DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

WELFARE EFFECTS OF PHARMACEUTICAL INFORMATIVE ADVERTISING

Paris Cleanthous

Discussion Paper 02-2011

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: ++357-22893700, Fax: ++357-22895028
Web site: http://www.econ.ucy.ac.cy

WELFARE EFFECTS OF PHARMACEUTICAL INFORMATIVE ADVERTISING[1]

PARIS CLEANTHOUS[2]

December 2010

Abstract

Pharmaceutical markets are characterized by a high degree of innovation, complexity and uncertainty, especially markets of idiosyncratic symptomatolgy and response to treatment such as the antidepressant market. It may, therefore, be unreasonable to assume that consumers are aware of all antidepressants for sale at the time of purchase, as is the case in traditional models of consumer choice. Such an assumption will bias demand curves towards being more elastic and the evaluation of consumer welfare downwards. This paper, therefore, aims at analyzing and evaluating the effects of promotions by pharmaceutical firms on patient welfare taking into account the interaction of multiple agents (patients, physicians, insurance companies and pharmaceutical companies) in the decision process.

I present an empirical discrete-choice model of limited information, where advertising influences the set of drugs from which a purchase choice is made. The estimation technique incorporates both macro- and micro-level data. Estimation results indicate that pharmaceutical firms use advertising media to target high-income households and households with more comprehensive prescription drug insurance schemes through their physicians or directly. Model comparison shows that limited information leads to less elastic demand curves and larger estimates of patient welfare due to pharmaceutical innovation that exacerbate the moral hazard issue that coexists with insurance coverage.

KEYWORDS: Advertising, Health, Information, Moral Hazard, Pharmaceuticals, Welfare
JEL classification: M37, I11, L15, D82, L65, D6

[1]I have greatly benefited from discussions with Steve Berry, Bart Bronnenberg, Sofronis Clerides, Tulin Erdem, Michelle Sovinsky Goeree, Wonjoon Kim, Pinelopi Goldberg, Valanta Milliou, Fiona Scott Morton, Emmanuel Petrakis, Nikos Vettas, and seminar participants at Athens University of Economics & Business, Columbia University, KAIST, New York University, Northwestern University, UCLA, USC, University of Chicago, University of Cyprus, Yale University. I also thank IMS Health for graciously providing me with the dataset. I have benefitted from excellent research assistance by Tryfonas Chrisodoulou and Kyriakos Petrou. Financial support is gratefully acknowledged from a New York University Stern School of Business Data Grant, and a University of Cyprus Young Faculty Research Grant.

[2]Department of Economics, University of Cyprus, paris@ucy.ac.cy

I INTRODUCTION

Annual sales of pharmaceuticals through retail pharmacies have climbed to $220.2 billion in 2010.[3] In 2001 over 4 billion antidepressant drugs were sold in the U.S., generating over $12 billion in revenues — over $150 million of which was spent on advertising. In the pharmaceutical industry innovation occurs rapidly. In 2010 a patient in the U.S. can be prescribed one of 64 antidepressants (brands and generics) relative to 47 drugs in 2001 and 19 in 1980. Patients, therefore, are unlikely to be aware of all drugs available to them. This is ether because patients are unaware of the existence of a drug, because patients are unaware of the possibility that a drug could be used in their treatment, or simply because patients are unaware of their illness. More importantly, patients' choice sets are created with limited information, a common attribute of many industries. This problem is exacerbated in the antidepressant market, where symptomatology of the disease, major clinical depression, and response to pharmacological treatment is idiosyncratic.

I formulate an empirical methodology that quantifies patient welfare from pharmaceutical innovation in the U.S. antidepressant market and I isolate the effect of promotional efforts by pharmaceutical firms on welfare. I utilize a structural discrete choice model with observed and unobserved consumer heterogeneity and limited information on the part of consumers adding to the existing literature.[4] In doing so, I show, first, that traditional models, which assume consumers are aware of all products for sale at the time of purchase, generate inconsistent estimates of drug-specific demand curves that are biased towards being too elastic. I, therefore, use data on promotional activity by pharmaceutical firms directed towards physicians and towards patients to account for informative advertising. Second, I employ additional information on media exposure to estimate a model of limited information which improves the estimated price elasticities. In many industries, data on individual exposure to advertising are difficult to obtain. However, variation in advertizing exposure across households is an important source of consumer heterogeneity. I combine macro-level advertising data with micro-level data relating consumer attributes to media exposure, thereby permitting a model which allows for individual heterogeneity in choice sets and advertising media exposure while having limited data connecting consumers to purchases and advertising.

[3] 12 months to March 2010; source: IMS Health Inc.,

[4] Berry (1994), Berry, Levinsohn and Pakes [BLP] (1995, 2004), Petrin (2002), Cleanthous (2003, 2009), Sovinsky Goeree (2008).

Recent structural studies of advertising incorporate micro purchase and advertising exposure data.[5] Sovinsky Goeree (2008) augments the stuctural work by Berry, Levinsohn, Pakes [BLP] (1995) and Petrin (2002) to add limited information and relax the assumption of full awareness. In the pharmaceutical industry, Berndt et al. (1994) examine non-structurally product-level demand for anti-ulcer medications. They concentrate on marketing variables, and distinguish between 'industry-expanding' and 'rivalrous' marketing efforts by looking at a natural experiment: the introduction of Tagamet and, later, Zantac.

The limited information model as introduced by Sovinsky Goeree (2008) incorporates three important sources of consumer heterogeneity: choice sets, tastes, and advertising media exposure. The results suggest that advertising has very different informative effects across types of agents, individual consumers and across media, and that allowing for heterogeneity in patient information yields more realistic estimates of demand elasticities. The results also suggest that assuming full information may lead to incorrect conclusions regarding welfare effects of innovation and, consequently, the size of the moral hazard that arises due to the existence of prescription drug insurance coverage. Indeed, I found larger welfare effects in the antidepressant market than in the traditional, full information model, suggesting that the market is less competitive than expected. Finally, informative advertising exacerbates moral hazard and in the presence of prescription insurance, combined with high incomes, demand for antidepressants becomes very inelastic.

For demand estimation, I use a simulated method of moments algorithm since demand aggregation involves the computation of multi-dimensional integrals for which there is no analytical solution. The estimated demand parameters provide marginal utilities or disutilities of drug side effects and help compute own- and cross-price and advertising elasticities of demand (with full and limited information), which describe patient substitution patterns. To estimate welfare gains from a new drug, I calculate the upper bound for the average patient surplus when all welfare gains at the time of introduction are attributed to that innovation. I then compute a lower bound when the new drug is excluded from the choice set at the time of innovation. The latter is a closer representation of the true welfare gains due to innovation. Gains per average daily dosage help evaluate the patients' willingness-to-pay in excess of the price charged. Annual prescription gains represent the additional amount patients are willing to forgo in a year in order to afford each drug. Relative

[5] Erdem and Keane (1996), Anand and Shachar (2010), Ackerberg (2001, 2003), Shum (2004).

gains help evaluate the importance and success of different innovations in the antidepressant market. Finally, comparing different models of full and limited information, I show that promotional efforts by pharmaceutical firms lead to increases in patient welfare and at the same time increases in the moral hazard gap.

The remainder of the paper is organized as follows. Section II analyzes the characteristics of the market for antidepressants and the pertinent characteristics of the pharmaceutical industry. Section III discusses the data. Section IV presents the empirical methodology in estimating a limited information random coefficients multinomial logit model for antidepressants. The results are presented and discussed in Section V. Section VI uses the demand estimation results to infer welfare implications of the varied promotional activity by pharmaceutical firms in antidepressants. Section VII concludes.

II MARKET BACKGROUND

The pharmaceutical industry is characterized by an impressive stream of new products, especially over the latter half of the twentieth century, due to rigorous research and development.[6] In fact, the pharmaceutical industry is the most research-intensive U.S. manufacturing industry. The patent system is in place to ensure that there is sufficient incentive for innovation to take place and that the high costs of research and development can be recouped. As Table I shows research and development expenditure for the industry increased from $1.5 billion in 1980 to $40 billion in 2008.

INSERT TABLE I ABOUT HERE

During the life of the patent, the innovator firm has a legal monopoly on the sale of a particular drug. Following the expiration of a patent, generic competitors may enter the market following FDA approval. To obtain this approval, a generic manufacturer must demonstrate that its product is biologically equivalent to the innovator drug.[7] Prior to patent expiration and the advent of generic competition, an innovator drug may experience competition from pre-existing or new drugs of different chemical make-up and which offer a therapeutic substitute in the treatment of the

[6] Scherer (2010) presents in detail the workings of the pharmaceutical industry.

[7] Biological or therapeutic equivalence means a drug acts on the body with the same strength and similar bioavailability as the same dosage of a sample of another drug of the same active ingredient when the route of administration is the same.

relevant condition. The latter could be *me-too* entry, that is, the new drug fights the disease in a manner copied from and closely similar to that of the rival. This would categorize drugs as being of the same 'type'.[8]

Estimating the demand for pharmaceutical products is challenging for two reasons. First, most pharmaceutical products in the United States must be prescribed by a physician. This implies that a third party makes the product choice most of the time. Second, most patients have some sort of insurance that may or may not include drug-reimbursement, and may or may not cover all drugs in the choice set. Moreover, the demand for pharmaceuticals is highly price-insensitive, and the more acute the illness the higher the insensitivity. The insensitivity is exacerbated by higher income and by insurance coverage.

Another unique industry characteristic is the unusually vigorous advertising and varied promotional activity. Spurred by product novelty, trademarking, and the difficulty consumers and prescribing physicians have in becoming informed about the efficacy of drug products, expenditure on promotional activities for pharmaceutical products ranks high among industries for both prescription and over-the-counter drugs. As reported in Table I, advertising expenditure has been in an upward trend since 1980 but has been in decline recently reaching a total of $39 billion in 2008 ranking second to the automotive industry. Advertising as a percentage of sales has also been declining in the 2000s dropping from a 6.9% advertising-to-sales ratio in 2000 to 4.2% in 2008. Similarly, advertising-to-margin ratios have been declining dropping from 9.5% in 2000 to 5.5% in 2008. The picture for pharmaceutical sales is different, still in an upward trend starting at $32.1 billion in 1980 and reaching $928 billion in 2008 an almost 30-fold rise.

Advertising in pharmaceuticals can be directed towards physicians, in the form of sales representative detailing, professional journal advertising and samples, or towards patients, also know as direct-to-consumer (DTC) advertising. Detailing is when a sales representative of a manufacturer of drugs calls on office-based physicians, hospital-based physicians, directors of pharmacies, and other professional distributors to promote new drugs. It is reportedly the primary information source to 57 percent of physicians; 85 percent give the process a "strong vote of confidence," because of

[8] Me-too entry into the market can also be by trademarked drugs of the same chemical entity as the innovator drug that nevertheless differ in the type of administration, in strength, and might specialize in attacking specific symptoms of the disease.

the valuable information it provides.[9] However, in terms of total advertising spending, detailing ranks second to samples, as shown in Table I. In the 2000s, it has decreased as a proportion of total advertising, to 17%, and a total of $6.5 billion by 2008. Professional journal advertising reflects advertising expenditures for prescription drugs appearing in medical journals. It is the lowest spender among the advertising methods. Though it has dropped to 1% of total advertising spending by 2008, it has not changed in value ($0.4 billion) Samples are prescription drugs given to physicians to disseminate freely to patients. These are reported in Table I in terms of their retail value, factoring in the opportunity cost of the firms that give them up. However, this may be upsetting the reported percentages of total advertising spending. In fact, samples have climbed to 71% of total spending by 2008. It is possible that the rising drug prices as is evident also from the steep increases in sales revenues, are causing this.

INSERT TABLE II ABOUT HERE

Direct-to-consumer advertising includes advertising for prescription drugs on television, radio, magazines and newspapers, as well as internet and outdoor advertising. DTC advertising, though in existence before 1997, it has increased rapidly ever since the law was relaxed to allow prescription drug advertising direclty to consumers. As reported in Table I, DTC advertising climbed to 16% of total advertising by 2000 and dropped to 11.3% by 2008, though in value terms it has increased from $4.2 to $4.4 billion. Table II reports the breakout of DTC advertising into the different media over time. Television holds the largest portion with 63.6% of total spending on media in 2008, a drop from 66.3% in 1995; print advertising has increased from 29.6% in 1995 to 30.9% in 2008; radio and outdoor advertising have both decreased from 0.7% and 3.6% in 1995 to 0.3% and 2% in 2008, respectively. The internet has only been used as a medium since the early 2000s, and even though it accounts for the decreases in other media, it has also seen a drop relative to 2005, standing at 3.2% in 2008.

INSERT TABLE III ABOUT HERE

Table III reports similar results to Table II on media breakout but for the top 10 firms in the pharmaceutical industry. Total advertising is heavily concentrated among the top 10 pharmaceutical firms. It increased from 53.3% in 1995 to 62.7% in 2000 and dropped to 40.9% in

[9] Scherer, F. M. (2010).

2005. Television advertising is the biggest spender among the different media for the top 10 firms throughout the years, but has dropped from 78.4% in 1995 to 65.5% in 2005.[10] Print advertising is second, and has similarly been on the rise: from 21.6% in 1995, top 10 firms accounted for 29.6% of total print advertising in 2005. The rest is spent on the internet, outdoor and radio advertising. There is a rise in the latter, which is more evident when we look at individual firms. Most of the increase is due to rising internet advertising, where besides improvements in technology and the ease it offers, the need for anonymity has driven a lot of patients to fill their prescriptions online.

i The Market for Antidepressant Drugs[11]

I concentrate on the market for antidepressants that includes prescription drugs[12] that are FDA-approved to be used in the pharmacological treatment of clinical depression. I take the antidepressant class as described by IMS Health Inc., USC codes 64300-64399. Treatment of depression does not require combinations of drugs from different categories and antidepressants are not used to treat diseases other than depression, thus elimintating market interaction that would complicate modeling demand.

Antidepressant drugs are used in the pharmacological treatment of clinical depression, a highly prevalent disease with a lifetime and annual prevalence of 17% and 10%, respectively.[13] Approximately fifty percent of Americans suffering from major depression seek professional care during a year and of those only about half go to psychiatrists.[14] Under-diagnosis and under-treatment may be due to various causes: Patients may not link their symptoms to a disease; public comprehension of mental diseases is generally poor; depression still constitutes a social stigma and primary care physicians miss diagnosing depression half of the time.[15] The lack of information is evidenlty key in this market and research on informative advertising in antidepressants becomes very important. There currently exists no definitive biological test for the diagnosis of depression. Consequently, the psychiatrist diagnoses depression with only the symptoms of a patient, the patient's medical

[10] Note that in 2000, the data reported TV and radio together, but in 2005 TV was reported independently and radio was grouped together with outdoor and internet advertising. Therefore, though, we observe a decrease in the proportion of TV advertising, the decrease is less than what is reported in the table.

[11] Cleanthous (2003) describes the market for antidepressants in full detail.

[12] There are no over-the-counter antidepressants.

[13] National Comorbidity Survey with data updated as of July 19, 2007.

[14] Miranda (1994), Badamgarav et al (2003).

[15] Salmans (1997), Badamgarav et al (2003).

history and the medical history of the patient's family, since depression is believed to be genetic. Symptomatology of depressed patients is idiosyncratic.

A therapeutic subdivision also involves categorizing drugs into types according to the way they act in curing a disease (mechanism of action). There exist 7 main types of antidepressant drugs, for example, Selective Serotonin Reuptake Inhibitors (SSRI). Types are further subdivided into collections of drugs with the same molecule (active ingredient), for example, fluoxetine (generic), Prozac, Sarafem, Prozac Weekly. The first two antidepressants were introduced in the late 1950s. Expansion in the market continued steadily with the introduction of new drugs, molecules and types. With the entry of the first SSRI, fluoxetine (Prozac) in 1988, unprecedented media attention proclaimed Prozac "a wonder drug," due to the marketing efforts of Lilly and its less severe side effects.

Treatments other than the pharmacological treatment of depression using antidepressants will be collectively referred to as the outside option. This option also includes the possibility of no treatment at all. Once a decision has been made in favor of a pharmacological treatment for depression using antidepressant medication then the choice is one among the available antidepressants at the time of choice. Table IV lists a combination of all the possible choices in antidepressants that appeared at least once over the 22-year period of the dataset used. The table divides the antidepressant medications into their different types and molecules. For instance, a choice of a specific drug among Prozac, Sarafem, Prozac Weekly or the generic alternative presupposes a choice of molecule, in this case Fluoxetine Hydrochloride, which in turn presupposes a choice of type of antidepressant medication, here SSRI. Note that the choice in antidepressants should be viewed as simultaneous rather than hierarchical. The divisions into groups are market segmentation characteristics and help the choice maker in matching tastes and preferences to drug characteristics.

