DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

EVALUATING INNOVATION AND MORAL HAZARD IN PHARMACEUTICALS

Paris Cleanthous

Discussion Paper 03-2011

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

EVALUATING INNOVATION AND MORAL HAZARD IN PHARMACEUTICALS[1]

PARIS CLEANTHOUS[2]

March 2010

Abstract

This paper formulates an empirical methodology that evaluates pharmaceutical innovation in the American antidepressant market by quantifying patient welfare benefits from innovation. While evaluating pharmaceutical innovation in antidepressants, I uncover and address the moral hazard issue that arises due to the existence of prescription drug insurance coverage. A combination of market-level data, drug and patient characteristics are used to estimate demand for all antidepressants between 1980 and 2001. The paper estimates large and varied patient welfare gains due to innovation and helps explain a detected divergence between social and private patient benefits by the existence of insurance.

KEYWORDS: Health, Innovation, Moral Hazard, Pharmaceuticals, Welfare
JEL classification: D6, I1, L65, O31

[1]I have greatly benefited from discussions with Patrick Bayer, Steven Berry, Luis Cabral, Sofronis Clerides, Iain Cockburn, Jean-Pierre Dubé, Nicholas Economides, Sarah Ellison, Nancy Epling, Gautam Gowrisankaran, Jerry Hausman, Günter Hitsch, Pinelopi Goldberg, Elena Krasnokutskaya, Sangin Park, Fiona Scott Morton, Subrata Sen, Scott Stern, K. Sudhir, Christopher Timmins, and seminar participants at Boston University, Columbia University, ITAM, London Business School, New York University, Northwestern University, SUNY Buffalo, Texas A&M, UCLA, University of Chicago, University of Cyprus, Yale University. I also thank Steven Propper, John Kennedy, Susan Schmitter Jack and many others for introducing me to this industry, IMS Health and anonymous consultants for graciously providing me with their datasets. Financial support is gratefully acknowledged from the John Perry Miller Fund Award at Yale University, 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

Returns to innovation constitute a major component of social welfare. To evaluate innovations requires use of a methodology that isolates their precise effect on welfare. This is particularly important in the pharmaceutical industry where rapid innovation occurs and drug development costs are high and increasing. In this paper, I formulate an empirical methodology that quantifies patient welfare benefits from pharmaceutical innovation in the U.S. antidepressant market and I address the moral hazard issue caused by the existence of prescription drug insurance coverage.

The antidepressant market has experienced an impressive stream of innovations over the last four decades and available data on antidepressants are exceptionally rich and accurate. The paper employs an original dataset that consists of annual observations on prices, quantities and drug characteristics for every antidepressant medication sold in the U.S. market from 1980 to 2001. Data also include information on the segmentation of the therapeutic area of antidepressants into different categories of drugs as well as information on branded and generic entry of antidepressants in the U.S. market. Sales data are mostly from IMS Health Inc.; the main sources for drug characteristics data are the Food and Drug Administration (FDA) and the Drug Information Handbook; patient characteristics data come from the National Center for Health Statistics. The latter are time-varying demographic data on the distribution of patient income and out-of-pocket prescription drug expenditures.

The paper utilizes a structural discrete choice model of demand for procuct characteristics to estimate the changes in patient welfare due to antidepressant introduction. To obtain correct substitution patterns between drugs, the model includes unobserved drug characteristics. These, in turn, make necessary the use of instrumental variable techniques to correct for the endogeneity of prices. I estimate a full random coefficients multinomial logit model, which contributes to the literature in several ways. The model allows for patient observed and unobserved heterogeneity in both patient willingness-to-pay and taste for branded drugs over generics. Draws from a joint distribution of income and prescription drug insurance coverage model the observed patient heterogeneity. Draws from an assumed multivariate normal distribution approximate the unobserved heterogeneity of patient preferences.

In addition, the model allows for unobserved patient heterogeneity in the valuation of dif-

ferent drug characteristics that reflects the idiosyncrasy of antidepressant side effects. 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 elasticities of demand, which describe patient substitution patterns. Finally, parameter estimates allow me to estimate patient willingness-to-pay for hypothetical drugs that might be introduced in the future.

