Should We Prevent Off-Label Drug Prescriptions? Empirical Evidence from France

Tuba Tunçel*

October 2019†

Abstract

Regulation of off-label prescriptions, when a physician prescribes a drug that is not approved for the diagnosed disease, differs across countries as it raises questions on its costs and benefits. Evaluating the role of off label prescriptions needs to account for its effect on physician behavior, patients health, treatment costs and pharmaceutical firms pricing, which I ravel out with a structural model of demand and supply. Exploiting rich panel data on physicians' activities and all their office visits over a nine-year period, I first study the decision to prescribe approved vs. off-label drugs in depression treatment. Using a model of prescription choice and health outcome with unobserved patient-level heterogeneity, I identify the demand for on-label and off-label drugs as well as the effect of prescription choice on health outcome. I find that there is significant patient-level heterogeneity impacting both the treatment choice and the treatment outcome. The results show that off-label drugs, prescribed in 21% of cases, are not worse than on-label alternatives in the treatment of depression. On the supply side, I develop a Nash-in-Nash bargaining model between the government and the pharmaceutical companies that allows me to identify the marginal cost of drugs and simulate the new equilibrium outcome that would occur if the bargaining on drug prices were taking place under a strict ban on off-label prescriptions. Counterfactual simulations show that removing off-label drugs from the choice set of physicians, while keeping prices of on-label drugs constant, would lead to a substantial increase in the expenditure on prescription drugs due to substitution effect. Allowing drug prices to be different under a ban on off-label prescriptions would further increase the treatment cost due to both the substitution and price effects 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.

1 Introduction

Regulatory agencies approve drugs for the specific indications requested by pharmaceutical companies, not for general use, and the approval process requires substantial evidence of efficacy and safety. 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]. This practice is very common. It can happen that a good candidate drug for the treatment of a disease is not approved for it, as is the case for Avastin in the treatment of age-related macular degeneration (AMD). Avastin is approved for cancer treatment, and there is scientific evidence that it is effective for the treatment of AMD; however, it is not approved in Europe for treatment of AMD. The average price of Avastin in Europe is 40 euros per injection, whereas the approved drug for AMD treatment (Lucentis) is around 900 euros per injection. Another example is the off-label use of acetylsalicylic acid, famously known as Aspirin, to prevent secondary myocardial infarction (heart attack). Preliminary evidence began to come out in the 1960s and 1970s that aspirin could lower the risk of a second heart attack. However, aspirin was already a cheap generic drug back then, as its patent had expired in 1917 and no company had the incentive to conduct clinical trials to prove that aspirin prevents heart attacks. It was thanks to government financing of such studies, and only in 1985, that aspirin was approved for secondary myocardial infarction (Weisman and Healy, 1987). From the point the preliminary evidence started to emerge until the approval date, over decades, aspirin was used off-label as a heart attack preventive[2].

In the pharmaceutical market, which accounts for 12% of total health spending in the U.S. (roughly 2.1% of US GDP) and 14% of total health spending in France (1.6% of French GDP) (OECD, 2017), off-label use is very common. The most comprehensive study of off-label prescriptions, using nationally representative data, finds that among the 160 most commonly prescribed drugs in the US, off-label prescriptions account for approximately 21% of overall use (Radley et. al., 2006). 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

[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 for an indication the drug has never received approval. To avoid conflicts with patient population-based off-label use, I 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/.

inpatients in a single hospital on a single day (Martin-Latry et. al. 2007).

Regulation of off-label use differs across countries; some countries permit off-label use, some partially restrict it, and some others restrict it through 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 using 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.

The data we use are from France during 2000 to 2008 which is a period where there was no restriction on off-label prescriptions in France. Physicians were free to choose among drugs regardless of their label status. While there is an extensive literature on the determinants of physicians' treatment decisions, studies of physicians' choices between approved and off-label alternatives are sparse. Bradford et al. (2015) provide an important first step documenting off-label prescription rates and identifying some determinants of such decisions. Shapiro (2016) investigates the impact of detailing on physicians' prescriptions of on-label versus off-label use. Bradford et al. (2015) analyses the share of off-label prescriptions of a given drug, that is, the frequency with which a drug is prescribed for its approved indications (on-label) versus for other indications (off-label). Instead, we analyze the share of off label prescriptions conditional on a diagnostic, that is the propensity of the physician to prescribe an off label treatment conditional on the diagnosed disease, namely depression[3].

This unique French dataset provides longitudinal information over nine years on a representative sample of physicians and all their patients. The longitudinal dimension of the prescription data allows following patients' visits to their physician and thus observing treatment outcomes. Hence, we can investigate whether treatment outcomes are different for patients treated with approved drugs than 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 with patient-level unobserved heterogeneity allowed to be correlated with treatment outcome, we

[3]The existing literature on off-label use provides information on which drugs have the highest number off-label prescriptions. Our analysis is based on depression treatment for which off-label use is indeed prevalent among French general practitioners.

identify the demand for on-label and off-label drugs in depression treatment. We separately identify the impact of treatment choice from the impact of patients' unobserved state on treatment outcome using detailing (advertising) expenditures that affects prescription decisions of physicians and can credibly be excluded from the health outcome equation conditional on the treatment chosen. The results show there is significant patient level unobserved heterogeneity impacting both the treatment choice and the treatment outcome, and that off-label drugs are not worse than the on-label alternatives in terms of probability to recover. Then, counterfactual simulations using the structural model are performed.

Current regulations question policies that would restrict or fully prevent off-label prescriptions. In France, the current system adopted since 2011 aims at strictly regulating off-label prescriptions with "Temporary Recommendations for Use (TRU)", while in the US, the formulary drug lists of health insurance companies contribute to limiting off-label prescriptions. In order to shed light on the best regulation of off-label use, we thus evaluate the counterfactual effects of a ban on off-label prescriptions. Using the structural model parameters estimates, we can first evaluate the impact of such ban on choice probabilities, treatment outcomes, and costs of prescription drugs, keeping the prices of drugs fixed. The counterfactual simulations show that banning off-label prescriptions would lead to a substantial increase in the cost of treatment because of substitution to more expensive products, whereas it would not lead to substantial improvement in terms of health outcomes. Then, we 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 (Grennan, 2003) between the government and pharmaceutical companies allows to estimate 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 between consumer surplus and 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 government. We show how we can set identify the bargaining parameters and the weights the government puts on consumer surplus versus cost of treatment thanks to marginal cost restrictions. 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 price elasticity of demand and of the drug's additional value added in consumer surplus. The ban can impact drug prices through both of these channels; through its impact on demand and consumer surplus elasticities. An approved drug for depression treatment is not only used in its on-label market (depression) but also in markets of indications for which it is not approved, the off-label markets. Similarly, a drug approved for depression will face competition of 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 as, now, the drug will not face any competition in this market from off-label drugs. However, the drug will lose all the sales in the off-label markets where it would be used when there is no-ban. Therefore, whether the aggregate demand for the drug will increase or decrease with the ban relative to no-ban is ambiguous. Hence, the impact of the ban on the price of a drug through the channel of price elasticity can go in any direction depending 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 its elasticity in the off-label indication markets. We thus also estimate the demand of 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 consumer surplus of drugs. The added value of a drug on 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 consumer surplus elasticity tends to increase the price. Overall, the effect of the ban on drug prices will depend on whether the effect due to the price elasticity of demand goes in the same direction or is stronger and in the opposite direction of the effect of the consumer surplus channel.

We estimate the impact of the ban on prices for the values of the bargaining parameters and weights in the identified set. Results show that prices of some approved drugs for depression increase when off label use is banned whereas prices of some others decrease. Overall, the ban on off-label drug prescriptions would increase treatment expenditures even more under the new bargaining equilibrium prices. Keeping prices unchanged, physicians substitute towards on-label drugs and because, on an average, on-label drugs are more expensive this substitution effect leads

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

to an increase in prescription expenses when off label prescriptions are banned. Using equilibirum prices negotiated under the ban, the prices of on-label drugs, on average, increase and hence; the prescription expenses increase 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 under no-ban benchmark but also if prices are negotiated under the ban, while the ban does not provide any improvement in patients' health outcomes.

In a different context, some recent literature has studied the effects of restricting supply in prescription drug plan choices in the US as done by some insurance companies. Lucarelli, Prince and Simon (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, Polyakova, and Ryan (2016) 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. Dickstein (2015) shows the effects of "demand-side" incentives (imposing costs on patients to limit moral hazard) and "supply-side" incentives (adjusting physicians' compensation to discourage spending) on health spending.