INSERT TABLE IV ABOUT HERE

Historical evidence indicates that no one antidepressant is clearly more effective than another in achieving the desired health outcome.[16] A major source of differentiation, therefore, is the mechanism of action of an antidepressant as this is identified by a drug's type. Another major source of differentiation is an antidepressant's side effect profile that is common to drugs of the

[16] Depression Guideline Panel (1993).

same active ingredient (molecule).

In most industries consumers choose the product, the quantity and the method of payment. In the case of prescription drugs the decision is shared by the patient, the physician and sometimes the prescription drug coverage provider. If a patient were left alone to make a decision, she would base that decision on the expected health outcome of a treatment and the cost of the treatment net of any insurance co-payment. A patient's expectation on a health outcome depends on her information about the treatment, which in turn depends on factors like health awareness, direct-to-consumer advertising, word-of-mouth, personal experience with antidepressants or medication for symptomatically similar diseases. However, legislation prevents and protects the patient from making an uninformed decision by requiring that a prescribing physician makes the treatment choice. The patient, therefore, can only participate in the optimization of her utility by trying to affect the physician's preferences. It is reasonable to assume that drug-prescribing physicians care about their patients and, thus, try to maximize their patients' utility.

In the case of depression, patients are highly heterogeneous in their response to treatment, hence, experience with other patients should only influence a physician's decision initially. For the same reason, existing protocols and guidelines for the treatment of depression are merely suggestive in nature.[17] What is more, existing formularies[18] only make a distinction between branded and generic antidepressants and not across types and molecules. The initial choice of an antidepressant type and molecule is based on the patient's own or her family's medical history. In the absence of a medical history, physicians start an experimentation phase; often, a physician will begin with antidepressants with the least overall side effects: some SSRI, TCA. Therapeutic effects appear within two to six weeks. Treatment of depression typically takes much longer. It may vary from a few years in the cases of mild depression to a person's life span. This implies that a patient's initial experimentation phase is short-lived and will not affect the long-term market shares in antidepressants. The brevity of the experimentation phase (six months on average) as compared to total treatment time justifies that annual data captures all learning.

Scientists do not currently have definitive biological tests that can be administered to humans to predict exact response to a particular treatment. Prescribing physicians have to rely on their

[17] Depression Guideline Panel (1993).

[18] The Lewin Group (2000).

patients to find out whether a certain pharmacological treatment is working out or not. As a result, patients influence the physician's choice in antidepressants. Moreover, it is highly unlikely that a physician would change types of antidepressants during the continuation phase of a treatment for price considerations due to the difference in the way different-type drugs are believed to fight depression.

The major effect of price in the case of antidepressants is in the choice between branded and generic drugs, where the difference in price is more pronounced. Interviews with physicians have revealed that in most cases a physician would prescribe a molecule, not a specific drug, especially when the generic is available. A physician would consider choosing the branded drug if the patient asks him to. With a molecule prescription, a patient could choose to buy the branded version at the pharmacy. Since all antidepressant drugs of the same molecule are bioequivalent they should be perfect substitutes in demand. The data show otherwise. This is because patients tend to perceive the physically identical branded and generic drugs as different in quality. The decision to buy brand over generic is influenced by the patient's perception of quality and the price difference (after insurance) between two drugs. This is exacerbated by DTC advertising.

Table I reports antidepressant sales and advertising trends over time. Similarly to the rest of the pharmaceutical industry, advertising-to-sales ratios in antidepressants have been steadily rising, reaching 9.7% in 2000 and dropping to 8.1% in 2005. Sales and advertising expenditures have been on the rise reaching $12.5 and $1.02 billion, respectively, in 2008. Regarding the different channels of advertising, the antidepressant market resembles the rest of the pharmaceutical industry when we look at trends over time, however, in absolute terms detailing and journal advertising are relatively more important than samples and DTC in antidepressants than in the rest of the industry.

INSERT FIGURE 1 ABOUT HERE

The evolution of DTC advertising in total spending and by medium is depicted in Figure 1. The graph shows a clear increase in all channels of advertising, apart from journal advertising. Two important kinks in the graph are in 1988 and 1997. Prozac was introduced in the market in 1988 and acted as a catalyst to the new era of advertising in pharmaceuticals. Being a new drug, new molecule and new mechanism of action, it was necessary for Eli Lilly to inform physicians about the 'wonder' drug and the way it tackles the disease in the brain. At the same time, the company

tried to inform patients about the drug and the disease through provision of free samples, but also through popular press releases and popular literature. In 1997, with the relaxation of laws on DTC advertising regarding prescription drugs, spending on DTC advertising increased and has been steadily on the rise. It is also evident from the graph, that most advertising expenditure in antidepressants goes into samples. Detailing is secondary in importance but also quite large and rising. DTC is also on the rise but still premature for the span of the dataset.

INSERT FIGURE 2 ABOUT HERE

Figure 2 analyzes the evolution of DTC advertising alone, in antidepressants, and separates drugs into pre-2001 market entry, therefore, included in this paper's estimation and post-2001 market entry, the newer antidepressants.[19] Two trends are evident. First, DTC advertising climbs high very fast and suddently drops. This is mostly evident in 2002 for older drugs, the year Prozac went off patent and there was generic introduction of its molecule, fluoxetine; for newer drugs, this is mostly evident in 2005. What is happening is that DTC advertising is used to inform patientss about the drug in order to include it in their choice set. When the drug is off patent, DTC advertising can no longer be effective and is no longer advertised to consumers. However, DTC advertising reduces much earlier than patent expiration. This agrees with the informative nature of DTC advertising. Firms use it to inform consumers about the existence of the drug and the disease. Once the drug enters consumers' information sets, patients will inform their physician about the drug, in case the physician fails to suggest it. There is no need for DTC advertising when most consumers have been exposed to the advertisement, unless there is a persuasive role, not so evident in the case of antidepressants. Second, DTC advertising follows a normal distribution over time for older drugs. This is due to the fact that firms spend to advertise their new drugs up until the point they go off-patent and generic introduction ensues. In fact, as mentioned above, DTC spending decreases sharply much before patent expiration. As all older drugs have been introduced before 2001, a large proportion of these drugs have gone off-patent by the year 2009, and for the rest, the informative role of the advertising diminishes.

INSERT FIGURE 3 ABOUT HERE

[19] Data on DTC advertising are monthly and are available upto the first quarter of 2010. However, these data cannot be used post-2001 since we lack data on quantity sales, revenues and other advertising media post-2001.

Finally, Figure 3 breaks up DTC advertising for all antidepressants into 3 media. As is the case in the rest of the pharmaceutical industry there is an upward trend in television advertising. The only drop is in 2005, which is characterized by the end of many older campaings and start of many new drug campaigns. Print advertising follows a wavy, yet slightly upward pattern as well. Internet advertising is mostly going up, as is the case in all of the industry.

III DATA

Data are for the pharmaceutical preparations industry (SIC 2834) and include market shares, advertising and prices, drug characteristics (physical or otherwise) and distribution on patient demographics. Antidepressant sales and advertising data come from IMS Health Inc. and are complete unlike data on other therapeutic areas. Sales data are national data on quantities and prices for each antidepressant drug reported on an annual basis: Quantities are in extended units (adjusted by preparation); prices are wholesale and are aggregated by drug. All values are deflated using the Consumer Price Index of the Bureau of Labor Statistics with 1980 as the base year.

Data for advertising expenses[20] and their breakout into different channels and media, in the case of DTC advertising, come from IMS Health Inc., Integrated Promotional Service, Competitive Media Reporting, the Strategy dataset of Kantar Media, LNA and Schonfeld & Associates via Advertising Age. Direct-to-physician advertising data consist of professional journal advertising, sales representative detailing and the samples. The former is reported as total expenditure; detailing is reported as the total number of contacts wth physicians, the total number of minutes spent during contacts, and the total cost; finally, the retail value of samples is reported. For direct-to-consumer advertising, the total value is reported as well as, spending by medium:[21] Sunday, local, and Spanish-language magazines; national and Spanish-language newspapers; network, national spot and local television; cable networks; internet; outdoor and radio. In order to cover the whole span of the available data, DTC advertising was regroupped into three bigger categories: television, print and internet, radio and outdoor expenditures.

Data run for 22 years from 1980 to 2001. In this period the market for antidepressants includes

[20] They exclude promotion spending for professional meetings and events.

[21] Data are for the 'medicines & proprietary remedies' category (aggregated from TNS classifications) by LNA, which include pharmaceutical houses, medicines & proprietary remedies, fitness, eye glasses, medical equipment. As this is a different categorization than 'pharmaceutical preparations', SIC code 2834, aggregate pharmaceutical industry results may differ in some tables.

a total of 47 drugs, which are sub-divided into 24 molecules. These are in turn grouped into 7 types of antidepressants according to their mechanism of action. These constitute market-segmentation characteristics, as does the distinction between branded and generic drugs. Model estimation uses dummy variables to account for these. 28 antidepressant innovations took place from 1980 to 2001, 12 of which were molecular innovations of different types, both branded and generic. Drug observed characteristics include side effects, a drug's half-life, market and revenue shares and they come from the Drug Information Handbook, Physician's Desk Reference, Depression Guideline Panel and the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). Side effects included separately in the estimation at certain specifications are the drug's fatality rate, weight gain of more than 6 kilograms, and insomnia and/or agitation. The rest of the side effects are averaged together in one variable. 'Side effects (5)' includes anticholinergic, drowsiness, orthostatic hypotension, cardiac arrhythmias, and gastrointestinal distress effects; 'side effects (6)' includes in addition insomnia and/or agitation; 'side effects (8)' includes all. Side effects are rated between 0 and 4+, that is, a range from absent (or rare) to very common side effects. These rates are averaged over individuals, dosage regimens and bioavailability (half life). Finally, the data include information on market segmentation characteristics. Media exposure data linked to purchases come from a match between ACNielsen Homescan(R) Rx/OTC Consumer Panel and Nielsen Media Research. In the former patient demographics are matched to purchases and the latter collects data on media habits and demographics.

Demographic data are available for the 22-year span. Income and population data are taken from the Bureau of Economic Analysis and the Current Population Survey whereas prescription drug insurance data were provided by the Centers for Disease Control and Prevention of the National Center for Health Statistics (NCHS), specifically, combined data on income and out-of-pocket prescription drug expenditures,[22] which helped construct the joint distribution of income and prescription drug insurance. These data are important for allowing patient heterogeneity in prices and preference for branded drugs.

The market size is assumed to be the number of possible consumers in the antidepressant market, specifically, the portion of the population that is estimated to be clinically depressed.[23]

[22] Prescription drug expenditures net of any prescription drug insurance coverage (private or public).

[23] Studies show that the portion of the clinically depressed population that actually seeks medical assistance is about half the portion of the population that is estimated to be clinically depressed. What is more, people who do

Time-varying data on the prevalence of clinical depression from the National Comorbidity Study is used to construct the annual market size and the correct market shares. This procedure is salient in estimating the correct welfare effects.

IV EMPIRICAL MODEL

I employ a structural estimation strategy[24] whereby I specify a model of patient choice with limited information just as in Sovinsky Goeree (2008) and derive the implied relationships among choice probabilities. The model is solved using data on price, quantity and other observed drug characteristics, data on patient and physician preferences, data on advertising quantities and expenditures across media, micro-level data linking patients to media and to the drugs they purchase, and the notion of equilibrium.

I model demand for antidepressant drugs as demand for their characteristics, where each drug is defined as a set of characteristics. Patients are modeled as having heterogeneous tastes, placing different utility weights on these characteristics. In the case of antidepressants, each patient only consumes one antidepressant at a time. Therefore, patient choice in antidepressants is best described by a discrete choice model of demand at the patient level where the key assumption is that each patient buys at most one unit of the good. The full random coefficient logit model allows for patient heterogeneity using both a multivariate normal distribution and demographic data on prescription drug insurance and income.

A utility maximizing patient i, where $i = 1, ..., I$, in a given time period t, where $t = 1, ..., T$, faces $J_t + 1$ alternatives: J_t different antidepressant drugs and the option of not purchasing any of the drugs, the outside option, $j = 0$. At every time period t, each patient maximizes her level of indirect utility as follows:

$$\max_{j \in \{0,1,...J_t\}} u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt}, \tag{1}$$

where $\delta_j \equiv \alpha p_j + x_j'\beta + \xi_j$ is the mean utility level, a drug-specific term common to all patients. The β are the marginal utilities of the drug's observed characteristics, x_j, the α is the marginal

not seek medical assistance are not always consciously doing so. They can be unaware of the fact that depression is not merely a mood but rather a debilitating disease and tend to ignore depressive symptoms altogether. Since doctors depend on their patients to confirm the existence of symptoms, people who are not clinically depressed may seek to purchase antidepressants for various other reasons like the treatment of mild anxiety disorders or entertainment. Information on the latter two cases is not widely available but I assume their net effect to be small.

[24] The model builds on BLP (1995), Petrin (2002), Cleanthous (2003).

disutility associated with price, p_j and ξ_j are drug characteristics unobservable to the researcher. Note that this formulation of utility specifies that the unobserved characteristic is identical for all patients. By letting the price coefficient vary across patients in the full random coefficients model the ξ_j captures the elements of vertical product differentiation in the antidepressant market. The term μ_{ijt} captures the heterogeneity in patient preferences for observed (by the econometrician) drug characteristics and ϵ_{ijt} is a mean zero random utility component distributed i.i.d. type I extreme value across both drugs and patients. The sum of the latter two components represents the deviation from the mean utility level for each patient i and is a measure of the idiosyncratic valuation of drug j's characteristics.

i Demand Model Including Information on Patient Heterogeneity

In the random coefficients logit (RCL) patient heterogeneity in preferences will be captured. A patient's substitution to a new drug due to an increase in the price of the initial drug chosen will depend on the attributes of her initial choice, her income and her prescription drug coverage. As already mentioned, the model will be estimated, first, by assuming a multivariate normal distribution of patient tastes. Then, more precise estimates will be obtained by using information on the distribution of patient preferences as it relates to some of the drug characteristics. The benefit of this methodology is that it does not require observations on patient purchase decisions to estimate the demand parameters. The use of demographics provide ample information on patients' substitution patterns and, hence, estimates of the parameters of the distribution of patient preferences are precise.

In the absence of patient-level data, I use aggregate-level information that relate average patient demographics to some of the drug characteristics. Individual patient characteristics are modeled as a combination of an observed component (patient demographics), D_i, and an unobserved component, τ_i assumed to be independent. This allows the inclusion of information about the distribution of the marginal disutilities of price and the preference for branded drugs obtained from demographic data.

Combining the demand parameters in δ_{jt} and μ_{ijt} the overall effect of observed characteristics

on utility can be encapsulated by α_i and β_i as expressed below:

$$\begin{pmatrix}\alpha_i \\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \Pi D_i + \Lambda\tau_i,\ D_i \tilde{} P_D^*(D),\ \tau_i \tilde{} P_\tau^*(\tau) \tag{2}$$

where $P_\tau^*(.)$ is a parametric distribution that has a standard multivariate normal distribution, and $P_D^*(.)$ a non-parametric distribution derived from the data. Π is a matrix of coefficients that measure the relation between demographics and patient preferences and Λ is a matrix of parameters. Hence, Λ allows a different variance for each component of τ_i and a correlation among these patient preferences. The utility can now be rewritten as follows:

$$\begin{aligned} u_{ijt} &= \delta_{jt} + \mu_{ijt} + \epsilon_{ijt} \\ \text{where } \delta_{jt} &= \alpha p_{jt} + x'_{jt}\beta + \xi_{jt},\ \mu_{ijt} = [p_{jt}, x_{jt}] * (\Pi D_{it} + \Lambda\boldsymbol{\tau}_{it}) \end{aligned} \tag{3}$$

and the probability that a patient chooses drug j at time t as expressed in equation can now be written using Bayes rule and under the distributional assumptions as $s_{jt}\left(x_{jt}, \xi_{jt}, p_{jt}; \theta_2\right) = \int_{A_{jt}(\delta)} dP^*(D,\tau,\epsilon) = \int_{A_{jt}(\delta)} dP^*(\epsilon|D,\tau)dP^*(\tau|D)dP^*(D) = \int_{A_{jt}(\delta)} dP_\epsilon^*(\epsilon)dP_\tau^*(\tau)dP_D^*(D)$, where $P^*(.)$ are population distribution functions. With an extreme value distribution for the random utility component, the RCL market shares become:

$$s_{ijt} = \frac{\exp\left(\delta_{jt} + \mu_{ijt}\right)}{1 + \sum_{k \in J_t} \exp\left(\delta_{kt} + \mu_{ikt}\right)}. \tag{4}$$

where s_{ijt} represents the probability that patient type i will purchase drug j at time t. The price elasticities of demand follow:

$$\eta_{s_{jt},p_{kt}} = \frac{\partial s_{jt}}{\partial p_{kt}} \cdot \frac{p_{kt}}{s_{jt}} = \left\{ \begin{array}{l} \frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) dP_\tau^*(\tau) dP_D^*(D),\ j = k \\ -\frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dP_\tau^*(\tau) dP_D^*(D),\ j \neq k \end{array} \right\} \tag{5}$$

Each patient type has a different price sensitivity and this varies by drug. The weighted average of this sensitivity is calculated using as weights the patient-specific purchase probabilities. Own and cross-price elasticities of demand are not just the result of the logit function and cross-elasticities are larger for products that are closer in terms of their characteristics.

ii Incorporating Consumer Awareness

I now allow for the purchase probability of each patient i to also depend on her awareness of each drug j and, consequently, on her awareness of drugs that are competing against j. When a patient becomes aware of the existence of a drug, the drug enters her choice set. Given her choice set, the patient then buys the product, or not, exactly as in the previous analysis. The implicit assumption is that pharmaceutical advertising alerts the patient to the drug's existence, directly or via a physician, and thereby increases the probability that the drug enters the patient's choice set.