The inclusion of prescription drug insurance as an observed patient characteristic lets me estimate patient willingness-to-pay separately for those with and without insurance, and draw implications for patient welfare. 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 gains help evaluate the importance and success of different innovations in the antidepressant market.

The findings of this paper are relevant to health care and pharmaceutical industry public policy. For instance, sky rocketing drug costs have emerged as a potent public policy issue. Consumers view rising pharmaceutical prices as a result of unfair pricing policies. The pharmaceutical industry argues that rising prices are due to the increases in the amounts of research and development required to find new medicines to cure diseases and relieve suffering and the surge in administrative costs from the time of innovation to market entry. In fact, public opinion indicates that price controls on pharmaceutical products are only favorable when these do not hurt the industry's ability to conduct research. The estimated magnitudes of patients' welfare gains from pharmaceutical innovations presented in this paper, especially after controlling for prescription insurance, can help evaluate the merits of competing claims. The paper finds positive and excess patient willingness-to-pay for every antidepressant drug in the choice set and evaluates its relative importance between different drugs.

The remainder of the paper is organized as follows. Section II reviews the relevant empirical

literature. Section III analyzes the characteristics of the market for antidepressants and the pertinent characteristics of the pharmaceutical industry. Section IV presents the adaptation of previous theory and the methodology in estimating demand for antidepressants. The paper focuses on the full random coefficients multinomial logit model and the inclusion of demographic data on the distribution of patient income and prescription drug insurance. The data, estimation procedure and results are presented and discussed in Sections V, VI and VII, respectively. Section VIII uses the demand estimation results to infer welfare implications of innovation in antidepressants. Section IX concludes.

II RELATED EMPIRICAL LITERATURE

Technological advances in the latter half of the twentieth century spurred an amazing stream of invention and innovation of new products. The continuous introduction of new products has motivated economists to search for methodologies that evaluate the economic importance of new goods. Bresnahan and Gordon (1997) provide a thorough review of the economics-of-new-goods literature in a collection of essays, which include historical treatments of new goods and their diffusion; practical exercises in evaluating innovations; and real-world methods of devising quantitative adjustments for quality change. Among notable work in the area Trajtenberg (1990) analyzes the welfare implications of innovation in computed tomography scanners; Hausman (1996) estimates welfare gains generated by a new brand of cereal; Bresnahan, Stem and Trajtenberg (1997) compute rents from innovation in personal computers; and Petrin (2002) studies the welfare gains from the introduction of the minivan. Petrin uses demand and cost side estimates from observed data to recompute equilibrium prices and quantities from a choice set that does not include the minivan. I, instead, recompute welfare given the estimated demand parameters only since cost data are not readily available. Then by excluding each innovative drug from the choice set at the time of innovation, I find the change in consumer welfare caused by the innovation.

In the pharmaceutical industry, work that attempts to quantify the economic value of innovation has been scarce. Lichtenberg (1996, 2000, 2001a, 2003) estimates the contribution of pharmaceutical innovation to consumer welfare through reductions in mortality, morbidity and total medical expenditure. In all papers, Lichtenberg estimates large gains to consumers. Whereas the first three papers do not deal explicitly with patient characteristics, Lichtenberg (2001a) uses data that

links medicine event-level data to patient-level data to provide evidence on the negative correlation between a drug's age and mortality, morbidity and total medical expenditure. The study supports that, though cheaper generic drugs might seem as an effective way to reduce health expenditure, branded drugs tend to be younger and, therefore, better so that their use reduces total treatment costs. Murphy and Topel (2003) and Lichtenberg (2001b) estimate the economic value of biomedical research via changes in life expectancy. They find that the economic return to improvements in health are greater for larger populations, when average lifetime incomes are higher, when existing levels of health are better, and as the age of the population approaches the typical age of the disease's onset.