The paper proceeds as follows. Section 2 summarizes the relevant information on the institutions and regulations of the French health care system, and in section 3 describes the data and provide descriptive statistics. Section 4 presents the econometric model and empirical findings. Section 5 describes the supply side model of price setting. 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 can issue marketing authorization for drugs in France. For authorization, they require substantial evidence of efficacy and safety determined by clinical trials. 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 (called SMR). The Commission also compares the new drugs with the existing drugs and assigns a score for the improvement in medical benefit called ASMR. 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.

After the Commission determines the reimbursement level of a new drug, the Economic Committee on Health Products negotiates prices with manufacturers. Pricing decisions depend on the medical benefit improvement rating (ASMR), 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 SMR and ASMR values of a drug are based on the main indication(s) of the drug. Potential off-label uses of the drug do not play a role in determining the SMR/ASMR values. Unlike in the US, drug samples and direct-to-consumer promotion of prescription drugs are tightly restricted in France (Gallini et. al. 2014).

2.2 Regulation of off-label prescriptions

After 2009, France went through major changes in how it regulated drugs, including post-marketing regulations. In the law passed on December 29, 2011, the main regulation regarding off-label use is a decree called "Temporary Recommendations for Use" (TRU). This is a process for temporarily supervising drug prescriptions for unlicensed indications. The objective with TRUs is to open an observation window (maximum 3 years) to assess the benefits and risks of marketed drugs for unlicensed indications. The drug-regulatory agency became able to authorize a TRU only under some conditions. For instance, there should not be any approved medication available for the particular indication for which a TRU is demanded. Another condition for approval of a TRU is the existence of high quality scientific evidence of the therapeutic efficacy of the drug for the

indication for which a TRU is demanded (Emmerich et. al. 2012). Since the introduction of this regulation on December 29, 2011, only three TRUs have been granted. On December 31, 2014, a new decree amending the rules on TRUs was published. According to the new regulation, authorization of a new TRU no longer requires the absence of a therapeutic alternative for the indication in question, so the regulatory agency can authorize a TRU for an off-label indication even though there is an approved drug for that indication[5]. The sample period of data used, 2000-2008, precedes the major changes in regulation concerning off-label use. During this period, physicians did not have any restrictions in off-label prescriptions.

2.3 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. A total of 90 % of the population has supplementary health insurance for complementary coverage of benefits not fully covered by the mandatory health insurance. 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 face a uniform incentive scheme. For instance, the heterogeneous constraints on doctor's 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 one, however detailing towards doctors by pharmaceutical companies may play a role in their prescription choices.

3 Data and Descriptive Statistics

3.1 Data

The main data is proprietary by 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 to follow individual patients over time. At the physician level, the data set includes age, gender

[5]Source: Article by Cynthia Burton on 'The National Law Review ' posted on January 21, 2015.

and region of operation. At the patient level, it includes socio-demographic information (age, gender, employment status) and information on health (chronic disease condition, BMI). 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 diagnoses-drug prescription pairs on each visit, including details such as 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 (HAS)[6]. 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 as off-label. We also use the DRUGDEX system (Thomson Micromedex, Greenwood Village, Colo), a nationally recognized pharmaceutical compendium in the US, 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[7]. Combining these data allows 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 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 since 1990s to today at CIP level (code that identifies

[6] http://www.sante.gouv.fr/medicaments,1969.html (official website of Ministry of Health) and http://www.theriaque.org (approved by French regulatory agency HAS)

[7] 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.

the presentation of a drug fromat, 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 World Health Organization (https://www.whocc.no/atc_ddd_index/).

Finally, we also use drug detailing data from 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 of May 2001-December 2008.

3.2 Descriptive Statistics

Table 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.

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 1 also reports the total number of depressed patients over the sample period.

Table 1: On-Label versus Off-label Prescriptions in Depression Treatment

	Share	Average Price
On-Label	79%	0.83€
Off-Label with Efficacy	3%	0.49€
Other Off-Label	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 World Health Organization. For each active ingredient, it is the price per mg times mg per day according to DDD. The third column shows the average of one-day treatment price for the drug group in the first column of the corresponding row.

Table 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, other off-label drugs. There are three columns showing three different time periods. 'One-year Period' represents the one-year period starting six months after the first diagnosis. 'Six-month Period' represents the six-month period starting six months after the first

diagnosis. 'Anytime' represents the whole period starting six months after the first diagnosis until the end of the sample period.

The share of patients that recover from the disease is higher among those who are prescribed off-label drugs: patients who are prescribed off-label drugs are around 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 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 are having another cycle of depression (relapse cases).

Table 3 reports statistics on recovery one year after the first diagnosis. Recovery rates after one year are around 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 2.

Table 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: 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 are having another cycle of depression (relapse cases).

These descriptive statistics show that there are clearly correlations between prescriptions and health status, that could be a combination of treatment effects of drugs and selection into treatments

by physicians. We will now develop a model allowing to disentangle causality effects from correlated effects due to heterogeneity. This will allow us to identify the counterfactual effects of choices and health outcomes when off label alternatives become unavailable to the prescriber.

4 A Joint Model of Prescription Choice and Treatment Outcome

4.1 Econometric Model and Identification

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 in order 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 a drug specific detailing expenditures stock, I_j is an unobserved patient state affecting the propensity to recover and which can affect drug preferences differently with coefficient λ_d and ε_{ijdt} is a deviation from the mean utility of d at t assumed independent of all other variables. $l(d)$ is a dummy variable indicating if 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. In order 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] $x_{dt} = \sum_{\tau=-\infty}^{t} (0.75)^{t-\tau} (\text{detailing expenditures})_{d\tau}$.

The unobserved state I_j is assumed to be discrete with two types of patients where $I_j = I_j^{high}$ with probability q and $I_j = I_j^{low}$ with probability $(1-q)$. The drug choice is denoted by $y_{ijt} \in \{1, .., D\}$ and alternative 0 corresponding to the group of other off-label drugs has normalized utility $U_{ij0t} = \gamma x_{0t} + \varepsilon_{ij0t}$ (details on choice alternatives are provided in Section 4.2).

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

$$P(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j) = \frac{\exp\left(\alpha_d(z_i, z_j) - \beta p_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d I_j\right)}{\sum_{d'=0}^{D} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} I_j\right)} \tag{1}$$

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

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

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

The recovery status of the 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 for individual i treated by physician j relative to the reference drug. This specification allows 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 more severe cases and *low-type* patients as the mild cases of depression. λ_d allows the unobserved patient state to affect drug preferences differently across drugs. Therefore, the patient state, which is 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 to recover conditional on prescription choice y_{ijt} and unobserved patient-state I_j is:

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

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

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

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

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

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

where

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

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

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

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

We can then use simulated maximum likelihood in order 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 the likelihood $l(y_{ijt}=d, r_{ijt}=1)$ is

$$\begin{aligned} l(y_{ijt} = d, r_{ijt}=1) = \frac{1}{S}\sum_{s=1}^{S} \Bigg\{ & q \left(\frac{\exp\left(\alpha_d\left(z_i, z_j\right) - \beta p_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d I_j^{high}\right)}{\sum_{d'=0}^{D} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} I_j^{high}\right)} \right) \times \\ & \varphi(z_j'\theta + \sum_{d=1}^{D} \left(\delta_d + \sigma_d \delta_{ij}^{ds}\right) \mathbf{1}_{\{y_{ijt}=d\}} + I_j^{high}) \\ & + \ (1-q) \left(\frac{\exp\left(\alpha_d\left(z_i, z_j\right) - \beta p_{dt} + \gamma_{l(d)} x_{dt} + \lambda_d I_j^{low}\right)}{\sum_{d'=0}^{D} \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \gamma_{l(d')} x_{d't} + \lambda_{d'} I_j^{low}\right)} \right) \times \\ & \varphi(z_j'\theta + \sum_{d=1}^{D} \left(\delta_d + \sigma_d \delta_{ij}^{ds}\right) \mathbf{1}_{\{y_{ijt}=d\}} + I_j^{low}) \Bigg\} \end{aligned}$$

where 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

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

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:

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

We use an exclusion restriction on drug advertising which is allowed to affect treatment choice but not treatment outcome conditional on the treatment choice. As direct-to-consumer promotion of prescription drugs is forbidden in France (Gallini et. al. 2014) patients are not exposed to drug advertising.

4.2 Estimation and Empirical Results

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 as 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 with scientific evidence from the medical literature showing their efficacy in treatment of depression, and 'Other Off-Label Drugs', which includes all the other off-label active ingredients and is the reference alternative in the estimation[9]. In total there are 8

[9] Details on the drugs aggregated under alternatives 'Other On-Label Drugs', 'Off-Label Drugs with Efficacy' and 'Other Off-Label Drugs' are provided in the appendix.

exclusive alternatives in the physicians' choice set[10]. To avoid dealing with learning about patient's response in 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[11]. The treatment outcome considered in the analysis is whether the patient recovers from the disease after six-months[12]. The recovery is defined by the patient not being diagnosed with depression during the one-year period after the six-months treatment period. Given the sample period ends in December 2008, only the patients that are diagnosed with depression until the end of June 2007 are included in the analysis[13]. The patients that stop visiting their family physician during or after the six-month treatment period are not included in the analysis. This share is small as it is unlikely that an individual does not visit his primary care physician for a long period and it is the same for patients treated with on-label drugs and those treated with off-label drugs (for details see the appendix).