Let C_j be the set of all possible choice sets that include drug j and let ϕ_{ijt} be the probability that patient i is informed about drug j at time period t. The information component, ϕ_{ijt}, describes the effectiveness of advertising at informing patients and physicians about drugs. Assuming patients are aware of the outside option with probability one, and given all assumptions mentioned above, the conditional probability that patient i purchases drug j is given by

$$s_{ijt} = \sum_{S \in C_j} \prod_{l \in S} \phi_{ilt} \prod_{k \notin S} (1 - \phi_{ikt}) \frac{\exp\left(\delta_{jt} + \mu_{ijt}\right)}{1 + \sum_{r \in S} \exp\left(\delta_{rt} + \mu_{irt}\right)}. \tag{6}$$

The outside sum is over all the different choice sets that include drug j. Advertising affects demand through the awareness function ϕ_{ijt}, which describes the effectiveness of advertising at informing the choice makers. The awareness is modeled as a function of drug j's advertising by medium, observed drug characteristics, and unobserved idiosyncratic patient-advertising-medium-specific effects. The awareness function for patient i is given by

$$\begin{aligned} \phi_{ijt}\left(\theta_{\phi}\right) &= \frac{\exp\left(\gamma_{jt} + \lambda_{ijt}\right)}{1 + \exp\left(\gamma_{jt} + \lambda_{ijt}\right)} \text{ where} \\ \gamma_{jt} &= \alpha'_{jt}\left(\varphi + \rho\alpha_{jt}\right) + x_j^{age}(\vartheta + \psi x_j^{age}) \\ \lambda_{ijt} &= \alpha'_{jt}\left(\Upsilon D_{it}^{n}\zeta + \kappa_i\right) + \widetilde{D'}_{it}\lambda, \ \ln\kappa_i \tilde{} N\left(0, I_m\right) \end{aligned} \tag{7}$$

where $\theta_{\phi} = (\lambda, \varphi, \rho, \zeta)$. The advertisements for drug j are grouped under the different direct-to-consumer and direct-to-physician advertising methods and are represented by the m-dimensional vector, α_{jt}. The $\widetilde{d}$-dimensional vector λ measures the fraction of patients of type $\widetilde{D}_i$ who are informed without seeing any advertising, a subset of observed patient characteristics D[25]; φ measures

[25] It consists of data such as a constant, dummy variables for prescription drug insurance coverage and various

advertising effectiveness; ρ captures the decreasing or increasing effectiveness of advertising. Together φ and ρ capture how variation in advertising effectiveness varies across drugs. ϑ captures the effect of a drug's age (patients may know a drug the longer it has been on the market) and ψ captures the decreasing or increasing effect of age; the hypothesis is that when a drug is too old, then the informational effect of age is decreasing. Υ captures how advertising media's effectiveness varies by observed patient characteristics. I estimate the parameters of the Υ first using the Nielsen data. I use these estimates in the model and estimate the other parameters matching observed market shares to predicted shares as in BLP. Therefore, ΥD_i^n is the exposure of patient i to medium m, and $\alpha_j' \Upsilon D_i^n$ is the exposure of i to advertisements for drug j: ζ measures the effect of this ad exposure on the information set. The matrix $\Pi_{m \times d}$ will now capture how advertising media effectiveness varies by observed patient characteristics D^n. The stochastic patient-medium-specific term κ_{im} represents additional unobserved patient heterogeneity with regard to advertising medium effectiveness. These include patient attributes that may influence the effectiveness of medium m at informing the patient, but that are not uncovered by observed demographic characteristics. I assume κ is independent of other unobservables.

As advertising increases, the awareness function approaches one but it is non-zero even for zero advertisements since some patients are informed even if there is no advertising. The magnitude of the probability that a patient is informed when no advertising occurs is determined by $\widetilde{D'}_{it}\lambda$. Patients may get informed through word-of-mouth or the internet, for instance; physicians may become aware about the existence of new drugs by reading articles in medical journals or interacting with other physicians at conferences.[26]

The conditional probability that i purchases j, s_{ij} , is now given in (6). Market share is a function of prices and advertising of all drugs. The smaller the ϕ_{ij}, the smaller the drug market share. If ϕ_{ij} were equal to one for all drugs, market share would be the standard full information choice probability. Demand for j at time t is $M_t s_{jt}$, where M_t is the market size given by the portion

income groups.

[26] Notice ϕ_{ij} depends upon own drug advertising only. I assume the probability a patient is informed about a drug, conditional on her attributes, is independent of the probability she is informed about any other drug, that is, information provided through advertising for one drug cannot spillover to another drug. This is an oversimplifying assumption, especially in pahrmaceuticals were advertising of the disease is common. Howerer, allowing informational spillovers would greatly complicate the model: first, the theoretical framework would have to address free-riding in advertising choices across firms and, second, one would need adequate variation in the data to empirically identify the spillover effect across drugs.

of the U.S. population that is estimated to be clinically depressed, which is directly observed.

iii Identification

I treat media exposure as exogenous to the purchase decision. To the extent that media exposure is endogenous the estimate of v in the information component will be overstated. Associated with each drug is a mean utility, which is chosen to match observed and predicted market shares. If patients were identical, then all variation in sales would be driven by variation in drug characteristics. Variation in market shares corresponding to variation in the observablr characteristics of those drugs (such as insomnia) is used to identify the parameters of mean utility (β).

While a drug may have characteristics that are preferred by many patients (high β's), it may also have characteristics that appeal to certain types of patients. Identification of the taste distribution parameters relies on information on how patients substitute. First, new drug introductions are common in the pharmaceutical industry. Variation of this sort is helpful for identification of Λ. The distribution of unobserved tastes, ν_i, is fixed over time, but the choice set of available drugs is changing over time. Variation in sales patterns over time as the choice sets change allows for identification of Λ. Second, I augment the market level data with micro data on drug choice. The extra information in the micro data allows variation in choices to mirror variation in tastes for product attributes. Correlation between $x_j D_i$ and choices identifies the Π parameters.

If patients were identical, then all variation in the information component, and induced variation in shares, would be driven by variation in advertising or the age of the drug. Variation in sales corresponding to variation in drug age identifies γ. Variation in sales corresponding to variation in advertising identifies the other parameters of γ_j . Returns to scale in media advertising (ρ_m) are identified by covariation in sales with the second derivative of a_{jm}.

One major drawback of aggregate advertising data is that I do not observe variation across patients. Normally observed variation in market shares corresponding to variation in household advertising media exposure would be necessary to identify Y and ν. The individual-level data contain useful information on media exposure across households. Variation in choices of media exposure corresponding to variation in observable consumer characteristics (D_i^n) identifies Y. Variation in sales and ad exposure ($a_j' Y D_i^n$) identifies the effect of advertising exposure on the information set (ν). Thus, the data allow me to side-step the need for observed advertising variation across

households. The other parameters of λ_{ij} which do not interact with advertising (λ) are separately identified from Π due to nonlinearities.

V ESTIMATION AND RESULTS

i Descriptive Statistics

Summary statistics on observed drug characteristics are presented in Table V. The reported characteristics show that MAOIs tend to have the most adverse side effect profile including high fatality rates, whereas newer antidepressants have better side effect profiles. It is apparent though that adverse side effects are lower for SSRIs overall. Depending on patient valuations of these different side effects, some drugs are more favorable than others on average. Due to the idiosyncracy of patient valuations, it is not possible to say which drug is more efficient for each patient. In other words, a side effect profile is used by individual patients to make decisions. Additionally, SSRIs have a closer correlation of side effect occurrence than do drugs in other types. TCAs have the highest average occurrence in five of the side effects and the lowest in two. SSRIs rank exactly the opposite to TCAs. New Generation antidepressants have the lowest average half-life and in addition the lowest fatality rates and occurrence of weight gain on average. The latter is shared with SSRIs.

INSERT TABLE V ABOUT HERE

TCAs controlled most of the market in the beginning of the period in 1980. However, their share has been steadily dropping ever since, more dramatically for revenues than for quantities. In 2001 TCAs still account for 14.5% of antidepressant sales, yet they only amass 1.2% of the revenues. MAOIs have had low revenue and quantity shares over the whole period. The advent of new generation antidepressants in 1982 is marked by an evident drop in the shares of TCAs and a negligible drop in the shares of MAOIs. Shares for new generations keep rising until 1988 when SSRIs enter the market. From then on, these shares fluctuate around the same mean until they stabilize in the latter half of the 1990s when some of the newer types are introduced (NDRI, SNRI, NaSSA). Prozac's introduction swept the market. Since the shares of both new generations and TCAs dropped, Prozac was seen as a possible substitute for both types of drugs. However, as time passed only TCA shares kept decreasing. This means that the new SSRIs being introduced were

no longer viewed as ameliorations to the side effect profiles of new generations but were viewed as better medications than TCAs. This also explains the fact that SSRI revenue shares kept rising sharply whereas their quantity shares increased at a decreasing rate.

INSERT TABLE VI ABOUT HERE

Consumer Characteristics averaged over the 22 years are presented in Table VI. Whereas costs covered by prescription drug plans follow a similar trend in the dataset as does income per capita, overall prescription drug costs rise faster. Average income per patient in 1980 dollars was $11,922 in 1980 and $17,044 in 2001 and prescription drug expenditures per patient in 1980 dollars were about $53 in 1980 and rose to about $239 in 2001. Out of these expenditures, about 30.6% was covered by a prescription plan in 1980 and 69.1% in 2001. Respectively, antidepressant expenditures per patient were $11 in 1980 and $386 in 2001. Insurance coverage for antidepressant medications changed from 17% in 1980 to 44% in 2001. This is mainly due to the fact that mental health is not included in some insurance plans. Also, depression is a chronic disease and depressed patients reach their cap much faster than the average patient. As shown in Table VI, between 1980-2001, 19.6% of patients were covered by public insurance, such as medicare and medicaid, 34.4% by private insurance, and 46% had no prescription drug insurance coverage. On average, 2% of the total U.S. population are on antidepressants over the 22 years.

Table VI shows that sample demographics are very close to actual population demographics when we average over the years with DTC data. This is because sample data strives to resemble the U.S. population. Media exposure data show that over the period in question television exposure was 4.3 hours per person per day; 2.6 hours for radio; 1 hour for newspapers and magazines; and 12 minutes on the internet. Note that the internet picked up only in the late 1990s. Average media exposure per person per day is 8.2 hours.

ii Estimation

I estimate demand parameters in the full information random coefficients model following the GMM approach introduced in BLP (1995), extended by Nevo (2000) and adapted for the pharmaceutical industry by Cleanthous (2003). For the limited information case I use the algorithm introduced

in Sovinsky Goeree (2008) extending BLP (1995).[27] To correct for the above-mentioned price endogeneity, I need to specify variables that can act as instruments for price in the demand equations. Variables that are correlated with specific functions of the observed drug prices, but are not correlated with the unobserved demand disturbances, ξ_j, are appropriate instruments. Valid instruments used are: the number of products in the market at each time period t, J_t; the number of products of the same type of antidepressants available at each time period t, J_{ct}; the number of products of the same molecule available at each time period t, J_{mt}; the time passed since generic entry took place in the same molecule.

When a patient makes a decision as to which drug to buy, she has to take into account her income and insurance status, specifically, her coverage for pharmaceutical products. Insurance can come in many types. Coverage can also take many forms within each insurance status. I incorporate these variables in the patient's decision process as described by observed demographics. These are the logarithm of a patient's income, the logarithm of squared income and insurance dummy variables simulated from the distribution of out-of-pocket prescription drug expenditures. Income and insurance variables are drawn from the joint distribution provided by NCHS. The random coefficient model is estimated with and without demographic information and results are compared.

Besides income and insurance, other idiosyncratic variables, unobserved to the econometrician, may enter a patient's decision process. The unobserved patient characteristics, τ_i, are random draws from a standard normal distribution. Draws are for 2000 individuals per time period. With the inclusion of insurance data, the estimated effect on utility of the various drug characteristics will be closer to their true values. Given the ns draws of the observed and unobserved characteristics I average over the implied logit shares. Simulation draws are held constant as the parameters change otherwise changes in the objective function would be due to simulation changes.

With an estimated model one can verify whether the estimated parameters carry the expected signs and from their magnitude infer the relevant significance of their role in a patient's decision process in choosing to purchase an antidepressant. The price sensitivity, α, is expected to be negative; it represents the disutility associated with the drug's price. The side-effect coefficients, β's, are also expected to be negative since they are taste parameters to undesirable side effects. A

[27] See Cleanthous (2003) for details on the full information model and Sovinsky Goeree (2008) for a detailed technical description of the computation algorithm for the limited information case.

drug's half-life is also expected to have a negative coefficient. The lower the coefficient the faster is takes for the drug to become available at the site of physiological activity after administration. However, some people dislike small half-lives because they increase the frequency at which they have to repeat the medication. It is ambiguous, therefore, what sign to expect for the coefficient of half-life.

In the random coefficient model, additional information is obtained from the inclusion of unobserved and observed patient characteristics. This model estimates mean effects, the means of the distribution of marginal utilities (α's and β's in equation (2)), by a minimum-distance procedure. It also estimates standard deviations (Λ in equation (2)) which are estimates of the unobserved heterogeneity about the mean effects. Finally, it estimates coefficients of demographic interactions with price and preference for 'brandness', that is, estimates of the observed heterogeneity about the mean effects (Π in equation (2)). To avoid obtaining positive values for price sensitivity in the tail of the distribution that would imply that the higher the price the higher the utility, I regress the negative of the logarithm of the α_i's on the observed and unobserved characteristics in equation (2). This restricts the overall price sensitivity to non-positive values.

iii Demand Estimation

Previous research[28] has estimated demand on antidepressants using traditional structural models, which assume that the choice maker has full information about all available products in the market. Consumer attributes that matter the most in the case of antidepressant demand are income and prescription drug insurance coverage. Research on the personal computer industry[29] has suggested that age, education, household size and marital status are significant. After various reduced form checks, I included some of these demographics in the limited information model estimated here. In the appendix, I estimate different models of demand to show that the instruments used in the full model address the endogeneity issues: the simple logit model and three nested multinomial logit models, described in the appendix, NML1a, NML1b, NML2. Two versions of each model were estimated: an OLS version and an IV version correcting for price endogeneity. The results are presented in Table A.I in the appendix. In the OLS results, estimated price sensitivity is

[28] Cleanthous (2003, 2009).

[29] Sovinsky Goeree (2008).

negative but small and decreases when moving from the logit model to either NML1 models or to NML2. Price sensitivity increases in the reverse manner in the IV counterparts, which correct for price endogeneity. This implies that correction for price endogeneity is important and necessary. When moving from each OLS model to its IV counterpart, estimated price sensitivity more than quantuples. In fact it is 17 times as big in the NML2 model. The disutility for price was correctly detected and is high; it ranges from -2.1 to -2.7.

