)	.830	.013 (1.6%)
Female	Mean	.762	.770	.008 (1.0%)	.787	.025 (3.3%)
Male	Mean	.765	.772	.007 (0.9%)	.793	.028 (3.7%)
Old	Mean	.781	.787	.006 (0.8%)	.799	.018 (2.3%)
Young	Mean	.753	.761	.008 (1.1%)	.783	.030 (4.0%)

Notes: See the notes in Table B.9.

Table B.19: *Expected Recovery Rate After Six Months (When $\overline{I} = 1.5$, $\underline{I} = -1.5$)*

Choice Set		When All Drugs are in the Choice Set (Benchmark)	When 'Other Off-Label Drugs' are Removed		When All Off-Label Drugs are Removed	
Patients		Recovery Probability	Recovery Probability	Change	Recovery Probability	Change
All	Mean	.516	.514	-.002 (-0.4%)	.509	-.007 (-1.4%)
	Min.	.287	.287	.000 (0.0%)	.286	-.001 (-0.3%)
	Max.	.784	.781	-.003 (-0.4%)	.759	-.025 (-3.2%)
Female	Mean	.507	.506	-.001 (-0.2%)	.500	-.007 (-1.4%)
Male	Mean	.537	.536	-.001 (-0.2%)	.529	-.008 (-1.5%)
Old	Mean	.432	.432	.000 (0.0%)	.429	-.003 (-0.7%)
Young	Mean	.563	.562	-.001 (-0.2%)	.554	-.009 (-1.6%)

Notes: See the notes in Table B.10.