As a baseline for the analysis, we estimate first 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 4 reports the parameters of the logit estimation of treatment choice in this case, and Table 5 provides the parameters of binary probit estimation of treatment outcome. In the estimations, the reference category is 'Other Off-Label Drugs'; therefore, coefficients are interpreted relative to 'Other Off-Label Drugs'. The positively significant coefficient for patients' age for the on-label drug 'Citalopram' means that, older patients are more likely to get a 'Citalopram' prescription rather than an 'Other Off-Label Drug'. The drug-specific constant terms control for the time invariant variation across drugs. Physician and patient characteristics (age, gender) are also introduced.

[10] Co-prescriptions are less than 3 % of the cases and they are excluded from the analysis.

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

[12] For robustness, the same analysis was done by allowing for combinations of six-month or one-year of treatment period with six-month or one-year of 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.

[13] For visiting a psychiatrist, patients need referral from their family physician. We can observe whether a patient is referred to a specialist. The patients that 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 that 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.

Table 4 shows that demand is price sensitive with a significantly negative coefficient for price[14]. Also, detailing expenditures positively affect the demand in the on-label market, whereas its impact on the demand in the off-label market is limited similar to the findings in Shapiro (2018). This is coherent with the fact that pharmaceutical firms can advertise their drugs for the on-label use but are not allowed to advertise them for the off-label use.

Table 4: Logit Estimation of Treatment Choice

	Alternative Specific Parameters				
	Patient's		Physician's		Constant
	Age	Sex (female=1)	Age	Sex (female=1)	
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)	(0.252)
	Parameters Common Across Alternatives				
Price (β)	Advertising*On-Label			Advertising*Off-Label	
-1.042	0.227			0.004	
(0.278)	(0.014)			(0.023)	
Observations	37,510				

Notes: Standard errors in parentheses.

Advertising is the natural logarithm of stock of advertising.

Table 5 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_d's in the model, the treatment effect of drug d relative to the control group. As a baseline for the analysis, we do not allow for heterogeneity in treatment impact across patients hence $\sigma_d = 0$. δ_d's are significantly negative for all of the approved drugs, which means patients treated with approved drugs have lower recovery rates than patients treated with 'Other Off-Label Drugs'. The recovery

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

rates with alternative 'Off-Label Drugs with Efficacy' are not different than the recovery rates with 'Other Off-Label Drugs'.

Table 5: Binary Probit Estimation of Treatment Outcome

	Parameter	Std. Error	Marginal effect	Std. Error
Patient's Age	-0.013	(0.000)	-0.005	(0.000)
Patient's 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 in parentheses.

Table 6 provides the results of the joint estimation of treatment choice and treatment outcome, taking into account unobserved patient heterogeneity which can potentially affect both the treatment choice and the treatment outcome. Detailing expenditures to general practitioners at drug-month level are excluded from the health outcome equation. Advertising affects drug choices but, conditional on the drug choice, it would not impact the health outcome. If patients treated with drugs that are more advertised obtain better health outcomes, the model attributes this effect to the drug choice.

The first part of Table 6 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 outcome. The drug-specific coefficients in part 2 of Table 6 are the δ_d's 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[15].

[15] 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).

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_d's show the 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 72% of the patients are high-types and hence 28 % are low-types. The negatively significant estimates of λ_d for all the on-label alternatives show that these active ingredients are less (more) likely to be prescribed to high (low) types which are the patients that are more (less) likely to recover from the disease due to their unobserved state. The severity of the disease can be an example of this type of heterogeneity: i.e. relatively more severe cases 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 treatment impact parameter for most of the on-label alternatives is either positive or not statistically different than zero. The alternatives with a significantly positive δ_d, 'Sertraline' and 'Paroxetine', on an average, lead to higher recovery rates than the reference category, 'Other Off-Label Drugs', and the alternatives 'Citalopram' and 'Escitalopram' lead to the same recovery rates as the reference category in contrast to the results presented on Table 5 which does not take into account this unobserved heterogeneity. The alternatives 'Fluoxetine' and 'Other On-Label' lead to, on an average, slightly lower recovery rates than 'Other Off-Label Drugs' despite the selection into treatment based on unobservables. Overall, the results in Table 6 show that there is unobserved patient-level heterogeneity that impacts both the drug choice and the recovery status. Statistically significant estimates of σ_d show that the treatment effect of each drug, relative to the reference group, is heterogenous across patients, which means, for some patients, treatment with 'Citalopram', for instance, leads to lower recovery rates than treatment with 'Other Off-Label Drugs', whereas for some other patients it leads to higher recovery rates, even though the average relative impact, δ_d, is not statistically different than zero. Table 7 reports the quantiles of the marginal treatment effects of the drugs relative to the reference alternative 'Other Off-Label Drugs'. For instance, with 'Citalopram' treatment, recovery rates are, on an average, 0.5 percentage points lower than recovery rates provided by 'Other Off-Label Drugs'. However, for 46% of the patients, 'Citalopram' leads to higher recovery rates than the reference category. For all the on-label drugs, there are some patients for whom the

treatment effect relative to 'Other Off-Label Drugs' is positive, the lowest being 41% for 'Fluoxetine' treatment. The coefficients of observed patient characteristics in Table 6 show that older patients and female patients are less likely to recover from the disease.

Table 8 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 the patients are treated with 'Citalopram' the average recovery rate is 0.14 for *low-type* patients and 0.60 for the *high-types*. The average recovery rate is at the highest when all the patients are treated with 'Sertraline' or 'Paroxetine', and the lowest with 'Fluoxetine' for *high-type* patients and with 'Other Off-Label Drugs' for the *low-types*.

Table 6: Joint Estimation of Treatment Choice and Treatment Outcome

Part 1: Treatment Choice Equation

	Alternative Specific Parameters				
	Patient's		Physician's		λ_d
	Age	Sex	Age	Sex	
On-Label drugs					
Citalopram	0.015	-0.088	-0.015	-0.115	-2.256
	(0.001)	(0.044)	(0.003)	(0.052)	(0.491)
Sertraline	0.004	-0.281	0.007	-0.410	-4.472
	(0.003)	(0.092)	(0.005)	(0.120)	(0.628)
Paroxetine	0.011	-0.325	0.011	-0.451	-4.811
	(0.003)	(0.093)	(0.005)	(0.121)	(0.630)
Fluoxetine	0.012	0.033	0.001	-0.166	-1.915
	(0.001)	(0.042)	(0.003)	(0.050)	(0.535)
Escitalopram	0.011	-0.131	0.006	-0.430	-2.826
	(0.002)	(0.060)	(0.004)	(0.076)	(0.578)
Other	0.022	-0.198	0.005	-0.066	-2.415
	(0.001)	(0.036)	(0.002)	(0.043)	(0.519)
Off-Label drugs with Efficacy	-0.011	-0.258	-0.001	-0.132	0.750
	(0.002)	(0.067)	(0.004)	(0.083)	(0.500)

Parameters Common Across Alternatives			Share of
Price (β)	Advertising*On-Label	Advertising*Off-Label	*High-Type* (q)
-1.068	0.268	0.061	0.720
(0.397)	(0.018)	(0.024)	(0.072)

Part 2: Treatment Outcome Equation

	δ_d	Std. Err.	σ_d	Std. Err.
On-Label drugs				
Citalopram	-0.098	(0.074)	1.065	(0.245)
Sertraline	1.514	(0.122)	0.741	(0.341)
Paroxetine	1.701	(0.084)	1.274	(0.229)
Fluoxetine	-0.178	(0.045)	0.778	(0.211)
Escitalopram	0.248	(0.154)	1.435	(0.466)
Other	-0.127	(0.052)	1.067	(0.186)
Off-Label drugs with Efficacy	0.109	(0.148)	0.769	(0.426)

	Coef.	Std. Err.
Patient's Age	-0.019	(0.001)
Patient's Sex	-0.086	(0.021)
Constant	0.404	(0.053)
Observations	37.510	

Notes: Standard errors in parentheses. Advertising is the natural logarithm of stock of advertising. The "Sex" variable is 1 for females and 0 for males. The estimation includes a dummy which 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 Tuncel, 2019).