In general, statistical significance in coefficients increases when moving from the OLS logit models rightward and up to the NML2 model. This means that market segmentation, and the additional structure it imposes in NML1 and NML2, improves on the model's prediction. The magnitude and direction of the coefficients of product characteristics is of the right order. The low and sometimes insignificant coefficients are reason to favor the full random coefficient model which places no restriction on the correlation between antidepressants. The observed high significance of most of the estimated random coefficients leads Wald tests to favor the random coefficient models (both with and without demographics) over the more restrictive specifications of the models in Table A.I.

The model that offers the best estimates in Table A.I is NML2. I, therefore, use this model to test 3 different specifications that use different aggregations of the side effects variable. Results are presented in Table A.II in the appendix. In specification (1) I use 'side effects (5)', fatality, weight gain and insomnia/agitationg; in specification (2) I use 'side effects (6)', fatality and weight gain; and in specification (3), I use 'side effects (8)' as explained in section III. Branded drug with(out) generics are dummies that take the value one for branded drugs where there has (not) been generic introduction in the respective molecule. In all specifications, the parameter estimates carry the expected signs and are similarly significant. The instruments appear to address the endogeneity of price and result in estimates for the price coefficient that are significantly higher in absolute value. The first-stage F-statistic for the IV regressions are high suggesting the instruments have power.

INSERT TABLE VII ABOUT HERE

Table VII presents the results of the OLS-NML2, IV-NML2 and RCL models together. The results of the latter are given both when no demographics are used and when demographic information is incorporated. Also the RCL models are run with full and with limited information. Columns

3-4 of Table VII show the mean effects, α's and β's, of the RCL estimates with no demographics and columns 5-6 the RCL estimates with demographics. Models in columns 3 and 5 are estimated in the traditional manner, under the full information assumption, and models in the other four columns are estimated under the limited information assumption. The standard deviations from the addition of the unobserved characteristics and estimated parameters from the addition of demographics for the limited information RCL model with demographics are then presented in Table VIII.

The price coefficient in all models is negative and significant as expected. The magnitude of the coefficient increases a lot when going from the OLS to the IV version of the NML2, indicating the power of the instruments. In the RCL models, the estimated price coefficients decrease. A decrease is also observed when we move from the no-demographics models to the corresponding models with demographics. Finally, a further decrease in the coefficients occurs when we take limited information into account. Results suggest that limited information about a drug is a contributing factor to differences in purchase outcomes and that information is distributed across patients in a nonrandom way.

The signs of fatality, weight gain and side-effects coefficients are negative as expected and statistically significant for all versions of the model. The magnitude of the fatality rate coefficient is largest in the full-information-no-demographics version and smallest in the limited-information-demographics version. Similar results are obtained for the side-effect coefficient. Insomnia/agitation obtains different results: positive significant in the NML2 models, negative insignificant in the no-demographics RCL models, and negative significant in the RCL models with demographics. Similarly to the other three side-effect coefficients, the limited information disutilities of insomnia are smaller in absolute magnitude. It seems that including informational effects in the model mitigates the disutilities of the adverse side effects. Patients are now using advertising as a determinant for drug choice and even though adverse side effects still have negative marginal valuations they are relatively less important now in influencing consumer choice; information, therefore, does play a significant role in patient choice.

The coefficient on drug's age also changes sign while moving along the different models. With random or no consumer heterogeneity, age has a positive significant effect, that is, patients prefer drugs that have been on the market for a long time. When demographics are included, I obtain

the result that is consistent with the data, that newer drugs are preferable to consumers: negative significant coefficients. If older drugs were preferebale (say, they cure a disease) there would be no incentive for pharmaceutical firms to innovate. Innovation takes place to cover unmet demand, so there must be something that new drugs are offering that old existing drugs failed to address. A side-effect-age interaction effect was included to correct this detected divergence in age coefficients: it is meant to capture the fact that older drugs are preferred only when they have low side-effects, which is intuitive. The coefficient is negative and significant and outweighs in almost all cases the positive significant effect of age, or increases the disutility that comes with age in the demographics case. Again, the limited information models obtain smaller coefficients 75% of the time. The statistical and economic significance of this results is analyzed in more detail in the full RCL model.

Finally, the coefficient on brand preference is of great importance. With an ambiguous expected direction, it is interesting to observe that estimated coefficients are positive and significant, when a generic does not exist in the same molecule as the brand, and negative significant, when a generic does exist.[30] This says that people are not swayed by perceptions that favor branded over generic drugs when the two are therapeutically equivalent. It is possible that when patients have insurance plans that fully cover for branded drugs they may opt for brand over generic since they are indifferent among the two. What should still hold, even in the case of full insurance coverage, is that branded drugs with generics would be more elastic than branded drugs without. In the case of no generics, patients could opt for me-too drugs with generics, but the highly significant, positive coefficient shows they do not. Patients, in general, place great value in their health and would opt for the drug with the highest quality. In pharmaceuticals, 'brandness' is a good proxy for perceived quality. The utility obtained from whether a drug is branded or generic also depends on an individual's income and prescription drug insurance coverage. Therefore, the two brandness coefficients, similarly to the price coefficient, are allowed to vary with demographics in the RCL model. One more time, comparing the RCL estimates between the full and limited information models, we observe a decrease in the magnitude of the negative brand effect but an increase in

[30] In the relevant literature, 'brandness' is a dummy in itself. Cleanthous (2009) did not make the distinction between branded drugs with and without generics and a higly significant, positive result was detected, which was increasing with income and prescription drug insurance. In this model, the effect for brand without generic is much higher, and significant as there is no competition and, in the case of generic competition negative significant.

magnitude in the positive coefficient. Patients still favor brand over competitive generics and with more intensity when we include informational effects in the model: patients preference for brandness is reinforced by advertising.Yet patients still opt for generic over brand when the two coexist for the same molecule.

The observed high significance of most of the estimated random coefficients leads Wald tests to favor the random coefficient models (both with and without demographics) over the more restrictive specifications of the logit models. Dummy variables for type and 'brandness' are included as drug characteristics in the RCL model to capture any correlation that may exist. The results are presented in Table VIII. The mean coefficients of the RCL models without demographics (columns 3 and 4 of Table VII) are economically and statistically significant. In the models with demographics, the price disutility is close to unity for the full information model (column 5) and decreases to below unity in the limited information model. Estimated standard deviations from the addition of the unobserved characteristics were mostly insignificant which means that the normality assumption for the distribution of the unobserved characteristics is not valid under the limited information assumption. These results reinforce the need for the use of demographics to correctly model patient decisions, especially when it comes to the choice of characteristics such as price and brandness.

INSERT TABLE VIII ABOUT HERE

Table VIII presents demand estimation results for the full random coefficients model using demographic income and prescription drug insurance described above. The model was estimated using instrumental variables to correct for price endogeneity. The first column lists the means of the distributions of marginal utilities and disutilities, α's and β's, of antidepressant characteristics. The only mean effect with an insignificant coefficient is the side-effect-age interaction. The rest have significant coefficients both statistically and economically. The estimates of the heterogeneity around these means are presented in the other columns of the table. The second column tests the standard deviations which are parameters that capture the effect of the unobserved patient preferences. These effects are half of the times statistically and economically significant. The last three columns present the effects of demographics (observed patient characteristics) on the mean coefficients. These estimates are almost all statistically and economically significant.

All adverse side effects have negative mean coefficients and two have relatively large and in-

significant standard deviations. The negative coefficients suggest that the average patient gets more disutility the more these side-effects occur. The estimated standard deviations are estimates of the random patient heterogeneity around these means. Since some of these are relatively large, this means that some of the adverse effects are not viewed as adverse by some patients in the simulated sample. There is no immediate explanation behind the insignificant standard deviations. A possible explanation for this result is that most depressed people are very similar in their valuation of these side-effects.The negative coefficient on fatality differs. It has a small but significant standard deviation, thus, all patients have a disutility from this side effect. Though a large negative coefficient was obtained for insomnia/agitation implying high disutility for the average consumer, it is interesting to see that a relatively large standard deviation was estimated implying that some patients obtain positive utility from occurrence of this characteristic.

Half-life has a negative coefficient, but a relatively large and significant standard deviation. As explained before, the expected coefficient on half-life is ambiguous as some people who experience severe adverse side effects prefer a fast reaction to the medication whereas others, who consider taking medication often a hassle, prefer a longer half-life. The negative coefficient shows that the severe side-effects effect won over the hassle effect. This means that, by allowing variability in patient preferences, patients who experience severe adverse effects and prefer shorter half-lives are more in the randomly chosen sample. The large standard deviation implies that the coefficient is positive for many patients, that is the sample includes those patients that consider small half-lives a hassle as they have to keep remembering to frequently retake their medication. The negative and statistically significant coefficient of age, as explained before, is due to the fact that patients prefer newer more innovative drugs to old antidepressants. The small standard deviation, though statistically insignificant, is economically significant and shows that for all patients age provides a disutility.

The parameters of most importance in this final model are the coefficients on the preference of brandness and price sensitivity. As presented in equation (2), given the assumption on the independence of the distributions of unobserved and observed patient characteristics (that is, τ_i and D_i are independent), the total price sensitivity is a combination of the mean effects and the effects prescribed by the interaction with unobserved and observed characteristics. The mean effect on price is now below unity, -0.872. This is the disutility obtained by the average patient. The

relatively small estimated standard deviation suggests that most of the heterogeneity (87%) in patient preferences is explained by the included demographics. In other words, the inclusion of these observed demographics improved the model's predictability. Estimates imply that wealthier patients and patients with prescription drug insurance tend to be less price sensitive. In fact, when deriving the combined effect of income and insurance on the mean price disutility the total marginal valuation of price comes closer to zero. When taking the standard deviation from the unobserved patient characteristics into account as well, one concludes that many patients have price sensitivities not far from zero.[31] This result uncovers the moral hazard problem that arises due to the presence of prescription drug insurance coverage. The question then becomes one of distinguishing between a patient's private marginal willingness-to-pay or the marginal social willingness-to-pay when estimating welfare.

INSERT TABLE IX ABOUT HERE

The estimates of the coefficients on the preference for brandness (without generic competition) are all positive as expected. The mean effect is a high positive value and says that the average patient prefers branded drugs over generics. The marginal valuation of brand preference increases with income and insurance. This means wealthier patients and patients with prescription insurance coverage get even more utility from consuming branded drugs over generics. This reinforces the result of associating brandness to quality. In other words, the coefficient on preference for branded drugs is a proxy for patient-perceived drug quality. Again, the relatively small estimated standard deviation suggests that most of the heterogeneity (86%) in patient preference for this brandness is explained by the observed demographics.

The estimates of the coefficients on the preference for brandness (with generic competition) are as expected. The mean effect is large, negative and significant. The standard deviation is tiny suggesting that almost all of the heterogeneity (98%) in patient preference for this brandness is explained by the observed demographics. More importantly, the marginal valuation of this brand preference increases with income and decreases with insurance. This means wealthier patients get less disutility from consuming branded drugs over generics when the two coexist in the same molecule. This reinforces the result of associating brandness to quality. In other words, the coefficient on

[31] Recall that I restricted the model not to allow for positive price sensitivity, even in the tail of the distribution. This became necessary when making demand-based patient welfare assessments.

preference for branded drugs is a proxy for patient-perceived drug quality and when income is high the patient is indifferent and opts for the marginally higher quality of the branded drug. On the other hand patients with prescription insurance coverage get even more disutility from consuming branded drugs over generics when the two coexist in the same molecule. Though counterintuitive at first glance, this finding may be related to the fact that formularies in most insurance schemes cover only the generic, or at most the generic cost, when that exists. Hence, individuals with insurance are likely to be confined in their choices by formularies.

INSERT TABLE X ABOUT HERE

Tables IX and X show the combined effect of demographics on price and brand sensitivity, respectively, and compare results for the full information (FIM) and limited information (LIM) models. Wealthier patients who also have full prescription coverage are almost insensitive to changes in the price of a drug and have a higher preference for branded drugs (with or without insurance). Poorer patients without prescription drug coverage are the most sensitive to price changes and have the lowest preference for brandness (without generic competition) and the highest disutility for brandness (with generic competition). Moving from the latter extreme case to the former, the off diagonal results show that poorer but insured patients are less price sensitive, have a higher preference for brandness (without generic competition) and lower disutility for brandness (with generic competition) than do wealthier, uninsured patients. When comparing the two models together LIM obtains lower price disutilities in all demographic combinations. In the case of brandness (without generic competition), LIM obtains lower marginal valuations. Finally, in the case of brandness (with generic competition) LIM obtains lower disutilities for the uninsured and turns the disutilities into positive utilities in the case of the insured.

Traditional FIM capture all differences in information through the ξ_j or the ε_{ij}, both of which are independent across patients. Information heterogeneity indirectly captured by the i.i.d. error will be restricted such that each patient-drug pair has its own realization that is independent of patient and drug attributes (such as advertising) and of all other patient-drug pairs. This does not permit correlation in information across patients nor does it permit informational advantages to depend on patient and drug observables. Alternatively, information heterogeneity can be indirectly captured via ξ_j. In LIM, a product with little advertising is unlikely to be in many patients' choice sets and

will have a low market share. In FIM, a small market share could be explained by a low value for ξ_j. Again, the unobserved term is independent of patient characteristics. Not explicitly allowing for informational asymmetries is particularly restrictive in rapidly changing, complex markets where patients are likely to have limited information, and hence where heterogeneity in the distribution of information across patients and drugs explains (perhaps a significant) part of the variation in sales across products. The results indicate that relying on ξ_j or ε_{ij}, to explain differences in information across patient-drug pairs can generate inconsistent estimates of drug–specific demand curves that are biased toward being too elastic.

iv Consumer Information Heterogeneity and Advertising Effectiveness

As expected, results indicate that advertising has very different effects across patients and that exposure to advertising significantly impacts the information set. Table XI presents estimates of how media exposure varies with observed demographic characteristics (Υ). These coefficients proxy for effectiveness of advertisements in reaching consumers through various media. The results indicate print media (magazines and newspapers) are most effective at reaching high income, insured, married individuals who are above the age of 34, but are less likely to reach low-educated white males. It is the most effective medium to reach high-income, fully-insured individuals who also have the lowest price elasticities. TV advertising is the most effective medium for reaching low income households and is also effective at reaching married, less-educated individuals over 54, although not as effective as print media. Most DTC pharmaceutical advertising is on television (63.6% in 2008) suggesting pharmaceutical firms target low-income, insured, older individuals, possibly people on medicare and medicaid. Internet advertising is the most effective medium for reaching high-income individuals and it has recently been picking up. Perhaps pharmaceutical firms are trying to catch up and approach the wealthier, price-insensitive population, who are also very likely at being insured.

INSERT TABLE XI ABOUT HERE

The results confirm that variation in advertising media exposure across households is an important source of consumer heterogeneity. The variation in exposure translates into variation in information sets as evidenced by the positive and highly significant estimate for ζ. The estimates highlight the importance of considering the differential effects of advertising both across households

and across media. Sovinsky Goeree (2008) was the first to do so in a structural empirical industrial organization setting. Parameter estimates of λ suggest other means of information provision, such as word-of-mouth or experience, play a role in informing certain types of consumers. The coefficient on income less than \$25,000 (0.59) indicates these individuals are likely to be informed about 41% of the products without seeing an advertisement, coefficient on income between \$25,000-\$49,000 (0.43) indicates these individuals are likely to be informed about 57% of the products without seeing an advertisement, whereas having a high income ($>$ \$100,000) is not significantly different from having a middle income in terms of being informed without seeing an ad. Possibly, lower-income individuals have lower opportunity costs and thus more time to search for information. the difference between the two low-income groups is explained by the fact that individuals with income less than \$25,000 may be on medicaid and have no incentive to search.