Berry, Levinsohn and Pakes [BLP] (1995) examine equilibrium in the automobile industry and eliminate price endogeneity from the model by allowing for the existence of unobserved characteristics. The inversion of the market share function in Berry (1994) used to obtain the mean utility level of a product allows the use of standard instrumental variable techniques to estimate demand parameters. The use of individual-level data greatly complemented the results of studies that use market-level data to estimate demand. For instance, Goldberg (1995) estimates the demand for automobiles to investigate trade policy issues; BLP (2004) show how use of micro-level data augments the importance of the use of market level data; Nevo (2001) estimates demand in the cereal industry; Sudhir (2001) estimates consumer pricing behavior in the automobile industry.

A substantial economic literature exists on pharmaceuticals, and antidepressants especially. Berndt et al. (1996) examines how changing market conditions such as generic entry and entry of new products affects price index calculation and interpretation and Berndt et al. (2002) investigate the changes in treatment price indices for acute phase major depression. Berndt et al. (1997) examine product-level demand for anti-ulcer medications. They concentrate on marketing variables, which are an important part of the new-good commercialization process in prescription drug markets, but not the only determinant of changes in consumer welfare. Their analysis also distinguishes between 'industry-expanding' and 'rivalrous' marketing efforts by looking at a natural experiment: the introduction of Tagamet and, later, Zantac.

III 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.[2] In fact, the pharmaceutical industry is the most research-intensive U.S. manufacturing industry. The quality of its products has been subjected to especially close regulation by the FDA, which regulates entry and maintains high product quality standards. In order to be approved by the FDA for marketing to the public, a drug must go through difficult and lengthy pre-clinical and clinical trials.

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. 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.[3]

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 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'.[4]

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.

[2] Scherer (2010)

[3] 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.

[4] 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.

i The Market for Antidepressant Drugs

I concentrate on the market for antidepressants because it contains a considerable number of drugs and has experienced substantial qualitative innovation over a long time, so there exists a benchmark against which to compare innovation. The therapeutic area of antidepressants includes prescription drugs[5] that are FDA-approved to be used in the pharmacological clinical depression. Drugs, therefore, that are therapeutically bioequivalent but are FDA approved for the treatment of different diseases will fall into different therapeutic areas. I take the antidepressant class as described by IMS Health Inc., USC codes 64300-64399.[6] Moreover, treatment of depression does not require combinations of drugs from different categories and antidepressants are not used to treat diseases other than depression. This elimintates market interaction that would complicate modeling demand.

Antidepressant drugs are used in the pharmacological treatment of clinical depression, a serious psychotic and incapacitating condition. The market of treating clinical depression is potentially quite lucrative. This is due to the fact that depression is highly prevalent and debilitating. According to the National Comorbidity Study the lifetime and annual prevalence of major depression are 16.9 percent and 9.7 percent respectively.[7] According to the World Health Organization depression will be the leading cause of disability and premature death in the industrial world by the year 2020. Moreover, it is chronic, has a high degree of recurrence and requires maintenance drug therapy. Depression, therefore, tends to be costly to treat as well as to have, especially when left untreated for a long period of time.

Moreover, the market could have even been more profitable were depression well-diagnosed and treated. Approximately fifty percent of Americans suffering from major depression seek professional care during a year and of those only about half go to psychiatrists.[8] Under-diagnosis and under-treatment may be due to various causes: Patients may not link their symptoms to a disease; public

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

[6] Suslow (1996) warrants against the possibility that the economic definition of a market may not coincide with the IMS definition. In the case of antidepressants, the two definitions are a close match. Some drugs that contain active ingredients of antidepressants and have fulfilled the FDA bioequivalence requirements are included in other categories by IMS since they are used primarily for the pharmacological treatment of other diseases. For example, Zyban has the antidepressant active ingredient Bupropion Hydrochloride, an antidepressant, but is used as an aid in the cessation of tobacco smoking. For the purposes of this study Zyban is not an antidepressant drug.