Table 7: Quantiles on Marginal Effect of Treatment

	Quantiles					Percentage with positive relative treatment effect
	25%	50%	75%	95%	Mean	
On-Label						
Citalopram	-9.6	-1.3	14.5	39.2	-0.5	46 %
Sertraline	20.1	30.1	46.5	75.9	34.4	98 %
Paroxetine	17.3	29.7	52.8	88.0	35.0	91 %
Fluoxetine	-9.8	-2.2	7.8	26.9	-3.1	41 %
Escitalopram	-7.2	5.1	25.2	65.1	7.3	57 %
Other On-Label	-10.0	-1.6	13.9	38.5	-1.0	45 %
Off-Label						
with Efficacy	-5.4	2.0	14.8	33.8	3.2	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, for 25 % of the patients recovery rates when treated with 'Citalopram' are 9.6 percentage points lower than recovery rates when treated with 'Other Off-Label Drugs'.

Table 8: Quantiles on Recovery Probabilities

	Quantiles				
	25%	50%	75%	95%	Mean
If all *low-type* patients are treated with:					
On-Label					
Citalopram	.01	.05	.20	.59	.14
Sertraline	.30	.50	.71	.91	.50
Paroxetine	.25	.58	.86	.99	.55
Fluoxetine	.01	.05	.13	.38	.10
Escitalopram	.01	.10	.40	.88	.24
Other On-Label	.01	.05	.19	.58	.14
Off-Label					
with Efficacy	.03	.08	.20	.49	.14
Other	.04	.07	.10	.14	.07
If all *high-type* patients are treated with:					
On-Label					
Citalopram	.36	.65	.87	.99	.60
Sertraline	.93	.98	.99	1.0	.94
Paroxetine	.91	.99	1.0	1.0	.91
Fluoxetine	.40	.62	.81	.95	.60
Escitalopram	.40	.77	.96	1.0	.66
Other On-Label	.35	.64	.87	.99	.60
Off-Label					
with Efficacy	.52	.73	.88	.98	.68
Other	.62	.70	.76	.83	.68

5 A Supply Side Model of Price Setting

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

5.1 Pricing Equilibrium with Off-Label Drugs

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 treatment. Bargaining models have been used to represent negotiations between insurers and hospitals in the US (Gowrisankaran et al. 2015) and between hospitals and medical device providers (Grennan, 2013).

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

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

where $q_{dt}^{tot}(\mathbf{p}_t^{on}, \mathbf{p}_t^{off})$ is the total sales of drug d in markets for both on-label and off-label indications, at the vector of prices $\mathbf{p}_t = (\mathbf{p}_t^{on}, \mathbf{p}_t^{off})$ where $\mathbf{p}_t^{on}$ is the vector of prices of all the drugs in the relevant on-label market and $\mathbf{p}_t^{off}$ is the vector of prices of all the drugs in the relevant off-label market[16].

As we assumed that ε_{ijdt} is i.i.d. with type I extreme value, the consumer surplus on the on-label market of drug d has the standard closed form (Small and Rosen, 1981)[17]

$$\begin{aligned} CS_t(\mathbf{p}_t^{on}) &= \frac{1}{\beta} \sum_{j \in J_{on}} \left\{ \ln \left[\sum_{d' \in D^{on}} \left(q \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \lambda_{d'} I_j^{high} \right) \right. \right. \right. \\ &\quad \left. \left. \left. + (1-q) \exp\left(\alpha_{d'}(z_i, z_j) - \beta p_{d't} + \lambda_{d'} I_j^{low} \right) \right) \right] \right\} \end{aligned}$$

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 drug d (on-label indication).

[16] The price and the reimbursement level of a drug do not change depending on the indication it is prescribed for. 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 vector of prices p_t^{on} and p_t^{off} are different vectors.

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

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 involves that all contracts remain the same if another negotiation fails (see e.g. Crawford and Yurukoglu (2012), Gowrisankaran et al. (2015), Dubois and Saethre (2019)). Assuming a Nash in Nash equilibrium (Horn and Wolinsky, 1988), active ingredient by active ingredient, this amounts to maximizing the Nash product given the other prices:

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

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), $\Delta_d CS_t(\mathbf{p}_t^{on}) \equiv CS_t(\mathbf{p}_t^{on}) - CS_{t,-d}(\mathbf{p}_t^{on})$ and $TC_{dt}(\mathbf{p}_t^{on}) = p_{dt}q_{dt}$ are respectively drug d's additional value added in consumer surplus for the on-label indication and the cost of treatment with the drug d, where q_{dt} is the aggregate demand for drug d in the on-label market. w is the weight the government puts on consumer surplus, and hence $(1-w)$ is the weight the government puts on cost of treatment with drug d.

Note that the consumer surplus absent drug d is:

$$\begin{aligned} CS_{t,-d}(\mathbf{p}_t^{on}) &= \frac{1}{\beta}\sum_{j\in J_{on}} \left\{ \ln\left[\sum_{d'\in D^{on}\setminus\{d\}} \left(q\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't} + \lambda_{d'} I_j^{high}\right)\right.\right.\right. \\ &\qquad \left.\left.\left. + (1-q)\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't} + \lambda_{d'} I_j^{low}\right)\right)\right]\right\} \end{aligned}$$

The first order conditions of this Nash equilibrium are

$$\frac{\mu}{1-\mu}\frac{\partial \ln \Pi_{dt}(\mathbf{p}_t^{on}, \mathbf{p}_t^{off})}{\partial p_{dt}} + \frac{\partial \ln[w\Delta_d CS_t(\mathbf{p}_t^{on}) - (1-w)TC_{dt}(\mathbf{p}_t^{on})]}{\partial p_{dt}} = \mathbf{0} \tag{2}$$

and thus the firm's marginal cost is

$$c_{dt} = p_{dt} + \frac{1}{\left(\frac{\partial \ln q_{dt}^{tot}(\mathbf{p}_t^{on}, \mathbf{p}_t^{off})}{\partial p_{dt}}\right) + \frac{1-\mu}{\mu}\left(\frac{\partial \ln[w\Delta_d CS_t(\mathbf{p}_t^{on}) - (1-w)TC_{dt}(\mathbf{p}_t^{on})]}{\partial p_{dt}}\right)} \tag{3}$$

where $\frac{\partial \ln q_{dt}^{tot}(\mathbf{p}_t^{on}, \mathbf{p}_t^{off})}{\partial p_{dt}}$ 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

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

and

$$\frac{\partial TC_{dt}(\mathbf{p}_t^{on})}{\partial p_{dt}} = \frac{\partial p_{dt} q_{dt}(\mathbf{p}_t^{on})}{\partial p_{dt}} = q_{dt}(\mathbf{p}_t^{on}) + p_{dt}\frac{\partial q_{dt}(\mathbf{p}_t^{on})}{\partial p_{dt}}$$

Moreover, the profit of the firm in both markets is

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

implying that

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

where

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

and

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

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 treatment of the following diseases: alcoholism, anguish, anxiety, asthenia, bipolar, dementia, headache and migraine, 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.

The first-order conditions (2) depend on demand, but also 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 of drugs. Indeed, using (3), the firm's marginal cost depends on the demand parameter vector denoted $\mathbf{\Lambda}$ and the supply side parameters μ and w as follows:

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

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

We then assume that true marginal costs c_d^0 are time invariant and that ζ_{dt} such that

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

is mean independent of true cost c_d^0. This implies the moment condition

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

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

$$0 \leq E\left(c_{dt}(\mathbf{\Lambda}, \mu, w, \mathbf{z}_t)\right) \leq \underline{p}_d \qquad \forall d = 1, .., D$$

Remark that we could slightly generalize by allowing marginal costs c_d^0 to vary overtime and then use time specific upper bounds on the cost. In practice, the minimum price $\underline{p}_d$ will be the minium price of the corresponding drug d over a period of observation that goes beyond the period of estimates of costs, until 2018.

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

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

and can be empirically estimated using

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

Remark that demand parameters are unknown and, while independently estimated, we should also account for the uncertainty in $\mathbf{\Lambda}$ estimates and obtain random sets $\hat{S}_{(\mu,w)}(\mathbf{\Lambda})$. The asymptotic theory of these identified sets can be studied using Chernozoukov, Hong, Tamer (2007) or Andrews and Jia 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 check with respect to $\widehat{\mathbf{\Lambda}}$.

In the identified set, the bargaining power of the firm, μ, takes values between 0.55 and 0.65 and the weight the government puts on consumer surplus, w, is between 0.9 and 1. Table 9 provides the average prices during the sample period, prices in 2018 and the marginal cost of on-label alternatives for all the combinations of μ and w in the identified set.