In addition, the probability of being informed without seeing any advertising the fully insured is low (24%). Possibly these individuals are indifferent in knowing which drug is the best for them since they are covered whatever the cost by insurance. This finding reinforces the existence of moral hazard and in fact shows that with limited information the moral hazard issue is more salient. Finally, in the bottom of Table XI there are estimates of the parameters that are the same across individuals (the γ_j parameters). Patients are significantly more likely to know a drug the longer it has been on the market (0.14), with stong diminishing returns (-0.08). This is intuitive, for the longer it has been on the market, the more opportunity consumers have had to learn of it by word-of-mouth or through advertising and diminishing returns are due to forgetting. There are also decreasing returns to advertising in print (-0.42), television (-0.62) and internet (-0.07) media, but they are decreasing at a faster rate for television.

v Substitution Patterns and Information Provision

Table XII presents own- and cross-price elasticities of demand for selected antidepressants. These are weighted averages for the years 2000 and 2001. The reported cross-price elasticities are averaged over drugs of the same type. The selection includes one NewGen, all SSRIs, NDRIs and SNRIs. The selected antidepressants help show the superiority of the RCL model over the other estimated models in describing substitution patterns and, moreover, help distinguish the differences between the FIM and LIM models. Drugs with similar characteristics have larger substitution patterns,

ceteris paribus. Drugs within the same molecule that are both branded should essentially have almost identical cross-price elasticities. For instance, Prozac, Prozac Weekly and Sarafem have very close cross-price elasticities of demand with respect to other drugs. Similarly Wellbutrin and Wellbutrin SR. Note that drugs of the same molecule but not both branded have similar relative cross-price elasticities but the elasticities are not similar in magnitude. The estimated strong preference for brandness exacerbated by the inclusion of demographics accounts for the difference. The table can be used to show, for example, that SSRIs tend to be closer substitutes to other SSRIs, less so to NDRIs and MAOIs and much less to TCAs. Moreover, the table shows that the only generic in SSRIs has high cross-price elasticities with all other SSRIs. This is explained by the fact that patients are more willing to substitute towards a me-too drug when that is much cheaper.

INSERT TABLE XII ABOUT HERE

Advertising elasticities work in the opposite fashion. Own advertising elasticities are positive, that is. more advetising raised own market share; cross-advertising are negative, that is, more advertising reduces market shares for other brands, more so of brands that are in close competition.[32] Therefore, we observe highest cross-advertising elasticities with drugs of the same molecule, followed by drugs of the same type. As Sovinsky Goeree (2008) has shown, substitution among drugs could be induced by changes in choice sets, which is significantly impacted by advertising with varying effects across consumers. When advertising changes, the impact on the choice set is more pronounced for those consumers who are more sensitive to advertising. The firms' decisions of what prices to charge and how much information to provide through advertising depend on these price and advertising elasticities of demand.

Finally, price elasticities were estimated for both the FIM and LIM models. The extent to which a firm can exercise market power depends on the elasticity of its drug's demand curve. The greater the number of competitors or the larger the cross-elasticity of demand with other drugs, the greater the elasticity of the firm's demand curve and the less its market power. The major difference in the FIM and LIM models is that the former predicts higher elasticities implying that the antidepressant market is quite competitive. Without the full information assumption, however, the industry is less competitive and products are less substitutable than in FIM. This arises because

[32] Note that generics are not advertised so no elasticities can be computed.

less of information by consumers implies smaller choice sets, and consequently less subsitutability and less competitiveness.

VI WELFARE IMPLICATIONS

The underlying assumption for a demand-based assessment of patient WTP is that consumer surplus can be measured by the revealed preferences of consumers through their observed choices.[33] In a market for a single, homogeneous drug, only patients that value the drug above its price purchase the drug. Patient surplus, therefore, is the area between the demand curve and the price and incremental welfare from product innovation is the before-and-after the innovation difference in the area under the demand curve.

Patient welfare associated with each drug, conditional on the prices and characteristics of available substitutes is equivalent to:

$$W_{jt} = \int_{i=1}^{ns} -\frac{1}{\alpha_i} \int_{p_j}^{\infty} s_{ijt}\left(q_j | q_k = p_k \forall k < j, q_k = \infty \forall k > j\right) dq_j dF(\alpha_i, \sigma_\alpha) \tag{8}$$

where each s_{ijt} is computed using the estimated parameters as in equation (4) summing over the estimated distribution of varying patient price sensitivities, $F(\alpha_i, \sigma_\alpha)$. Dividing the computed patient welfare by the price sensitivity in equation (8) gives the monetary amount a patient would be willing-to-pay to be faced with a choice set J_t prior to observing the realization of her idiosyncratic utility.

i Moral Hazard

Patients insured against prescription drug expenditures are willing to pay higher prices for their medications than they would be willing to pay when uninsured. This is reflected in the very low estimated marginal disutility of price that results from the presence of prescription drug coverage. The moral hazard problem appears more severe when we relax the full information assumption, as limited information has led to more inelastic demand curves. To address the moral hazard issue, welfare should be estimated both when patients are insured and when patients are uninsured against prescription costs. The former estimate reflects the social willingness-to-pay, the latter the

[33] Similar analyses in Trajtenberg (1990), Ellickson, Stern and Trajtenberg (2001), Cleanthous (2003).

private willingness-to-pay. The difference in the two is attributed to moral hazard.

INSERT TABLE XIII ABOUT HERE

Table XIII presents the welfare estimates for patients without insurance in 1980 dollars for selected antidepressants for both the full and limited information models. This is the 'true' patient's willingness-to-pay over the price charged, the moral hazard, that arises due to the inclusion of prescription insurance in the computation of welfare gains. The flexibility of the model allows me to remove the simulated individuals that have prescription insurance from the estimation of welfare gains. I, therefore, recalculate the gains that exclude prescription drug insurance and report them in the total welfare column. With the exclusion of insurance, estimated patient gains are insightful. Finding surplus per unit (average daily dosage) in the next column, shows a patient's willingness-to-pay above the price of the drug. The last column of Table XIII shows the excess willingness-to-pay accrued annually averaged over the data. In other words, an individual patient would be willing to pay \$8,929 in a year over the amount already spent to be able to use Prozac in FIM, compared to \$9,821 in LIM. Comparing this to the average annual cost of depression of an individual patient[34], \$3,351, a patient would be willing to pay 3.7 times more a year for a Prozac treatment under FIM and 3.9 times more a year under LIM.

As expected, when limiting consumer information, the estimated inelastic demands for drugs produce larger welfare estimates than in the full information model, and consequently, a steeper moral hazard problem. As pharmaceutical firms use informative advertising to extract monopoly power, they exacerbate the moral hazard that arises due to prescription drug insurance. The use of the limited information model as introduced by Sovinsky Goeree (2008) has aided in calculating the effect of informative advertising on moral hazard.

In general, demand estimation has shown that insured patients tend to be less price sensitive than uninsured patients, and more so when we take into account that patients are not informed about all the drugs available to them. This leads to a moral hazard, which I estimate by evaluating the willigness-to-pay for treatment had there been no insurance. These results on welfare gains and patient willingness-to-pay are useful to pharmaceutical companies to target drug characteristics in new innovations, to extract the extra willingness- to-pay through pricing, to advertise to the

[34] Annual deflated average of Greenberg et al (1993) and Badamgarav et al (2003).

right demograhic group and through the correct medium. Public policy can also use these welfare estimates. For instance, comparing these results to research and development costs provide cost-benefit analysis of new drug introduction. Moreover, governments can evaluate the fairness of pharmaceutical pricing practices and the effect of direct-to-consumer advertising and direct-to-physician advertising on consumer welfare.

VII CONCLUSION

Pharmaceutical markets are characterized by a high degree of innovation, complexity and uncertainty, especially markets of idiosyncratic symptomatolgy and response to treatment such as the antidepressant market. In this paper, I show that it is unreasonable to assume that consumers are aware of all antidepressants for sale at the time of purchase, as is the case in traditional models of consumer choice. Such an assumption biases demand curves towards being more elastic and, in consequence, the full information assumption biases the evaluation of consumer welfare downwards. This paper, therefore, aims at analyzing and evaluating the effects of promotions by pharmaceutical firms on patient welfare taking into account the interaction of multiple agents (patients, physicians, insurance companies and pharmaceutical companies) in the decision process.

I formulate an empirical methodology that incorporates both macro- and micro-level data in the U.S. antidepressant market and takes into account the multi-agent interaction to estimate demand and welfare. I use an empirical discrete-choice model of limited information, where advertising influences the set of drugs from which a purchase choice is made. The paper employs an original dataset that consists of annual observations on prices, quantities, direct-to-physician and direct-to-consumer advertising, media exposure by demographic and drug characteristics for all antidepressants sold in the U.S. market from 1980 to 2001 and demographic data on the distribution of patient income and prescription insurance.

Estimation results indicate that pharmaceutical firms use advertising media to target high-income households and households with more comprehensive prescription drug insurance schemes through their physicians or directly. Comparison of the full and limited information models shows that limited information leads to less elastic demand curves and more reasonable substitution patterns. As a result, I derive larger estimates of patient welfare due to pharmaceutical innovation that exacerbate the moral hazard issue that arises due to the existence of prescription drug insurance

coverage. Patients are willing-to-pay more for a drug when they are covered by prescription drug insurance and even higher when their choice set is smaller due to limited information.

These results imply that the pharmaceutical industry is not as competitive as a full information model would predict. Firms use informative advertising to maintain market power and monopoly profits for their product. The paper estimates large and precise patient welfare gains due to innovation and explains the detected divergence between social and private patient benefits by the presence of insurance in the full information model and show how this is much bigger with limited informations. These findings aid in public policy decision making on health care and pharmaceutical industry concerns.

Demand estimates correctly detect marginal disutilities for drug side effects and estimated drug substitution patterns accurately reflect differences in patient tastes for drug attributes. I find a large mean price disutility, which varies with income and insurance demographics. The estimated price sensitivity decreases with patient income, when patients are insured against prescription drug expenditures and with limited information. Moreover, patients demonstrate a high preference for branded drugs when these do not face generic competition and a disutility to brandness when generic competition exists. The wealthier the patients and the more insurance coverage they have, the higher the preference and, therefore, the more they will be targetted by pharmaceutical firms.

VIII APPENDIX

The simple logit model can be estimated with and without instruments using the following equation:

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt}. \tag{9}$$

Three nested multinomial logit models are also estimated with and without instruments. The estimation equations for the three models (NML1a, NML1b, NML2) are, respectively:

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt} + \rho_c \ln(\widetilde{s}_{j/c,t}) \tag{10}$$

where the only naturally-occurring nested group is assumed to be the 'type' of the antidepressant and ρ_c is the within-type correlation coefficient of utility;

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt} + \rho_c \ln(\widetilde{s}_{j/m_c,t}) \tag{11}$$

where the only naturally-occurring nested group is assumed to be the 'molecule' of the antidepressant and ρ_m is the within-molecule correlation coefficient of utility;

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt} + \rho_{m_c} \ln(\widetilde{s}_{m_c/c,t}) + \gamma \ln(\widetilde{s}_{j/m_c,t}) \tag{12}$$

where the two naturally-occurring nested groups are assumed to be the 'type' and 'molecule' of the antidepressant and ρ_c, ρ_m are the within-type and within-molecule correlation coefficients of utility.

In these nested logit models, ρ_c, ρ_{m_c}, and γ are correlation coefficients and have to be between zero and one. In NML1a ρ_c is the within-type utility correlation parameter and is the coefficient of the market share of drug j within type c. In NML1b, ρ_{m_c} is the within-molecule utility correlation parameter and is the coefficient of the market share of drug j within type m_c. In NML2, ρ_c is obtained by inserting the estimates for ρ_{m_c} and γ in $\gamma = [1 - (1 - \rho_{m_c})(1 - \rho_c)]$. The closer these correlation coefficients are to one the more valid the assumption that the nesting groups are naturally-occurring.

The low and sometimes insignificant correlation coefficients in the nested multinomial logit models are reason to favor the full random coefficient model which places no restriction on the correlation between antidepressants.

INSERT TABLES A1 AND A2 ABOUT HERE

IX REFERENCES

Ackerberg, Daniel, 2001. "Empirically Distinguishing Informative and Prestige Effects of Advertising," *RAND Journal of Economics* 32(2):100-118.

Ackerberg, Daniel, 2003. "Advertising, Learning, and Consumer Choice in Experience GoodsMarkets: A Structural Empirical Examination," *International Economic Review* 44:1007-1040.

American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*, Fourth Edition, Text Revision. American Psychiatric Association, Washington DC.

Anand, Bharat and Ron Shachar, 2010. "Advertising, the Matchmaker," Rand Journal of Economics, *Rand Journal of Economics*, Forthcoming.

Approved Drug Products with Therapeutic Equivalence Evaluations. (*The Orange Book*), 29^{th} Edition, 2009. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Office of Information Technology, Division of Data Management and Services.

Badamgarav, Enkhe, Scott R. Weingarten, James M. Henning, Kevin Knight, Vic Hasselblad, Anacleto Gano, Jr., and Joshua J. Ofman. 2003. "Effectiveness of Disease Management Programs in Depression: A Systematic Review," American Journal of Psychiatry 160: 2080-2090.

Berndt, Ernst R., L. T. Bui, D. H. Lucking-Reiley, and G. L. Urban, 1997. "The Roles of Marketing, Product Quality and Price Competition in the Growth and Composition of the U.S. Anti-ulcer Drug Industry." In Bresnahan, Timothy F. and Robert J. Gordon (eds) 1997, *The Economics of New Goods*, University of Chicago Press, Chicago.

Berry, Steven T., 1994. "Estimating discrete-choice models of product differentiation," *Rand Journal of Economics*, 25(2): 242-261.

Berry, Steven T., James Levinsohn and Ariel Pakes, 1995. "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4): 841-890.

Berry, Steven T., James Levinsohn and Ariel Pakes, 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112(1), 68–105.

Cleanthous, Paris, 2003. "Pharmaceutical Demand and Welfare Implications of Innovation," Ph.D. Yale University.

Cleanthous, Paris, 2009. "Evaluating Innovation and Moral Hazard in Pharmaceuticals," mimeo University of Cyprus.

Depression Guideline Panel 1993. Depression in Primary Care: Volume 2. Treatment of Major Depression. Clinical Practice Guideline, Number 5. Rockville, MD. U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. AHCPR Publication No. 93-0551. April.

Ellickson, Paul, Scott Stern and Manuel Trajtenberg, 1999. "Patient Welfare and Patient Compliance: An Empirical Framework for Measuring the Benefits from Pharmaceutical Innovation," in Ernst Berndt and David Cutler (eds.), 2001. *Medical Care Output and Productivity, Studies in Income and Wealth,* Vol. 62: 539-60. University of Chicago Press.

Erdem, Tulin and Michael Keane, 1996. "Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," Marketing Science 15(1): 1-20.

Greenberg P. E., Stiglin L. E., Finkelstein S. N., Ernst R. Berndt, 1993. "The economic burden of depression in 1990," Journal of Clinical Psychiatry, 54: 405–418.

Hausman, Jerry and Daniel McFadden, 1984. "Specification Tests for the Multinomial Logit Model," *Econometrica,* 52(5): 1219-1240.

Lacy, Charles F., Lora L. Armstrong and Morton P. Goldman, 2009. *Drug Information Handbook,* 18th Edition. Web Publication, Lexi-Comp.

Milgrom, Paul and John Roberts, 1986. "Price and Advertising Signals of Product Quality," Journal of Political Economy 94(4): 796-821.

Miranda et al., 1994. *Mental Disorders in Primary Care*, Jossey-Bass Publications.

National Comorbidity Survey, 2009. [http://www.hcp.med.harvard.edu/ncs]

Nevo, Aviv, 2000. "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 9(4): 513-48

Petrin, Amil, 2002. "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110(4): 705-29.

Physician's Desk Reference, 2002. Montvale, NJ. Medical Economics Data.

Salmans, Sandra, 1997. "More on Treatments." Depression: Questions You Have. Answers You Need: 145. People's Medical Society Publications.

Scherer, F. M., 2010. "Pharmaceutical Innovation," in Bronwyn Hall and Nathan Rosenberg (eds.), 2010. *Handbook of the Economics of Innovation*, Vol. 1, North Holland.

Shum, Matthew, 2004. "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast Cereals Market," Journal of Economics and Management Strategy 13: 241-272.

Sovinsky Goeree, Michelle, 2008. "Limited Information and Advertising in the U.S. Personal Computer Industry," *Econometrica*, 76(5): 1017-1074.

The Lewin Group, 2000. *Access and Utilization of New Antidepressant and Antipsychotic Medications*, U.S. Department of Health and Human Services.

Trajtenberg, Manuel, 1989. "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *Journal of Political Economy*, 97(2): 444-79.