[7] National Comorbidity Survey with data updated as of July 19, 2007. Badamgarav et al (2003) review the psychiatric literature and find that estimates of the prevalence of depression vary from 15% to 25% for lifetime prevalence and from 10% to 20% for 12-month prevalence.

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

comprehension of mental diseases is generally poor; depression still constitutes a social stigma and primary care physicians miss diagnosing depression half of the time.[9]

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 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. 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. Table I reports in detail the historic entry of antidepressants into the market from 1958 to the present broken down by type and level of entry. 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. Other than that, the reported therapeutic advantage of Prozac was not any different to the existing drugs.

INSERT TABLE I ABOUT HERE

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 II 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

[9] Salmans (1995), Badamgarav et al (2003).

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 II ABOUT HERE

Historical evidence indicates that no one antidepressant is clearly more effective than another in achieving the desired health outcome.[10] 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 same active ingredient (molecule). Examples of side effects include a drug's fatality, dry mouth, blurred vision and drowsiness.

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.[11] 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.[12] What is more, existing formularies[13] only make a distinction between branded and

[10] Depression Guideline Panel (1993).

[11] Wosińska (2005) has a discussion about the interaction of the various agents when making a choice in pharmaceuticals, in general. The discussion, here, focuses on antidepressants.

[12] Depression Guideline Panel (1993).

[13] The Lewin Group (2000).

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 diagnose depression or to predict exact response to a particular treatment. Prescribing physicians have to rely on their patients to find out whether a certain pharmacological treatment is working out or not. As a result, in the case of depression, 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 of the existence of spurious product differentiation, very common in pharmaceuticals, where 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. In the case that a physician chooses to prescribe the brand rather than the molecule when a generic exists, a patient can alter this prescription via a pharmacist

in most states.[14]

IV METHODOLOGY

i A Discrete Choice Model of Demand

As is the case in most industries, pharmaceutical products are not homogeneous. Consequently, there is no single demand curve characterizing all drugs. A system of single-drug demand curves needs to be derived taking into account the attributes associated with the drugs' therapeutic area and other pertinent characteristics that might enter a patient's choice decision in antidepressants, such as the prices of other drugs in the market. Moreover, pharmaceutical firms set their own prices, even when facing competition. It becomes plausible to think of the pharmaceutical industry as oligopolistic where producers of each product face downward-sloping demand curves. Following the tradition of hedonic demand formulation,[15] 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 drug. The advantage of using a discrete choice model is that demand is built from a well-specified utility for drug characteristics. In contrast to representative consumer models where the 'representative' patient would have been modeled to have a 'taste' for consuming a variety of drugs, in discrete choice models the variety in antidepressant drugs is represented by the variety in individual patient preferences for drug characteristics. Market-level demand is then obtained by aggregating individual demands. Parameter estimates of the demand system use only drug level data on prices, quantities and characteristics. I, eventually, incorporate information on individual patient income and prescription drug insurance coverage.

[14] State legislation exists that incites the use of generic medication when that exists, for example, in Maine, Ellison and Snyder (2010).

[15] The alternative tradition in deriving demand is to model a representative patient who has a taste for consuming a variety of products. For example, see Dixit and Stiglitz (1977) and Ellison et al (1997). In the case of pharmaceuticals and antidepressant drugs especially this method is not viable. Depressive symptoms for the same antidepressant vary across individual patients. Moreover, patients are heterogeneous in their willingness-to-pay and their perceptions about product quality.

ii Model Overview

Assume that 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, $j = 0$.[16] For a given drug j, where $j = 0, 1, ..., J_t$, the level of utility that an individual patient i derives is represented by the general conditional indirect utility function, $u(\nu_{it}, x_{jt}, \xi_{jt}, p_{jt}; \theta_d)$. That is, utility is a function of a vector of individual patient characteristics ν, a vector of drug characteristics (x, ξ, p) and θ_d. Here, x and ξ represent the observed and unobserved (by the econometrician) drug characteristics, respectively, and p denotes the drug's price. θ_d is the vector of demand parameters to be estimated. Patients are assumed to observe all the drug characteristics.