Table 9: Marginal Cost Estimates of On-label Drugs

	Average Price During Sample Period (€)	Prices in 2018 (€)	Marginal Cost Estimate (€)				
Bargaining parameter (μ)			0.55	0.60	0.60	0.65	0.65
Weight on CS (w)			1	0.95	1	0.90	0.95
Citalopram	.836	.271	.268	.252	.219	.245	.201
Sertraline	.804	.286	.199	.161	.150	.124	.112
Paroxetine	.776	.203	.091	.051	.041	.012	.001
Fluoxetine	.744	.267	.102	.097	.046	.101	.037
Escitalopram	.685	.209	.088	.103	.031	.132	.043
Other On-Label Drugs	.928	.526	.297	.335	.233	.407	.268

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

5.2 Pricing Equilibrium when Off-Label Drugs are Banned

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 for d is

$$\begin{aligned} P\left(y_{ijt}=d|z_i,z_j,p_{dt},x_{dt},I_j\right) &= q\frac{\exp\left(\alpha_d\left(z_i,z_j\right)-\beta p_{dt}+\gamma_{l(d)}x_{dt}+\lambda_d I_j^{high}\right)}{\sum\limits_{d'\in D_{on}^{on}}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\gamma_{l(d')}x_{d't}+\lambda_{d'}I_j^{high}\right)} \\ &\quad +(1-q)\frac{\exp\left(\alpha_d\left(z_i,z_j\right)-\beta p_{dt}+\gamma_{l(d)}x_{dt}+\lambda_d I_j^{low}\right)}{\sum\limits_{d'\in D_{on}^{on}}\exp\left(\alpha_{d'}(z_i,z_j)-\beta p_{d't}+\gamma_{l(d')}x_{d't}+\lambda_{d'}I_j^{low}\right)} \end{aligned}$$

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.

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_{dt}^{ban}, would be the solution of:

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

where $\Pi_{dt}^{ban}(\mathbf{p}_t^{ban})$ is the profit of the firm for active ingredient d when off-label drugs competing with d are banned, and $\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban}) \equiv CS_t^{ban}(\mathbf{p}_t^{ban}) - CS_{t,-d}^{ban}(\mathbf{p}_t^{ban})$ is the consumer surplus provided by drug d for the on-label indication where:

$$\begin{aligned} CS_t^{ban}(\mathbf{p}_t^{ban}) &= \frac{1}{\beta}\sum_{j\in J_{on}} \left\{ \ln\left[\sum_{d'\in D_{on}^{on}} \left(q\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't}^{ban} + \lambda_{d'} I_j^{high}\right)\right.\right.\right. \\ &\quad \left.\left.\left. + (1-q)\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't}^{ban} + \lambda_{d'} I_j^{low}\right)\right)\right]\right\} \end{aligned}$$

and

$$\begin{aligned} CS_{t,-d}^{ban}(\mathbf{p}_t^{ban}) &= \frac{1}{\beta}\sum_{j\in J_{on}} \left\{ \ln\left[\sum_{d'\in D_{on}^{on}\setminus\{d\}} \left(q\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't}^{ban} + \lambda_{d'} I_j^{high}\right)\right.\right.\right. \\ &\quad \left.\left.\left. + (1-q)\exp\left(\alpha_{d'}(z_i,z_j) - \beta p_{d't}^{ban} + \lambda_{d'} I_j^{low}\right)\right)\right]\right\} \end{aligned}$$

The firm's profits are now:

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

where $q_{dt}^{ban}(p_t^{ban})$ is the aggregate demand of drug d in case of an off-label ban and costs are estimated according to (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 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 the sales in the off-label markets.

The cost of treatment with the drug d is now:

$$TC_{dt}(\mathbf{p}_t^{ban}) = p_{dt}^{ban} q_{dt}^{ban}(\mathbf{p}_t^{ban})$$

The first-order condition can be written as

$$\frac{\mu}{1-\mu}\frac{\partial \ln \Pi_{dt}^{ban}(\mathbf{p}_t^{ban})}{\partial p_{dt}}+\frac{\partial \ln[w\Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})-(1-w)TC_{dt}^{ban}(\mathbf{p}_t^{ban})]}{\partial p_{dt}}=\mathbf{0} \tag{4}$$

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

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

where $\frac{\partial \ln q_{dt}^{on}(\mathbf{p}_t^{ban})}{\partial p_{dt}}$ is the price semi elasticity of demand at the new price equilibrium in the on-label market.

This allows us to compute 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 costs and recovery probabilities.

Note that price cost margin is a function of price elasticity of demand and also price elasticity of drug d's additional value added in consumer surplus at the new price equilibrium. The ban can impact drug prices through both of these channels; through its impact on demand and consumer surplus elasticities. As mentioned before, whether the aggregate demand for drug d will increase or decrease with the ban relative to no-ban is ambiguous because $q_{dt}^{tot}(p_t^{on},p_t^{off})=q_{dt}^{on}(p_t^{on})+q_{dt}^{off}(p_t^{off})$ is the sum of sales of drug d on the on-label market and all the off-label markets for which drug d is used, whereas $q_{dt}^{on}(p_t^{ban})$ 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 $\frac{\partial \ln q_{dt}^{on}(\mathbf{p}_t^{ban})}{\partial p_{dt}}$ is larger or smaller than $\frac{\partial \ln q_{dt}^{tot}(\mathbf{p}_t^{on},\mathbf{p}_t^{off})}{\partial p_{dt}}$ is ambiguous. Semi elasticities can be ranked in any direction depending on the price at which they are evaluated but also 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 also its elasticity in the off-label indication markets.

Similarly, the ban can impact the drug prices through the channel of the price elasticity of consumer surplus value added of drugs. 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 $\Delta_d CS_t^{ban}(p_t^{ban})>\Delta_d CS_t(p_t^{on})>0$. The order of derivatives of consumer surplus value added of drug d with respect to price is such that $\frac{\partial \Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}{\partial p_{dt}}<\frac{\Delta_d CS_t(\mathbf{p}_t^{on})}{\partial p_{dt}}<0$. It

can be shown that the semi elasticity of $\Delta_d CS$ is larger with less drugs than with more drugs in the choice set, $\frac{\partial \ln \Delta_d CS_t(\mathbf{p}_t^{on})}{\partial p_{dt}} < \frac{\partial \ln \Delta_d CS_t^{ban}(\mathbf{p}_t^{ban})}{\partial p_{dt}} < 0$. Therefore; the ban has an increasing impact on the price through the channel of consumer surplus elasticity. In short, the impact of the ban on drug prices through these two channels can go in the same direction; meaning through both channels the ban can increase drug prices. Similarly, the impact of the ban through the two channels can go in opposite directions; the ban can have a decreasing impact on prices through demand elasticity whereas it has an increasing impact on prices through consumer surplus-elasticity. In this case it will be the dominating channel which will determine whether the price of a given drug increases or decreases with the ban.

6 Counterfactual Simulations when Off-Label Prescriptions are Banned

Using the structural choice model, we perform counterfactual simulations of prescription choices when off-label drugs are not in the choice set of physicians. In France, the current regulation under "Temporary Recommendations for Use (TRU)" aims at strictly regulating off-label prescriptions. In the US, 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 change in reaction to a ban on off-label prescriptions and simulate the associated counterfactual expected demand and cost using the counterfactual prices.

6.1 Counterfactual Ban on Off-Label Prescriptions Keeping Prices Constant

Assuming prices would not change in case of a ban on off-label, we can simulate not only the counterfactual prescription choices but also their expected costs and the expected recovery rate of patients, simply by removing alternatives that are off label in the prescription possibilities of physicians. Two cases are worth consideration 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 evidence of efficacy in depression treatment, 'Other Off-Label Drugs', are banned.

Then, 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

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

where, as defined above, D_{on}^{on} is the set of on-label drugs for the on label indication considered. Note that the choice probability of the reference alternative after the ban is 0 because our reference alternative is off-label drugs.

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

$$\sum_{j \in J_{on}} \sum_{d \in D^{on}} P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j\right) p_{dt} \tag{5}$$

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

$$\sum_{j \in J_{on}} \sum_{d \in D_{on}^{on}} P_c\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j\right) p_{dt} \tag{6}$$

so the change in expected cost is the difference of (5) and (6).

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

$$E\left[r_{ijt}\right] = \sum_{d \in D^{on}} P(r_{ijt} = 1 | z_i, I_j, y_{ijt} = d) P(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j) \tag{7}$$

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

$$E\left[r_{ijt}^c\right] = \sum_{d \in D_{on}^{on}} P(r_{ijt} = 1 | z_i, I_j, y_{ijt} = d) P_c(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j) \tag{8}$$

Table 10 reports the statistics on the expected cost of one day treatment for the cases when off-label prescriptions are allowed, and also 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 by 7.4 % when we ban all off-label drugs, and by 5.7 % when we ban only 'Other Off-Label Drugs' and still allow for prescription of off-label drugs with efficacy.