TABLE I

PHARMACEUTICAL INDUSTRY AND ANTIDEPRESSANT MARKET SUMMARY STATISTICS

		Annual	1980		1990		1995		2000		2005		2008
		Average	Total	AD	Total	AD	Total	AD	Total	AD	Total	AD	Total
Sales	(Billion $)	434	32.1	0.13	102	0.96	163	3.04	229	10.1	654	12.5	928
Advertising-to-Sales Ratio	(%)	5.3	n.a.	15	n.a.	13	5.1	18	6.9	9.7	4.6	8.1	4.2
Advertising-to-Margin Ratio	(%)	7.3	n.a.	n.a.	n.a.	n.a.	7.1	n.a.	9.5	n.a.	6.0	n.a.	5.5
Advertising Expenses	(Billion $)	21.6	n.a.	0.02	n.a.	0.12	8.39	0.55	15.8	0.98	30.1	1.02	39.0
Direct-to-Consumer Advertising	(%)	13	n.a.	-	n.a.	-	4.5	-	16	12	14	12	11
Professional Journal Advertising	(%)	2.1	n.a.	22	n.a.	6.7	4.3	3.3	3.2	3.2	1.3	4	1.0
Sales Representative Details	(%)	25	n.a.	42	n.a.	47	35	45	30	34	24	34	17
Retail Value of Samples	(%)	60	n.a.	36	n.a.	46	57	52	51	51	61	49	71
Research & Development	(Billion $)	26	1.5	n.a.	6.8	n.a.	12	n.a.	21	n.a.	31	n.a.	40

Notes: Data are for the pharmaceutical preparations industry (SIC 2834). Note that numbers may differ for broader categorizations as in Table II, for example. Data for sales, advertising expenses and their ratios come from Schonfeld & Associates via AdAge.com; data for advertising expenses (exclude promotion spending for professional meetings and events) and their breakout come from IMS Health, Inc., Integrated Promotional Service, and Competitive Media Reporting; data for R&D come from PhRMA's annual industry review. 'Total' column refers to the whole pharmaceutical industry and 'AD' to the antidepressant market. Advertising-to-Sales Ratio=Advertising/Net Sales; Advertising-to-Margin Ratio=Advertising Expense/(Net Sales-Cost of Goods Sold). Direct-to-consumer advertising includes advertising for prescription drugs on TV, radio, magazines and newspapers, as well as internet and outdoor advertising; professional journal advertising reflects advertising expenditures for prescription drugs appearing in medical journals; sales representative details include costs associated with the sales activities of pharmaceutical representatives that are directed to office-based physicians, hospital-based physicians, and directors of pharmacies; samples are prescription drugs given to physicians to disseminate freely to patients.

TABLE II

MEDIA BREAKOUT IN THE U.S. PHARMACEUTICAL INDUSTRY

		1995	2000	2005	2008
Total Media Expenditure	(Billion $)	2.8	5.1	8.4	8.7
Magazine	(%)	25.1	26.0	27.3	28.4
Newspaper	(%)	4.5	3.4	2.6	2.5
Television	(%)	57.4	52.7	45.4	46.7
Cable Networks	(%)	8.9	15.3	16.4	16.9
Internet	(%)	-	-	4.6	3.2
Outdoor	(%)	0.7	0.1	0.2	0.3
Radio	(%)	3.6	2.6	3.4	2.0

Notes: Data are for the 'medicines & proprietary remedies' category (aggregated from TNS classifications) by LNA, which include pharmaceutical houses, medicines & proprietary remedies, fitness, eye glasses, medical equipment. This is a broader categorization than the SIC categorization in Table I, hence, the discrepancy in total DTC spending. Percentages calculated from measured media. Magazines include Sunday, local, and Spanish-language magazine; newspaper includes national and Spanish-language newspaper; television includes network, national spot and local.

TABLE III

MEDIA BREAKOUT OF DIRECT-TO-CONSUMER ADVERTISING FOR SELECTED PHARMACEUTICAL FIRMS

	Total Ad Spending	Print	TV & Radio	Internet & Outdoor	Total Ad Spending	Print	TV & Radio	Internet & Outdoor	Total Ad Spending	Print	TV	Internet, Outdoor, & Radio
	1995				2000				2005			
	(Billion $)	(%)	(%)	(%)	(Billion $)	(%)	(%)	(%)	(Billion $)	(%)	(%)	(%)
Top 10 Firms	4.47	21.6	78.4	0.02	9.90	23.3	76.3	0.4	12.32	29.6	65.5	4.9
Abbott Laboratories	0.13	17.9	82.1	0.00	0.28	12.7	87.5	0.0	0.41	44.6	53.9	1.5
AstraZeneca	n.a.	n.a.	n.a.	n.a.	0.20	38.4	61.5	0.1	0.80	40.6	50.3	9.1
Bayer AG	0.39	17.2	82.7	0.09	0.65	7.7	92.1	0.2	0.57	14.8	79.8	5.4
Bristol-Myers Squibb	0.44	22.9	77.1	0.05	1.19	33.2	66.5	0.3	0.58	57.8	37.3	4.9
Eli Lilly & Co.	n.a.	n.a.	n.a.	n.a.	0.10	29.9	63.8	6.4	0.48	14.6	80.2	5.2
GlaxoSmithKline	0.44	40.6	58.9	0.45	1.13	19.2	79.7	1.2	2.19	26.0	70.8	3.2
Johnson & Johnson	0.80	17.3	82.7	0.00	1.60	20.1	79.9	0.1	2.21	32.7	63.1	4.3
Merck & Co.	0.07	82.5	17.5	0.00	0.98	41.1	58.9	0.0	0.77	33.6	61.9	4.6
Novartis	0.36	19.0	81.0	0.00	0.57	19.5	80.5	0.0	1.16	18.5	74.7	6.7
Pfizer Inc.	0.16	26.9	73.1	0.00	2.27	24.4	74.8	0.8	2.15	33.6	61.3	5.1
Schering-Plough	0.20	48.7	51.3	0.00	0.51	22.7	76.6	0.7	0.85	14.5	82.3	3.2

Notes: Firms are presented alphabetically and they must have been included in the top 10 firms in at least one of the three reported years. Data are for the 'medicines & proprietary remedies' category (aggregated from TNS classifications) by LNA, which include pharmaceutical houses, medicines & proprietary remedies, fitness, eye glasses, medical equipment. Top 10 firms are according to each year. Print includes Sunday, local, and Spanish-language magazine and national and Spanish-language newspaper; TV includes network, national spot and local. Astrazeneca and Eli Lilly's advertising data are not available for 1995.

TABLE IV
CHOICE IN THE ANTIDEPRESSANT MARKET

Type	Molecule	Drug	Name	Generic Name
1	1	1	Marplan	Isocarboxazid
1	2	2	Nardil	Phenelzine
1	3	3	Parnate	Tranylcypromine
2	4	4	Elavil	Amitriptyline
2	4	5	Endep	Amitriptyline
2	4	6	Generic	Amitriptyline
2	5	7	Asendin	Amoxapine
2	5	8	Generic	Amoxapine
2	6	9	Anafranil	Clomipramine
2	6	10	Generic	Clomipramine
2	7	11	Generic	Desipramine
2	7	12	Norpramin	Desipramine
2	7	13	Pertofrane	Desipramine
2	8	14	Adapin	Doxepin
2	8	15	Generic	Doxepin
2	8	16	Sinequan	Doxepin
2	9	17	Generic	Imipramine
2	9	18	Janimine	Imipramine
2	9	19	Tofranil	Imipramine
2	10	20	Tofranil PM	Imipramine Pamoate
2	11	21	Generic	Maprotiline
2	11	22	Ludiomil	Maprotiline
2	12	23	Aventyl	Nortriptyline
2	12	24	Generic	Nortriptyline
2	12	25	Pamelor	Nortriptyline
2	13	26	Generic	Protriptyline
2	13	27	Vivactil	Protriptyline
2	14	28	Generic	Trimipramine
2	14	29	Surmontil	Trimipramine
3	15	30	Serzone	Nefazodone
3	16	31	Desyrel	Trazodone
3	16	32	Generic	Trazodone
4	17	33	Celexa	Citalopram
4	18	34	Prozac	Fluoxetine
4	18	35	Prozac Weekly	Fluoxetine
4	18	36	Sarafem	Fluoxetine
4	19	37	Generic	Fluvoxamine
4	19	38	Luvox	Fluvoxamine
4	20	39	Paxil	Paroxetine
4	21	40	Zoloft	Sertraline
5	22	41	Generic	Bupropion
5	22	42	Wellbutrin	Bupropion
5	22	43	Wellbutrin SR	Bupropion
6	23	44	Effexor	Venlafaxine
6	23	45	Effexor-XR	Venlafaxine
7	24	46	Remeron	Mirtazapine
7	24	47	Remeron Soltab	Mirtazapine

Notes: Types 1-7 stand for MAOI, TCA, NewGen, SSRI, NDRIs, SNRI and NaSSA respectively.

TABLE V

ENTRY AND AVERAGE CHARACTERISTICS BY TYPE OF ANTIDEPRESSANTS

Drug Type	No. Molecules	No. Brands	No. Generics	Molecuted Entry	Branded Entry	Generic Entry	Year of First Entry in the Market	Average Annual Revenue Shares (%)			Average Annual Quantity Shares (%)				Average Dosing Frequency	Half-Life	Side Effects							
								1980	1990	2001	1980	1990	2001		FREQ	HALF	Fatality FAT	Anti-Cholinergic AC	Drowsiness DR	Insomnia/ Agitation IA	Orthostatic Hypotension OH	Cardiac Arrhythmias CA	GI Distress GID	Weight Gain WTG
(1) MAOI	3	3	0	0	0	0	1959	1.51	1.74	1.07	1.51	1.76	1.09	Mean	1.7	19.3	4.00	0.00	1.67	2.00	4.00	0.00	1.00	1.67
														StdDev	*0.0*	*8.1*	*0.00*	*0.00*	*0.58*	*0.00*	*0.00*	*0.00*	*0.00*	*0.58*
(2) TCA	11	16	10	2	2	8	1959	98.5	76.5	44.1	98.5	87.1	59.6	Mean	1.2	26.2	1.83	2.57	2.78	0.70	2.78	2.43	0.43	2.13
														StdDev	*0.4*	*17.7*	*0.39*	*1.04*	*1.13*	*0.63*	*1.00*	*0.51*	*0.90*	*1.22*
(3) NewGen	2	2	1	2	2	1	1982	-	13.6	8.61	-	5.71	5.37	Mean	2.7	22.3	1.00	2.00	4.00	0.33	3.00	0.67	3.00	0.00
														StdDev	*0.0*	*16.7*	*0.00*	*1.73*	*0.00*	*0.58*	*1.73*	*0.58*	*0.00*	*0.00*
(4) SSRI	5	7	1	5	7	1	1988	-	38.4	65.4	-	23.2	48.4	Mean	1.0	71.6	1.00	0.50	0.75	2.25	0.25	0.25	3.00	0.13
														StdDev	*0.0*	*80.4*	*0.00*	*0.93*	*1.39*	*0.46*	*0.46*	*0.46*	*0.00*	*0.35*
(5) NDRI	1	2	1	1	2	1	1989	-	-	5.96	-	-	4.18	Mean	2.5	15.0	1.00	0.00	0.00	2.00	0.00	0.00	3.00	0.00
														StdDev	*0.0*	*0.0*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
(6) SNRI	1	2	0	1	2	0	1994	-	0.95	4.51	-	0.83	3.27	Mean	3.0	16.7	0.50	0.00	0.00	2.00	0.00	1.00	1.00	0.00
														StdDev	*0.0*	*12.7*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
(7) NaSSA	1	2	0	1	2	0	1996	-	-	1.95	-	-	0.57	Mean	1.0	2.0	1.00	1.00	3.00	0.00	0.00	0.00	0.00	3.00
														StdDev	*0.0*	*0.0*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Antidepressants (All Drugs)	24	34	13	12	17	11	1959							Mean	1.5	31.5	1.60	1.61	2.11	1.16	1.98	1.43	1.25	1.39
														StdDev	*0.8*	*40.4*	*0.85*	*1.43*	*1.54*	*0.91*	*1.61*	*1.17*	*1.33*	*1.37*

Notes: Data come from IMS Health, Food & Drug Administration, Depression Guideline Panel (1993), Physician's Desk Reference Generics (2009), Drug Information Handbook (2009).

TABLE VI
CONSUMER CHARACTERISTICS

	Sample		Population	
Variable Description	Mean	Std. Dev.	Mean	Std. Dev.
Male	0.489	0.002	0.488	0.002
White	0.835	0.017	0.835	0.017
Age (years)	35.32	2.88	34.94	2.77
< 35	0.528	0.036	0.535	0.036
35-54	0.108	0.015	0.109	0.015
> 54	0.328	0.014	0.332	0.014
Education (years)	13.53	2.64	13.55	2.45
Married	0.566	0.015	0.574	0.015
Household size (persons)	2.662	0.046	2.660	0.046
Employed	0.598	0.026	0.597	0.026
Income ($)	58,629	6,126	58,211	6,020
Under $25,000	0.273	0.020	0.274	0.020
$25,000 to $49,999	0.272	0.013	0.273	0.013
$50,000 to $99,999	0.314	0.015	0.312	0.015
$100,000 and over	0.142	0.032	0.141	0.032
Antidepressant Prescription	0.020	0.016	0.020	0.016
Prescription Drug Expenditures:				
Public	0.196	0.019	0.196	0.019
Private	0.344	0.075	0.344	0.075
Out-of-Pocket	0.460	0.093	0.460	0.093

Media Exposure	Mean	Std. Dev.
Television (hours per person per year)	1,586	42.85
Radio (hours per person per year)	964	15.31
Print (hours per person per year)	360	11.41
Internet (hours per person per year)	80	53.73

Notes : Variables are dummies unless units are specified. Data come from Nielsen Media Research, IMS Health, National Center for Health Statistics, Bureau of Labor Statistics: Consumer Population Survey, Veronis Suhler Stevenson: Communications Industry Forecast.

TABLE VII
DEMAND ESTIMATION WITH FULL AND LIMITED INFORMATION

	OLS	IV	Random Coefficients			
	NML2	NML2	(no demographics)		(demographics)	
	(limited info.)	(limited info.)	(full info.)	(limited info.)	(full info.)	(limited info.)
	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Price Coefficients (α's)					
Price	-0.155**	-2.709***	-1.795*	-0.996**	-1.108***	-0.872***
	[0.071]	[0.124]	[1.045]	[0.424]	[0.025]	[0.042]
	Coefficients of Characteristics (β's)					
Half-life	0.006**	0.004	0.005	-0.350***	-0.673*	-0.355***
	[0.002]	[0.002]	[0.007]	[0.053]	[0.505]	[0.047]
Fatal	-0.536***	-0.731***	-2.184***	-2.076***	-2.665**	-0.508**
	[0.132]	[0.142]	[0.451]	[0.368]	[1.064]	[0.172]
Weight Gain	-0.039	-0.213**	-0.303*	-0.294**	-0.453***	-0.322***
	[0.086]	[0.095]	[0.162]	[0.125]	[0.124]	[0.113]
Insomnia/ Agitation	0.492***	0.419***	-0.825	-0.770	-0.777***	-0.340*
	[0.099]	[0.105]	[0.571]	[0.512]	[0.137]	[0.205]
Side Effects	-1.975***	-1.299**	-4.148**	-0.756**	-1.140**	-1.076**
	[0.574]	[0.613]	[1.697]	[0.381]	[0.545]	[0.467]
Age	0.151***	0.163***	0.374***	0.160***	-0.068***	-0.222***
	[0.023]	[0.024]	[0.044]	[0.045]	[0.021]	[0.028]
Side-Effect-Age Ineraction	-0.156***	-0.170***	-1.276*	-0.354***	-1.017***	-0.219
	[0.046]	[0.050]	[0.667]	[0.069]	[0.147]	[0.143]
Branded Drug without Generic	0.865***	0.516**	1.629**	1.352***	2.038***	1.589**
	[0.203]	[0.221]	[0.747]	[0.313]	[0.540]	[0.684]
Branded Drug with Generic	-1.337***	-0.949***	-1.337***	-0.180*	-0.882***	-0.352***
	[0.173]	[0.221]	[0.173]	[0.101]	[0.179]	[0.131]
Constant	2.142***	2.015***	10.231***	1.887**	10.780***	3.382***
	[0.333]	[0.349]	[2.605]	[0.838]	[3.873]	[0.868]

Notes: Dependent variable: $\ln(\sigma_{ijt})-\ln(\sigma_{0t})$. Standard errors are in parentheses. *, **, *** indicate 10%, 5%, 1% statistical significance, respectively. Number of observations: 658. Side Effects include anti-cholinergic, drowsiness, cardiac arrhythmias, orthostatic hypotension and gastrointestinal distress effects. *Branded Drug with(out) Generics* refers to the drugs where there has (not) been generic introduction in the respective molecule. Price is in logs.