Table 10: Expected Prescription Expense of Drug Treatment (in Euros)

	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	
		Expected Expense	Expected Expense	Change	Expected Expense	Change
All Patients	Mean	.679 [.581,.710]	.718 [.713,.721]	.039 (+5.7%)	.728 [.723,.732]	.050 (+7.4%)
	Min.	.598 [.479,.649]	.666 [.660,.669]	.068 (+11.4%)	.696 [.690,.701]	.098 (+16.4%)
	Max.	.745 [.686,.761]	.763 [.761,.765]	.018 (+2.4%)	.769 [.765,.770]	.024 (+3.2%)
Female Patients	Mean	.676 [.576,.708]	.716 [.712,.719]	.040 (+5.9%)	.726 [.721,.730]	.050 (+7.4%)
Male Patients	Mean	.685 [.595,.714]	.721 [.716,.725]	.036 (+5.3%)	.733 [.727,.737]	.048 (+7.0%)
Old Patients	Mean	.698 [.608,.724]	.730 [.725,.732]	.031 (+4.4%)	.736 [.732,.739]	.038 (+5.4%)
Young Patients	Mean	.668 [.566,.702]	.711 [.706,.714]	.043 (+6.4%)	.724 [.718,.728]	.056 (+8.4%)

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. 'Price per day' is the price per mg times mg per day according to DDD for each active ingredient. Old patients are patients above 60. Confidence intervals at 90 % confidence level constructed by 300 bootstrap draws drawn from the estimated distribution of parameters are in square brackets.

Table 11 shows expected recovery rate six-months 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, unconditional recovery probability is calculated according to (7) in the benchmark and according to (8) 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 11 provides the statistics on the unconditional recovery probability; in the first column when all the drugs are in the choice set of the physicians; the second column shows the case when all the approved drugs and off-label drugs with efficacy are in the choice set of the physicians; and 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 slightly more than 1 percent (less than 1 percentage point).

Table 11: 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	
		Expected Recovery	Expected Recovery	Change	Expected Recovery	Change
All Patients	Mean	.514 [.433,.606]	.510 [.416,.606]	-.004 (-0.8%)	.508 [.415,.605]	-.006 (-1.2%)
	Min.	.274 [.186,.368]	.275 [.189,.369]	.001 (+0.4%)	.275 [.190,.368]	.001 (+0.4%)
	Max.	.786 [.751,.829]	.777 [.716,.827]	-.009 (-1.1%)	.768 [.700,.830]	-.018 (-2.3%)
Female Patients	Mean	.505 [.420,.601]	.501 [.404,.600]	-.004 (-0.8%)	.499 [.404,.600]	-.006 (-1.2%)
Male Patients	Mean	.536 [.462,.620]	.532 [.445,.619]	-.004 (-0.7%)	.530 [.444,.619]	-.006 (-1.1%)
Old Patients	Mean	.428 [.336,.529]	.427 [.330,.529]	-.001 (-0.2%)	.426 [.331,.528]	-.002 (-0.5%)
Young Patients	Mean	.563 [.488,.650]	.558 [.466,.650]	-.005 (-0.9%)	.555 [.464,.649]	-.008 (-1.4%)

Notes: Expected recovery in the counterfactual scenarios is different than the benchmark only because the counterfactual choice probabilities are different than the benchmark choice probabilities. The estimated recovery rates given the drug treatment are the same in the benchmark and in the counterfactuals. Old patients are patients above 60. Confidence intervals at 90 % significance level constructed by 200 bootstrap draws drawn from the estimated distribution of parameters are in parenthesis.

6.2 Counterfactual Ban on Off Label Prescriptions with Price Renegotiations

Now, we will investigate what prices of on-label drugs would be if the negotiations were taking place under a strict ban on off-label prescriptions. When off-label drugs are banned, the market share of on-label drugs in depression treatment will increase. However, the on-label drugs in depression treatment are used as off-label in treatment of other diseases. Therefore, with an off-label ban, their market shares in treatment of other diseases will decrease down to zero even though their market shares in depression treatment will increase. The impact of an off-label ban on the aggregate sales of the drugs approved for depression is thus ambiguous. We thus also estimate the sales of these drugs in the markets of off-label indications they are used for. The list of off-label indications antidepressants are used for is provided above.

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 now equal to

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

where p_{dt}^{ban} are the new price negotiated with the regulator after the ban on off-label use.

After removing off-label drugs from the choice set and allowing for negotiations on drug prices under a strict ban on off-label prescriptions, the treatment cost ex post on this on-label indication market is

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

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 12 presents the counterfactual prices estimated using the system of first-order conditions in (4) for different values of μ and w in the identified set. It also reports the 'Simulated Prices' which are the prices that would prevail had bargaining on prices of drugs in the benchmark scenario of no-ban happened with the given bargaining and weight parameters and the corresponding estimated marginal cost.

When off-label drugs are banned, we observe an increase in prices of all on-label drugs in three of the five combinations of μ and w. In two of the combinations there is a decrease in prices of two drugs: when $\mu = 0.55$ and $w = 1$ the prices of 'Sertraline' and 'Paroxetine' decrease, and when $\mu = 0.60$ and $w = 0.95$ the prices of 'Citalopram' and 'Escitalopram' decrease. In all the combinations of μ and w, on an average, the prices of approved drugs increase. The increase in price depends on the values of μ and w and ranges from 0.6 % to 32.5 %.

Table 12: Counterfactual Results in Case of a Ban on Off-Label Prescriptions

	When Bargaining parameter, μ=0.55 Weight on CS, w=1			When Bargaining parameter, μ=0.60 Weight on CS, w=0.95		
	Simulated Price	Counterfactual Price	Change in Price	Simulated Price	Counterfactual Price	Change in Price
Citalopram	.802	.807 [.587,1.381]	.005 (+0.6%)	.780	.732 [.585,1.326]	-.048 (-6.2%)
Sertraline	.761	.756 [.523,1.392]	-.005 (-0.7%)	.744	.767 [.501,1.351]	.023 (+3.1%)
Paroxetine	.760	.743 [.488,1.320]	-.017 (-2.2%)	.668	.718 [.455,1.368]	.050 (+7.5%)
Fluoxetine	.675	.740 [.494,1.227]	.065 (+9.6%)	.567	.611 [.504,1.346]	.044 (+7.8%)
Escitalopram	.663	.703 [.448,1.176]	.040 (+6.0%)	.674	.643 [.473,1.265]	-.031 (-4.6%)
Other On-Label Drugs	.904	1.015 [.716,1.453]	.111 (+12.3%)	.853	.941 [.778,1.871]	.088 (+10.3%)

Notes: Price is the price for one-day treatment calculated using 'Defined Daily Dose (DDD)' assigned by World Health Organization. For each active ingredient, it is the price per mg times mg per day according to DDD. Confidence intervals at 90 % confidence level constructed by 300 bootstrap draws drawn from estimated distribution of parameters are in square brackets.

Table 12(continued): Counterfactual Results in Case of a Ban on Off-Label Prescriptions

	When Bargaining parameter, μ=0.6 Weight on CS, w=1			When Bargaining parameter, μ=0.65 Weight on CS, w=0.90		
	Simulated Price	Counterfactual Price	Change in Price	Simulated Price	Counterfactual Price	Change in Price
Citalopram	.775	.807 [.560,1.400]	.032 (+4.1%)	.836	.873 [.600,1.635]	.037 (+4.4%)
Sertraline	.707	.755 [.496,1.498]	.048 (+6.8%)	.761	.767 [.495,1.383]	.006 (+0.8%)
Paroxetine	.656	.745 [.470,1.438]	.089 (+13.6%)	.735	.755 [.464,1.621]	.02 (+2.7%)
Fluoxetine	.591	.742 [.474,1.501]	.151 (+25.5%)	.695	.837 [.556,1.477]	.142 (+20.4%)
Escitalopram	.611	.708 [.435,1.488]	.097 (+15.9%)	.667	.809 [.508,1.324]	.142 (+21.3%)
Other On-Label Drugs	.861	1.027 [.715,1.934]	.166 (+19.3%)	.887	1.175 [.846,1.794]	.288 (+32.5%)

Table 12(continued): Counterfactual Results in Case of a Ban on Off-Label Prescriptions

	When Bargaining parameter, μ=0.65 Weight on CS, w=0.95		
	Simulated Price	Counterfactual Price	Change in Price
Citalopram	.770	.833 [.579,1.577]	.063 (+8.2%)
Sertraline	.640	.760 [.480,1.407]	.120 (+18.8%)
Paroxetine	.698	.750 [.451,1.269]	.052 (+7.4%)
Fluoxetine	.691	.783 [.501,1.468]	.092 (+13.3%)
Escitalopram	.674	.755 [.450,1.212]	.081 (+12.0%)
Other On-Label Drugs	.912	1.094 [.763,1.690]	.182 (+20.0%)

Table 13 provides a comparison of the expected prescription costs of drugs 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 13 % in the counterfactual scenario when drug prices are allowed to change which is higher than the increase in the case when drug prices are kept constant, 7.4 %. When off-label drugs are banned physicians substitute towards on-label drugs and because, on an 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 increase and hence; the prescription expenses increase due to both the price and substitution effects. So, in the second column of Table 14 the increase in prescription expenses is due to substitution effect whereas the increase in the third column is due to both the substitution and price effects. Similarly, Table 13 reports a comparison of the probability of recovery six-months after the first diagnosis in the same three cases. The decrease in average recovery rate in the counterfactual scenario with renegotiated prices, 1.9 %, is slightly larger than the decrease when drug prices are kept constant.

Table 13: Expected Prescription Expense of Drug Treatment (in Euros) ($\mu = 0.60, w = 0.95$)

	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 and prices are negotiated	
		Expected Expense	Expected Expense	Change	Expected Expense	Change
All Patients	Mean	.679 [.581,.710]	.728 [.723,.732]	.050 (+7.4%)	.766 [.601,1.422]	.088 (+13.0%)
	Min.	.598 [.479,.649]	.696 [.690,.701]	.098 (+16.4%)	.725 [.561,1.354]	.127 (+21.2%)
	Max.	.745 [.686,.761]	.769 [.765,.770]	.024 (+3.2%)	.818 [.652,1.539]	.073 (+9.8%)
Female Patients	Mean	.676 [.576,.708]	.726 [.721,.730]	.050 (+7.4%)	.764 [.600,1.420]	.088 (+13.0%)
Male Patients	Mean	.685 [.595,.714]	.733 [.727,.737]	.048 (+7.0%)	.772 [.602,1.426]	.087 (+12.7%)
Old Patients	Mean	.698 [.608,.724]	.736 [.732,.739]	.038 (+5.4%)	.777 [.612,1.445]	.079 (+11.3%)
Young Patients	Mean	.668 [.566,.702]	.724 [.718,.728]	.056 (+8.4%)	.760 [.594,1.408]	.093 (+13.9%)

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. 'Price per day' is the price per mg times mg per day according to DDD for each active ingredient. Old patients are patients above 60. Confidence intervals at 90 % confidence level constructed by 300 bootstrap draws drawn from the estimated distribution of parameters are in square brackets.

Table 14: 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)	
		Expected Recovery	Expected Recovery	Change	Expected Recovery	Change
All Patients	Mean	.514 [.433,.606]	.508 [.415,.605]	-.006 (-1.2%)	.504 [.418,.600]	-.010 (-1.9%)
	Min.	.274 [.186,.368]	.275 [.190,.368]	.001 (+0.4%)	.295 [.215,.395]	.021 (+7.7%)
	Max.	.786 [.751,.829]	.768 [.700,.830]	-.018 (-2.3%)	.733 [.669,.798]	-.054 (-6.9%)
Female Patients	Mean	.505 [.420,.601]	.499 [.404,.600]	-.006 (-1.2%)	.496 [.408,.595]	-.009 (-1.8%)
Male Patients	Mean	.536 [.462,.620]	.530 [.444,.619]	-.006 (-1.1%)	.523 [.443,.611]	-.013 (-2.4%)
Old Patients	Mean	.428 [.336,.529]	.426 [.331,.528]	-.002 (-0.5%)	.432 [.344,.535]	.004 (+0.9%)
Young Patients	Mean	.563 [.488,.650]	.555 [.464,.649]	-.008 (-1.4%)	.544 [.461,.637]	-.019 (-3.4%)

Notes: Expected recovery in the counterfactual scenarios is different than the benchmark only because the counterfactual choice probabilities are different than the benchmark choice probabilities. The estimated recovery rates given the drug treatment are the same in the benchmark and in the counterfactuals. Old patients are patients above 60. Confidence intervals at 90 % confidence level constructed by 300 bootstrap draws drawn from the estimated distribution of parameters are in square brackets.

7 Conclusion

Using a unique dataset which provides longitudinal information over nine years on a sample of physicians and all their patients, we estimate a model of prescription behavior with potential unobserved patient-level heterogeneity which is allowed to be correlated with treatment choice and also treatment outcome, and estimate the demand for on-label and off-label drugs in 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 an average, the probability of recovery with off-label drugs is as high as the probability of recovery with on-label alternatives.

We perform counterfactual analysis by simulating the demand and the associated treatment outcome in the case of a ban on off-label use. We consider two cases of counterfactual scenarios with respect to price. In the first one, we assume that drug prices are fixed, hence they are not affected by the off-label ban. The results suggest that banning off-label prescriptions would lead to a substantial increase in the cost of prescription drugs, whereas it does not lead to any improvement in terms of health outcomes. In the second counterfactual scenario, we allow drug prices to change by allowing the off-label ban to impact the bargaining outcome. We then simulate the associated counterfactual expected demand, treatment cost, and treatment outcome using the counterfactual prices.

Finally, a ban on off-label use would probably also impact the innovation behavior of pharmaceutical companies and/or their decisions regarding the indications for which they seek approvals. We leave such an interesting question for future research.

References

Andrews, D. and P. Jia Barwick. 2012. "Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure." *Econometrica*, 80(6): 2805–2826.

Bradford, W.D., J. L. Turner, and J. W. Williams. 2015. "Off-Label Use of Pharmaceuticals: Trends and Drivers." Working Paper.

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.M. 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(6): 502–505.

Chernozhukov, V., H. Hong, and E. Tamer. 2007. "Estimation and Confidence Regions for Parameter Sets in Econometric Models." *Econometrica*, 75(5): 1243–1284.

Cohen, J., L. Faden, S. Predaris and B. Young. 2007. "Patient access to pharmaceuticals: an international comparison." *European Journal of Health Economics*, 8: 253–266.

Crawford, G.S., and M. Shum. 2005. "Uncertainty and Learning in Pharmaceutical Demand." *Econometrica*, 73(4): 1137–1173.

Crawford, G. S., and A. Yurukoglu. 2012. "The Welfare Effects of Bundling in Multichannel Television Markets." *American Economic Review,* 102(2): 643–685.

Deb, P., and P.K. 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. forthcoming. "Subsidy Design in Privately-Provided Social Insurance: Lessons from Medicare Part D." *Journal of Political Economy.*

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

Dickstein M. 2016. "Physician vs. Patient Incentives in Prescription Drug Choice." Working Paper.

Drug database of French Social Security Health Care System ("Base de Médicaments et Informations Tarifaires" on http://www.codage.ext.cnamts.fr)

Dubois, P., and M. Saethre. 2019. "On the Effect of Parallel Trade on Manufacturers' and Retailers' Profits in the Pharmaceutical Sector." CEPR Working Paper DP12649.

Dubois, P., and T. Tunçel. 2019. "Identifying the Effects of Scientific Information and Recommendations on Physicians Prescribing Behavior." Working Paper.

Emmerich, J., N. Dumarcet,and A. Lorence. 2012. "France's New Framework for Regulating Off-Label Drug Use." *The New England Journal of Medicine*, 367(14): 1279–1281.

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. 2014. "Diffusion of Antipsychotics in the U.S. and French Markets, 1998–2008." *Psychiatric Services*, 64(7): 680–687.

Gowrisankaran G., A. Nevo, and R. Town. 2015. "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry." *American Economic Review*, 105(1): 172–203.

Grennan, M. 2013. "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices." *American Economic Review*, 103(1): 145–177.

Healthcare Series MICROMEDEX. Greenwood Village, Colo: MICROMEDEX.

Horn, H., and A. Wolinsky. 1988. "Bilateral Monopolies and Incentives for Merger." *The RAND Journal of Economics*, 19(3): 408–419.

IMS Health, Press releases 2001-2008, available at www.imshealth.com.

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(4): 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(3): 116–120.

OECD. 2016. "Pharmaceutical spending (indicator) and Health spending (indicator)." Available at data.oecd.org.

Radley D.C., S.N. Finkelstein, and R.S. Stafford. 2006. "Off-label Prescribing Among Office-Based Physicians." *Archives of Internal Medicine*, 166(9): 1021–1026.

Rodwin, V.G. 2003. "The Health Care System Under French National Health Insurance: Lessons for Health Reform in the United States." *American Journal of Public Health*, 93(1): 31–37.

Shapiro, B. 2018. "Informational Shocks, Off-Label Prescribing and the Effects of Physician Detailing." *Management Science*, 64(12): 5925–5945.

Small, K. and H. Rosen. 1981. "Applied welfare economics of discrete choice models." *Econometrica*, 49: 105–130.

Stafford, R.S. 2008. "Regulating Off-Label Drug Use-Rethinking the Role of the FDA." *New England Journal of Medicine*, 358(14): 1427–1429.