TABLE VIII

LIMITED INFORMATION RANDOM COEFFICIENT MODEL

Variable	Means (α & β's)	Standard Deviations (Λ)	Interactions with Demographics: Income	Income Sqrd	Prescription Insurance
Price	-0.872***	0.131***	0.134***	-0.041*	0.832*
	[0.042]	[0.020]	[0.038]	[0.013]	[0.488]
Half-life	-0.355***	0.288**	-	-	-
	[0.047]	[0.120]			
Fatal	-0.508**	0.127***	-	-	-
	[0.172]	[0.040]			
Weight Gain	-0.322***	0.896	-	-	-
	[0.113]	[0.998]			
Insomnia/ Agitation	-0.340*	0.350***	-	-	-
	[0.205]	[0.095]			
Side Effects	-1.076**	0.183	-	-	-
	[0.467]	[0.293]			
Age	-0.222***	0.014	-	-	-
	[0.028]	[0.013]			
Side-Effect-Age Ineraction	-0.219	0.131	-	-	-
	[0.143]	[1.004]			
Branded Drug without Generic	1.589**	0.144***	1.256**	0.059*	0.467***
	[0.684]	[0.013]	[0.503]	[0.033]	[0.000]
Branded Drug with Generic	-0.352***	0.020*	0.060*	0.002	-0.465*
	[0.131]	[0.011]	[0.036]	[0.093]	[0.252]
Constant	3.382***	0.676	-	-	-
	[0.868]	[0.995]			

Notes: Dependent variable: $\ln(\sigma_{ijt})$-$\ln(\sigma_{0t})$. Standard errors are in parentheses. *, **, *** indicate 10%, 5%, 1% statistical significance, respectively. Number of observations: 658. Side Effects include anti-cholinergic, drowsiness, cardiac arrhythmias, orthostatic hypotension and gastrointestinal distress effects. *Branded Drug with(out) Generics* refers to the drugs where there has (not) been generic introduction in the respective molecule. Price is in logs.

TABLE IX
RANDOM COEFFICIENT LOGIT ELASTICITIES: PRICE

	Price Sensitivities			
	Full Information		Limited Information	
	Low Income	High Income	Low Income	High Income
Without Any Insurance	-1.108***	-0.889***	-0.872***	-0.273***
	[0.025]	[0.012]	[0.042]	(0.080)
With Full Insurance	-0.297***	-0.165	-0.240**	-0.055***
	[0.101]	[0.138]	[0.330]	[0.019]

Notes: Standard errors are in parentheses. *, **, *** indicate 10%, 5%, 1% statistical significance, respectively.

TABLE X
RANDOM COEFFICIENT LOGIT ELASTICITIES: 'BRANDNESS'

	Brand Sensitivities							
	Full Information				Limited Information			
	Low Income		High Income		Low Income		High Income	
	(Generics)	(No Gen.)	(Generics)	(No Gen.)	(Generics)	(No Gen.)	(Generics)	(No Gen.)
Without Any Insurance	-0.882***	2.038***	-0.464*	3.082***	-0.352***	1.589**	-0.290	2.591***
	[0.179]	[0.540]	[0.256]	[0.284]	[0.131]	[0.684]	[0.260]	[0.331]
With Full Insurance	-0.588**	3.870**	-0.068*	4.759**	0.113	2.056***	0.175*	3.371**
	[0.299]	[1.681]	[0.056]	[1.860]	[0.077]	[0.684]	[0.102]	[1.346]

Notes: Standard errors are in parentheses. *, **, *** indicate 10%, 5%, 1% statistical significance, respectively. *Generics* column refers to branded drugs with generic entry in its molecule. *No Gen.* column refers to branded drugs with no generic entry within their molecule.

TABLE XI
INFORMATION TECHNOLOGY PARAMETER ESTIMATES

	Coefficients for Interactions With Media			
Variable	Print	Television	Internet	Radio
Consumer Information Heterogeneity Coefficients				
Media and demographic interactions (Y)				
Constant	-1.143***	-0.960***	-0.867***	-1.150***
	[0.035]	[0.039]	[0.044]	[0.040]
Age: 35-54	0.067***	0.020	-0.034	-0.034
	[0.025]	[0.027]	[0.023]	[0.026]
Age: > 54	0.281***	0.216***	0.005	-0.270***
	[0.024]	[0.022]	[0.023]	[0.026]
Married	0.092***	0.075***	-0.023	-0.013
	[0.021]	[0.021]	[0.018]	[0.016]
Income:	-0.299***	0.492***	0.002	-0.270***
< $25,000	[0.028]	[0.020]	[0.022]	[0.030]
Income:	-0.202***	0.131***	-0.182***	-0.123***
$25,000-$49,999	[0.018]	[0.019]	[0.018]	[0.024]
Income:	0.140***	-0.022	0.141***	0.155***
> $100,000	[0.032]	[0.032]	[0.034]	[0.030]
White male	-0.040**	-0.018	-0.070***	0.006
	[0.018]	[0.017]	[0.015]	[0.017]
Maximum Education:	-0.183***	0.252***	-0.108***	0.073***
12 years	[0.025]	[0.025]	[0.030]	[0.028]
Maximum Education:	-0.068**	0.270***	0.031	0.114***
some college	[0.031]	[0.030]	[0.026]	[0.032]
Maximum Education:	-0.052**	0.141***	-0.021	0.083***
college	[0.020]	[0.027]	[0.027]	[0.024]
Maximum Education:	-0.048***	0.028***	-0.032***	-0.012***
< 11 years	[0.004]	[0.002]	[0.004]	[0.003]
Prescription Drug	0.095***	0.016***	0.040***	0.010*
Insurance Coverage	[0.020]	[0.005]	[0.005]	[0.006]
Advertising media exposure (ζ)				
media exposure *	0.796***			
advertising	[0.064]			
Demographics (λ)				
Constant	0.208***			
	[0.006]			
Income:	0.587***			
< $25,000	[0.077]			
Income:	0.432***			
$25,000-$49,999	[0.087]			
Income:	0.142			
> $100,000	[0.370]			
Prescription Drug	0.759***			
Insurance Coverage	[0.029]			

Notes: Standard errors are in parentheses. *, **, *** indicate 10%, 5%, 1% statistical significance, respectively. Dummy variables unless units are specified

TABLE XI (continued)
INFORMATION TECHNOLOGY PARAMETER ESTIMATES

Variable	
Information Technology Coefficients Common Across Consumers	
Age of drug	0.143***
	[0.004]
(Age of drug)^2	-0.077***
	[0.005]
Direct-to-Consumer Media Advertising (φ, ρ)	
Print advertising	0.415***
	[0.029]
TV advertising	0.949**
	[0.443]
Internet advertising	0.706***
	[0.043]
(Print advertising)^2	-0.004
	[0.021]
(TV advertising)^2	-0.062***
	[0.014]
(Internet advertising)^2	-0.069***
	[0.015]
Physician-directed Advertising	
Detailing	0.414***
	[0.103]
Journal Advertising	-0.131
	[0.093]
Samples	0.433***
	[0.087]
(Detailing)^2	-0.004*
	[0.002]
(Journal advertising)^2	0.073
	[0.051]
(Samples)^2	-0.011**
	[0.005]

Notes: Standard errors are in parentheses. *, **, *** indicate 10%, 5%, 1% statistical significance, respectively. Dummy variables unless units are specified

TABLE XII

PRICE AND ADVERTISING ELASTICITIES FOR SELECTED ANTIDEPRESSANTS, 2000-2001

Type	Molecule	Drug	Drug	OPE	Cross-Price Elasticities: MAOI (3)	TCA (26)	NewGen (3)	SSRI (8)	NDRIs (3)	SNRI (2)	NaSSA (2)	ALL (47)
Price Elasticities Under Full Information												
3	15	30	Serzone	-1.1645	0.0252	0.4029	0.6825	0.0870	0.0817	0.1204	0.0147	0.0706
4	17	33	Celexa	-0.1470	0.1053	0.1685	0.1584	0.7437	0.3419	0.5034	0.0617	0.2615
4	18	34	Prozac	-0.5415	0.1689	0.2703	0.2541	0.8347	0.2742	0.4038	0.0989	0.4716
4	18	36	Sarafem	-0.8147	0.1051	0.1683	0.1582	0.7588	0.3414	0.5027	0.0616	0.2115
4	19	37	Fluvoxamine [G]	-2.4038	0.0189	0.0303	0.0285	1.1007	0.0615	0.4526	0.1109	0.0249
4	19	38	Luvox	-1.4660	0.0130	0.2086	0.1961	0.7129	0.4233	0.3116	0.0764	0.4041
4	20	39	Paxil	-0.7208	0.0549	0.2196	0.2064	0.7002	0.4455	0.3936	0.0321	0.4445
4	21	40	Zoloft	-0.5991	0.0186	0.2969	0.2791	0.6813	0.3012	0.4435	0.0109	0.4798
5	22	41	Bupropion [G]	-1.3992	0.0372	0.0595	0.0559	0.2140	1.5785	0.5332	0.2177	0.0996
5	22	42	Wellbutrin	-0.6952	0.0243	0.1299	0.1221	0.2103	0.8289	0.4656	0.1426	0.0737
5	22	43	Wellbutrin SR	-1.0743	0.0399	0.1595	0.1999	0.2296	1.1112	0.3813	0.0234	0.1037
6	23	44	Effexor	-0.4936	0.0276	0.1105	0.2077	0.2385	0.4482	2.6565	0.0162	0.1574
6	23	45	Effexor-XR	-1.8536	0.0747	0.1495	0.1124	0.2151	0.4851	1.3384	0.0438	0.1952
Price Elasticities Under Limited Information												
3	15	30	Serzone	-1.1032	0.0167	0.2959	0.6776	0.0458	0.0616	0.0832	0.0088	0.0651
4	17	33	Celexa	-0.1278	0.0929	0.1341	0.1185	0.6729	0.2564	0.4159	0.0362	0.2045
4	18	34	Prozac	-0.4766	0.1199	0.2464	0.1988	0.5402	0.2491	0.3413	0.0832	0.3160
4	18	36	Sarafem	-0.7313	0.1160	0.1309	0.0991	0.7051	0.3031	0.4442	0.0365	0.0200
4	19	37	Fluvoxamine [G]	-2.1277	0.0174	0.0253	0.0237	1.1742	0.0504	0.3377	0.0771	0.0229
4	19	38	Luvox	-0.6544	0.0113	0.1851	0.1684	0.4912	0.3919	0.3021	0.0668	0.2557
4	20	39	Paxil	-0.6788	0.0324	0.1774	0.1739	0.5788	0.3012	0.2743	0.0265	0.4680
4	21	40	Zoloft	-0.4851	0.0113	0.2550	0.2005	0.5201	0.0977	0.1739	0.0078	0.3688
5	22	41	Bupropion [G]	-1.2495	0.0265	0.0535	0.0474	0.1876	1.6126	0.3290	0.2384	0.0888
5	22	42	Wellbutrin	-0.4312	0.0186	0.1046	0.1160	0.1316	0.1610	0.3835	0.1226	0.0727
5	22	43	Wellbutrin SR	-0.9808	0.0316	0.1220	0.1121	0.1928	0.9625	0.3392	0.0160	0.0698
6	23	44	Effexor	-0.4011	0.0191	0.0622	0.1514	0.1398	0.4422	2.5552	0.0169	0.1629
6	23	45	Effexor-XR	-1.1656	0.0695	0.0883	0.0833	0.2042	0.5415	1.1880	0.0358	0.1563
Advertising Elasticities												
3	15	30	Serzone	0.0038	-0.0024	-0.0013	-0.0173	-0.0020	-0.0104	-0.0209	-0.0058	-0.0041
4	17	33	Celexa	0.0438	-0.0013	-0.0071	-0.0708	-0.0561	-0.0572	-0.0115	-0.0322	-0.0236
4	18	34	Prozac	0.0173	-0.0032	-0.0017	-0.0170	-0.0599	-0.0137	-0.0276	-0.0077	-0.0148
4	18	36	Sarafem	0.0301	-0.0015	-0.0080	-0.0798	-0.0096	-0.0645	-0.0130	-0.0363	-0.0175
4	19	37	Fluvoxamine [G]	-	-	-	-	-	-	-	-	-
4	19	38	Luvox	0.0043	-0.0003	-0.0002	-0.0017	-0.0024	-0.0014	-0.0028	-0.0079	-0.0012
4	20	39	Paxil	0.0413	-0.0027	-0.0143	-0.0143	-0.0427	-0.0116	-0.0233	-0.0651	-0.0208
4	21	40	Zoloft	0.0250	-0.0017	-0.0092	-0.0923	-0.0585	-0.0746	-0.0150	-0.0419	-0.0282
5	22	41	Bupropion [G]	-	-	-	-	-	-	-	-	-
5	22	42	Wellbutrin	0.0327	-0.0073	-0.0384	-0.0384	-0.0600	-0.0629	-0.0624	-0.0174	-0.0418
5	22	43	Wellbutrin SR	0.0302	-0.0053	-0.0279	-0.0279	-0.0436	-0.0187	-0.0454	-0.0127	-0.0287
6	23	44	Effexor	0.0100	-0.0074	-0.0039	-0.0390	-0.0061	-0.0315	-0.0265	-0.0177	-0.0100
6	23	45	Effexor-XR	0.0257	-0.0072	-0.0038	-0.0379	-0.0059	-0.0306	-0.0975	-0.0172	-0.0128

Notes: Average of elasticities for the last two years of the dataset: 2000 and 2001, calculated from raw data. Specifically, column OPE carries the own-price elasticities and the other columns the cross-price elasticities of each drug displayed against all other drugs averaged by type. If the drug is compared to its own type, it is excluded from the average. Types 3-6 stand for NewGen, SSRI NDRI, and SNRI respectively. The number of drugs in each type of antidepressants is displayed under each type heading. [G] indicates generic drugs, which are not advertised, hence, no advertising elasticities.

TABLE XIII

MORAL HAZARD EFFECT IN SELECTED ANTIDEPRESSANTS 1981 - 2001

Type	Molecule	Drug No.	Drug Name	Generic Name	Entry	Patient Surplus					
						Full Information			Limited Information		
						Total Welfare (thousand $)	Per Unit Welfare ($)	Annual Moral Hazard ($)	Total Welfare (thousand $)	Per Unit Welfare ($)	Annual Moral Hazard ($)
3	15	30	Serzone	Nefazodone	1995	212	0.01	3	177,684	6.20	2,263
4	17	33	Celexa	Citalopram	1998	1,356	0.11	38	37,827	2.93	1,070
4	18	34	Prozac	Fluoxetine	1988	1,995,377	24.46	8,929	2,194,698	26.91	9,821
4	18	36	Sarafem	Fluoxetine	2000	3,353	0.84	305	62,583	15.62	5,700
4	19	37	Generic	Fluvoxamine	2000	8,669	7.82	2,856	8,917	8.05	2,937
4	19	38	Luvox	Fluvoxamine	1994	461	1.81	661	414	1.63	595
4	20	39	Paxil	Paroxetine	1993	734	0.01	5	1,003,593	19.00	6,935
4	21	40	Zoloft	Sertraline	1992	36	0.00	0	6,006,786	52.83	192,827
5	22	41	Generic	Bupropion	1999	459	0.49	180	2,920	3.14	1,145
5	22	42	Wellbutrin	Bupropion	1989	1,430	0.32	117	25,461	5.72	2,089
5	22	43	Wellbutrin SR	Bupropion	1996	1,362	1.49	545	1,671	1.83	668
6	23	44	Effexor	Venlafaxine	1994	297	0.01	3	48,444	1.20	4,394
6	23	45	Effexor-XR	Venlafaxine	1997	1,436	0.34	125	4,906	1.17	428

Notes: Tables shows value of innovation for selected antidepressants. Types 3-6 stand for NewGen, SSRI, NDRI and SNRI, respectively.