Weisman, H.F. and B. Healy. 1985. "Myocardial Infarct Expansion, Infarct Extension, and Reinfarction: Pathophysiologic Concepts." *Progress in Cardiovascular Disease, XXX(2): 73–110.*

8 Appendix

8.1 Details on Aggregated Choice Alternatives

Table A1 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 % of them 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 84 % are nervous system drugs.

Table A1: Drug Classification of Off-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 code in parenthesis

8.2 Clarification on Defining Off-Label Prescriptions

As mentioned before, in this study off-label use refers to the use of a drug for an indication the drug has never received approval. In the data, I observe which diagnoses are made at each office visit for a given patient. I also observe which drug is prescribed for each diagnosis. Using the three cases below as three examples of possible diagnosis-prescription pairs at a given office visit,

I explain how off-label prescriptions are determined in this study and how I had been conservative in determining them.

In case 1, the patient is diagnosed with depression and alcoholism. A drug that is approved for depression is prescribed in the depression diagnosis (on-label drug for depression) and another drug is prescribed in the alcoholism diagnosis. In the analysis, the patient of case 1 is considered as 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 alcoholism diagnosis). A drug that is not approved for depression, but approved for alcoholism, is prescribed in the depression diagnosis (off-label drug for depression). In the analysis, the patient of case 2 is considered as 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 in the depression diagnosis and another drug is prescribed in 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 thinks 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 presccriptions I exclude the cases like case three. Determining these cases requires that for every drug not approved but prescribed for depression I check for which indications the drug is approved and also 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

8.3 Dropouts

Table A2 shows the percentage of patients that stop visiting their physicians among all the patients and also among patients prescribed approved vs. off-label drugs. I call them 'dropouts', however one has to keep in mind that the objective of the data is to follow physicians over time, not the patients. Therefore, when a patient changes her generalist (for instance because of moving to another place) she is not in the dataset anymore. It would be worrying if dropout rates were different among patients receiving approved drugs than among those receiving off-label drugs. Share of patients that stop visiting their physician among patients that are prescribed approved drugs is not much different than the share among patients that are prescribed off-label drugs.

Table A2: 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 %
Off-Label Drugs without Efficacy	25.98 %

8.4 Additional Descriptive Statistics and Robustness Checks on Health Outcome

Table A3a 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. 96 % (87 %) of the patients that receive an on-label (off-label) drug on the first-visit are also prescribed an on-label (off-label) drug majority of the time during the six-month period.

Table A3a: 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 on the first row first column of the table is the share of patients who mostly get on-label drugs during the treatment period among the patients who get an on-label drug on the day of the first-diagnosis.

Table A3b reports the same statistics as in Table A3a for one-year treatment period.

Table A3b: 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: For explanation on the table see the note in Table A3a.

Table A4a provides the relationship between the prescription on the first visit the patient is diagnosed by 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 that receive a Citalopram on the first visit also receive a Citalopram majority of the times they are prescribed a drug in the six-month period, whereas 1.2 % of them receive a Sertraline majority of the times they are prescribed a drug.

Table A4a: 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 on the first row first column of the table is the share of patients who mostly get Citalopram during the treatment period among the patients who get a Citalopram on the day of the first-diagnosis. Similarly, the element on the first row second column is the share of patients who mostly get Sertraline during the treatment period among the patients who get a Citalopram on the day of the first-diagnosis.

Table A4b reports the same statistics as Table A4a for one-year treatment period.

Table A4b: 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: For explanation on the table see the note in Table A4b.

Table A5 reports the binary probit estimation of 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 day diagnosis.

Table A5: Binary Probit Estimation of Treatment Outcome

	Parameter	Std. Error	Marginal effect	Std. Error
Patient's Age	-0.01***	(0.00)	-0.00***	(0.00)
Patient's 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 in parentheses.

*, **, *** mean significance at 10 %, 5 % and 1 % levels.

One may be concerned that the lower recovery rates with on-label drugs are due to aggregation of off-label drugs under one choice alternative. To check if this is the case, I disaggregate the

choice alternative 'Other Off-Label Drugs' and estimate the binary probit treatment outcome model with more alternatives of off-label drugs. The off-label active ingredients I 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', 'Zopiclone'. Table A6 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', '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', 'Valeriane' than with 'Other Off-Label Drugs', and recovery rates with 'Zolpidem', 'Prazepam', '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 results in Table A6 should be considered as robustness check for results in Table 5.

Table A6: Binary Probit Estimation of Treatment Outcome

	Parameter	Std. Error	Marginal effect	Std. Error
Patient's Age	-0.01***	(0.00)	0.00***	(0.00)
Patient's 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)
Off-Label drugs without Efficacy				
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 in parentheses. *, **, *** mean significance at 10 %, 5 % and 1 % levels

9 Appendix for Online Publication

Table A1a presents the percentage of patients that are not diagnosed by depression anymore 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 A1a: 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 are having another cycle of depression (relapse cases).

Table A1b presents the percentage of patients that are not diagnosed by depression anymore one-year after they are diagnosed with depression for the first time.

Table A1b: 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: See notes in Table A1a.

Table A2 shows the statistics on percentage of patients that switch to another drug on the second visit, for each choice alternative. For instance, 20.4 % of the patients that get a Citalopram prescription on the first-visit switch to another drug on the second visit.

Table A2: Switching at the Second Visit

	Percentage of Switchers
First Visit	
Citalopram	20.4 %
Sertraline	22.1 %
Paroxetine	19.1 %
Fluoxetine	20.1 %
Escitalopram	17.5 %
Other On-Label	17.2 %
Off-Label w. Efficacy	37.1 %
Other Off-Label	22.3 %

9.1 Elasticities

9.1.1 Price Elasticity of Demand

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

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

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

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

Table A3a reports the own and cros price elasticities of demand for all the choice alternatives.

Table A3a: Own and Cross-Price Elasticities

	Citalopram	Sertraline	Paroxetine	Fluoxetine	Escitalopram	Other On-Lab.	Off-Lab. w. Eff.	Other Off-Lab.
Citalopram	-.75	.01	.01	.02	.02	.02	.03	.03
Sertraline	.01	-.66	.09	.01	.03	.02	.01	.01
Paroxetine	.04	.27	-.39	.02	.09	.05	.01	.01
Fluoxetine	.12	.03	.03	-.47	.10	.12	.14	.14
Escitalopram	.13	.13	.13	.13	-.57	.13	.13	.13
Other On-Lab.	.35	.19	.18	.36	.31	-.57	.35	.36
Off-Lab. w. Eff.	.02	.00	.00	.02	.02	.02	-.46	.03
Other Off-Lab.	.01	.00	.00	.01	.01	.01	.02	-.15

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

9.1.2 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:

$$\frac{\partial \ln s_{dt}}{\partial x_{kt}} = \begin{cases} \gamma_{l(d)} \frac{1}{s_{dt}} \frac{1}{J} \sum_{j \in J_{on}} \left\{ qP\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j^{high}\right)\left[1 - P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j^{high}\right)\right] \right. \\ \left. +(1-q)P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j^{low}\right)\left[1 - P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j^{low}\right)\right] \right\} & \text{if } d = k \\ -\gamma_{l(d)} \frac{1}{s_{dt}} \frac{1}{J} \sum_{j \in J_{on}} \left\{ qP\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j^{high}\right) P\left(y_{ijt} = k | z_i, z_j, p_{dt}, x_{dt}, I_j^{high}\right) \right. \\ \left. +(1-q)P\left(y_{ijt} = d | z_i, z_j, p_{dt}, x_{dt}, I_j^{low}\right) P\left(y_{ijt} = k | z_i, z_j, p_{dt}, x_{dt}, I_j^{low}\right) \right\} & \text{otherwise} \end{cases} \quad (9)$$

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 A3b reports the own and cros advertisement elasticities of demand for all the choice alternatives.

Table A3b: Own and Cross-Advertising Elasticities

	Citalopram	Sertraline	Paroxetine	Fluoxetine	Escitalopram	Other On-Lab.	Off-Lab. w. Eff.	Other Off-Lab.
Citalopram	.26	-.00	-.00	-.01	-.01	-.01	-.01	-.01
Sertraline	-.01	.23	-.03	-.00	-.01	-.01	-.00	-.00
Paroxetine	-.02	-.11	.15	-.01	-.04	-.02	-.00	-.00
Fluoxetine	-.05	-.01	-.01	.21	-.05	-.05	-.06	-.06
Escitalopram	-.05	-.05	-.05	-.05	.22	-.05	-.05	-.05
Other On-Lab.	-.10	-.05	-.05	-.10	-.09	.17	-.10	-.11
Off-Lab. w. Eff.	-.00	-.00	-.00	-.00	-.00	-.00	.06	-.00
Other Off-Lab.	-.01	-.00	-.00	-.01	-.00	-.00	-.01	.05

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