TABLE A.I
LIMITED INFORMATION DEMAND ESTIMATION ACROSS MODELS

Variable	OLS Logit	IV Logit	OLS NML1a	IV NML1a	OLS NML1b	IV NML1b	OLS NML2	IV NML2
	Price Coefficients (α's)							
Price	-0.421***	-2.560***	-0.379*	-2.066***	-0.144***	-2.620***	-0.155**	-2.709***
	[0.091]	[0.164]	[0.200]	[0.163]	[0.068]	[0.116]	[0.071]	[0.124]
	Group Correlation Coefficients (ρ's)							
Type (ρ_t)			0.513***	0.377***			0.788***	0.630***
			[0.035]	[0.043]			[0.080]	[0.088]
Molecule (ρ_m)					0.815***	0.773***	0.812***	0.753***
					[0.035]	[0.037]	[0.035]	[0.038]
	Coefficients of Characteristics (β's)							
Half-life	0.004	0.001	0.008***	0.005	0.006**	0.005*	0.006**	0.004
	[0.003]	[0.003]	[0.003]	[0.003]	[0.002]	[0.002]	[0.002]	[0.002]
Fatal	-0.491***	-1.031***	-0.711***	-1.091***	-0.545***	-0.761***	-0.536***	-0.731***
	[0.178]	[0.207]	[0.155]	[0.180]	[0.131]	[0.142]	[0.132]	[0.142]
Weight Gain	-0.300***	-0.609***	-0.207*	-0.348***	-0.033	-0.171*	-0.039	-0.213**
	[0.115]	[0.133]	[0.123]	[0.121]	[0.085]	[0.093]	[0.086]	[0.095]
Insomnia/Agitation	0.554***	0.516***	0.831***	0.726***	0.506***	0.493***	0.492***	0.419***
	[0.130]	[0.145]	[0.115]	[0.128]	[0.096]	[0.099]	[0.099]	[0.105]
Side Effects	-3.596***	-2.454***	-4.131***	-3.062***	-2.040***	-1.658***	-1.975***	-1.299**
	[0.758]	[0.855]	[0.660]	[0.746]	[0.561]	[0.586]	[0.574]	[0.613]
Age	0.182***	0.234***	0.229***	0.259***	0.154***	0.176***	0.151***	0.163***
	[0.030]	[0.034]	[0.026]	[0.029]	[0.022]	[0.023]	[0.023]	[0.024]
Side-Effect-Age Ineraction	-0.204***	-0.274***	-0.244***	-0.282***	-0.242***	-0.218***	-0.156***	-0.170***
	[0.056]	[0.045]	[0.044]	[0.049]	[0.045]	[0.042]	[0.046]	[0.050]
Branded Drug without Generic	0.427***	0.654**	0.988***	0.904***	0.894***	0.682***	0.865***	0.516**
	[0.116]	[0.289]	[0.239]	[0.126]	[0.195]	[0.206]	[0.203]	[0.221]
Branded Drug with Generic	-2.563***	-2.031***	-1.880***	-1.121***	-1.315***	-1.044***	-1.337***	-0.949***
	[0.201]	[0.274]	[0.195]	[0.262]	[0.173]	[0.229]	[0.173]	[0.221]
Constant	3.822***	2.953***	2.085***	1.842***	2.104***	1.840***	2.142***	2.015***
	[0.432]	[0.491]	[0.394]	[0.439]	[0.325]	[0.341]	[0.333]	[0.349]
R^2	0.343	0.311	0.381	0.237	0.558	0.525	0.559	0.517
Adjusted R^2	0.334	0.302	0.372	0.225	0.552	0.518	0.551	0.509

Notes: Dependent variable: $\ln(\sigma_{jt})-\ln(\sigma_{0t})$. Standard errors are in parentheses. *, **, *** indicate 10%, 5%, 1% statistical significance, respectively. Number of observations: 658. Side Effects include anti-cholinergic, drowsiness, cardiac arrhythmias, orthostatic hypotension and gastrointestinal distress effects. *Branded Drug with(out) Generics* refer to the drugs where there has (not) been generic introduction in the respective molecule. Price is in logs.

TABLE A.II
LIMITED INFORMATION DEMAND ESTIMATION WITH DIFFERENT SPECIFICATIONS FOR NML2

	Ordinary Least Squares			Instrumental Variables		
	(1)	(3)	(5)	(1)	(2)	(3)
Variable	Price Coefficients (α's)					
Price	-0.155**	-0.170**	-0.091***	-2.709***	-1.745***	-1.475***
	[0.071]	[0.072]	[0.022]	[0.124]	[0.125]	[0.120]
	Group Correlation Coefficients (ρ's)					
Type (ρ_t)	0.788***	0.720***	0.724***	0.630***	0.570***	0.593***
	[0.080]	[0.079]	[0.081]	[0.088]	[0.088]	[0.089]
Molecule (ρ_m)	0.812***	0.807***	0.820***	0.753***	0.747***	0.776***
	[0.035]	[0.036]	[0.036]	[0.038]	[0.039]	[0.038]
	Coefficients of Characteristics (β's)					
Half-life	0.006**	0.005**	0.007***	0.004	0.006**	0.007***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Fatal	-0.536***	-0.432***		-0.731***	-0.635***	
	[0.132]	[0.126]		[0.142]	[0.137]	
Weight Gain	-0.039	-0.172**		-0.213**	-0.344***	
	[0.086]	[0.081]		[0.095]	[0.090]	
Insomnia/Agitation	0.492***			0.419***		
	[0.099]			[0.105]		
Side Effects (5)	-1.975***			-1.299**		
	[0.574]			[0.613]		
Side Effects (6)		-2.298***			-1.400*	
		[0.728]			[0.779]	
Side Effects (8)			-2.549***			-1.587*
			[0.779]			[0.909]
Age	0.151***	0.133***	0.078***	0.163***	0.144***	0.073***
	[0.023]	[0.026]	[0.022]	[0.024]	[0.027]	[0.022]
Side-Effect-Age	-0.156***	-0.176***	-0.098***	-0.170***	-0.174***	-0.086*
Ineraction	[0.046]	[0.022]	[0.021]	[0.050]	[0.025]	[0.047]
Branded Drug without	0.865***	0.446**	0.784***	0.516**	0.844***	0.620***
Generic	[0.203]	[0.180]	[0.166]	[0.221]	[0.127]	[0.174]
Branded Drug with	-1.337***	-1.328***	-1.023***	-0.949***	-1.039***	-0.757***
Generic	[0.173]	[0.162]	[0.162]	[0.221]	[0.196]	[0.195]
Constant	2.142***	1.963***	1.711***	2.015***	1.831***	1.496***
	[0.333]	[0.353]	[0.366]	[0.349]	[0.371]	[0.379]
R^2	0.559	0.546	0.521	0.517	0.501	0.494
Adjusted R^2	0.551	0.539	0.516	0.509	0.493	0.487
1st Stage F-stat				92.61	100.25	103.98
Prob > F				0.000	0.000	0.000
Log Likelihood	-1,204	-1,213	-1,230			

Notes: Dependent variable: $\ln(\sigma_{jt})-\ln(\sigma_{0t})$. Standard errors are in parentheses. *, **, *** indicate 10%, 5%, 1% statistical significance, respectively. Number of observations: 658. Side Effects (5) include anti-cholinergic, drowsiness, cardiac arrhythmias, orthostatic hypotension and gastrointestinal distress effects. Side Effects (6) include, in addition, insomnia/agitation. Side Effects (8) include in addition fatal and weight gain. *Branded Drug with(out) Generics* refer to the drugs where there has (not) been generic introduction in the respective molecule. Price is in logs.

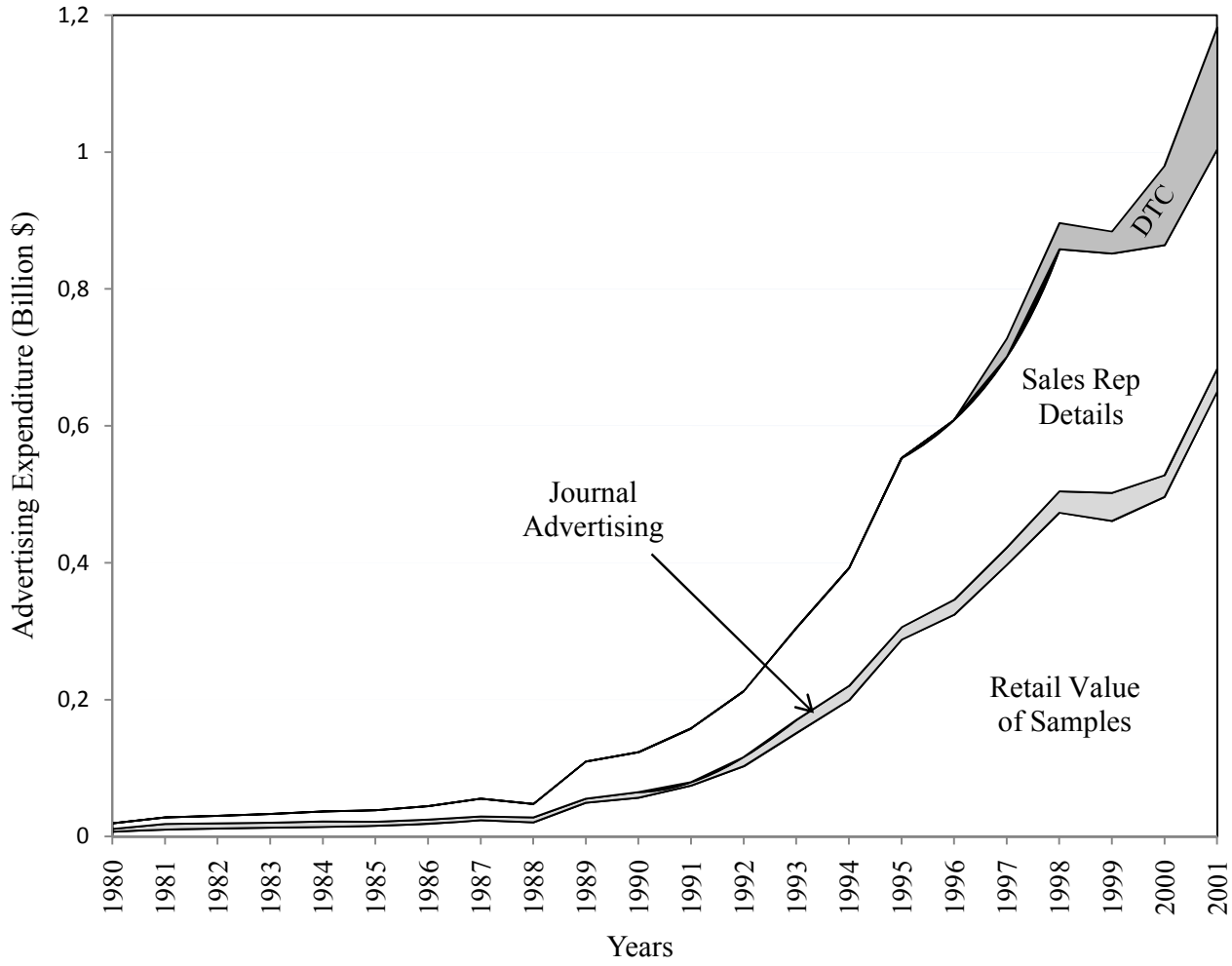

Figure 1
Advertising Expenditure Evolution by Type in the U.S. Antidepressant Market

Notes: DTC includes advertising for prescription drugs on TV, radio, magazines and newspapers, as well as internet and outdoor advertising; professional journal advertising reflects advertising expenditures for prescription drugs appearing in medical journals; sales representative details include costs associated with the sales activities of pharmaceutical representatives that are directed to office-based physicians, hospital-based physicians, and directors of pharmacies; samples are prescription drugs given to physicians to disseminate freely to patients.

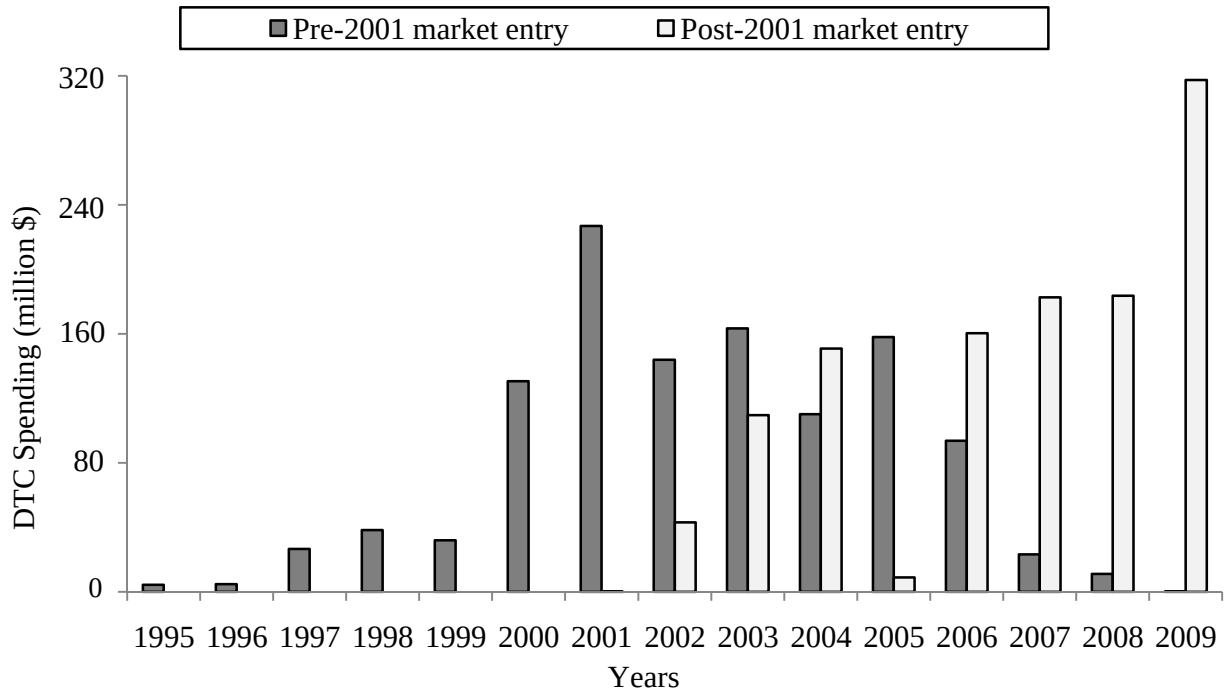

Figure 2
Direct-To-Consuming (DTC) Advertising Evolution For Older and Newer Antidepressants

Notes: DTC includes internet, outdoor, print, radio, and television advertising.

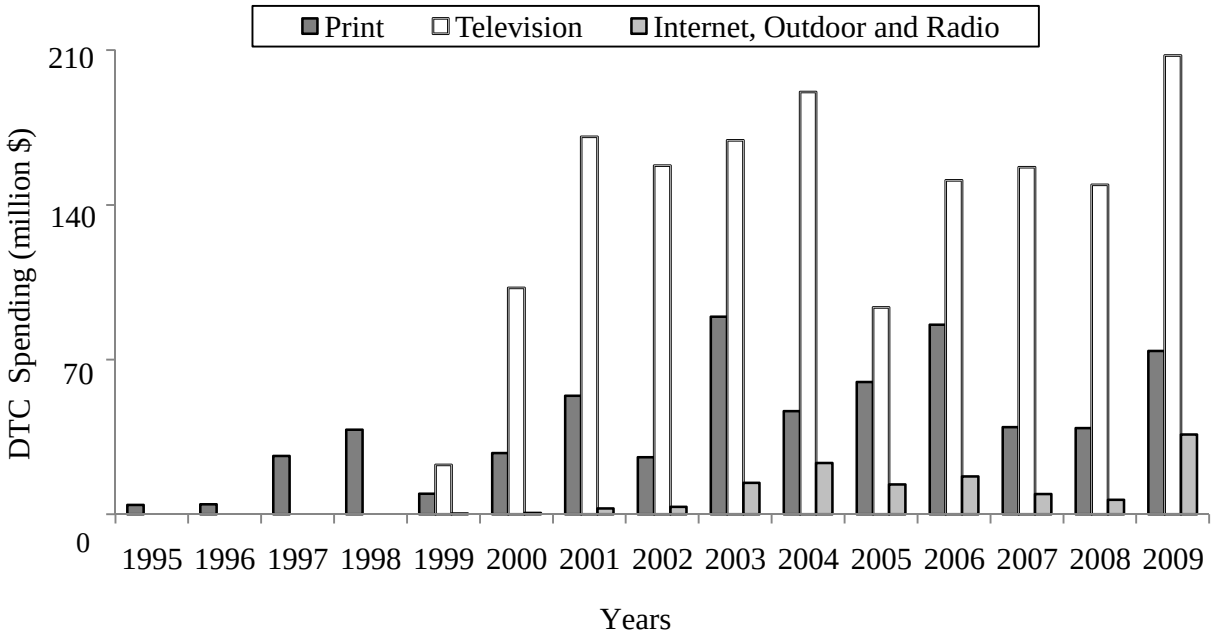

Figure 3
Direct-To-Consumer (DTC) Advertising in the U.S. Antidepressant Market by Medium

Notes: Print includes Business-to Business, Consumer, Hispanic Local and Sunday Magazines and Hispanic, Local and National Newspapers; Television includes Cable, Network, Spanish Language and Spot Television and Syndication; Radio includes Local, National Spot and Network Radio.