As in BLP (1995) and Petrin (2002) the utility function at time period t is:

$$u_{ijt}(\nu_{it}, x_{jt}, \xi_{jt}, p_{jt}; \theta_d) = \delta_{jt}(\nu_{it}, x_{jt}, \xi_{jt}, p_{jt}; \theta_d) + \mu_{ijt}(\nu_{it}, x_{jt}, \xi_{jt}, p_{jt}; \theta_d) + \epsilon_{ijt} \quad (1)$$

where δ_{jt} is the mean utility level, that is, a drug-specific term common to all patients; μ_{ijt}, is a term that captures the heterogeneity in patient preferences for observed (by the econometrician) drug characteristics. The third component, ϵ_{ijt}, is a random utility component that is assumed to be independent and identically distributed across both drugs and patients and follows an extreme value distribution. 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.

At every time period t, each patient will purchase one unit of the drug that provides her with the highest utility. Conditional on observable and unobservable drug characteristics (x_j, ξ_j) and price p_j, patient i will, therefore, choose to purchase one unit of drug j, at time t, if and only if $u_{ijt}(\theta_d) - u_{ikt}(\theta_d) > 0$, $\forall k \geq 0$, $k \neq j$, $\forall t$. Considering a population of patients that consume each drug j, I can estimate the drug market shares, s_j, which will represent the different drug demands. Define the set of unobservable characteristics that will induce the choice of drug j as $A_j(\delta) = \{\mu_i \,|\, \delta_j + \mu_{ij} >$

[16] The existence of the outside option is of importance. In its absence, patients would be forced to choose among the J 'inside' options only. Consequently, demand for each drug j would depend on relative drug prices only. This would mistakenly imply that with an overall increase in antidepressant drug prices the total sales of antidepressant drugs would stay the same. Though the existence of drug reimbursement within health insurance packages makes the demand for pharmaceuticals highly insensitive to marginal price changes, demand is nevertheless not perfectly inelastic. Information on the distribution of insurance coverage and detailed information on drug reimbursement will help identify the correct price elasticities of demand. For example, not all Americans have insurance coverage and not all insurance packages include full drug reimbursement for mental health diseases.

$\delta_k + \mu_{ik}, \forall k \neq j\}$. Given a joint distribution $F(\mu;\sigma)$ for patient characteristics with density $f(\mu;\sigma)$, then s_j, is the probability that μ_i falls within $A_j(\delta)$: $s_j\left(\nu_i, x_j, \xi_j, p_j; \theta_d\right) = \int_{A_j(\delta)} f(\mu;\sigma)d\mu$. A closed form solution may exist depending on the density function chosen. Otherwise, simulation methods can be used to estimate the drug market shares.

I assume that the market size is directly observed and is equal to the portion of the U.S. population that is estimated to be clinically depressed. Given M, the observed output quantity of the firm producing product j, $q_j = M \cdot s_j(x, \xi, p; \theta_d)$, where the s_j represent market shares. Shares can thus be calculated from the data: $\widetilde{s}_j = q_j/M$. To solve for the parameters that enter the market share function, I set $\widetilde{s}_j = s_j(\delta(x, p, \xi); \theta_d)$, $\forall j$, which should hold exact at the true values of δ.

The mean utility level is as in Berry (1994):

$$\delta_{jt} \equiv \alpha p_{jt} + x_j'\beta + \xi_{jt} \tag{2}$$

where the β are the marginal utilities of the drug's observed characteristics and α is the marginal disutility associated with price. Note that this formulation of utility specifies that the unobserved characteristic ξ_j 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. However, a problem arises: drug prices are endogenous because pharmaceutical companies observe the ξ_j and take them into account when setting prices. Thus prices are expected to be correlated with the ξ_j. This requires use of instrumental variable (IV) econometric techniques in the estimation. The problem is exacerbated as the ξ_j enter the above equation nonlinearly and prohibit the use of standard IV techniques. It has been proven in Berry (1994) that with a known distribution of unobservable patient characteristics, f, market shares depend only on mean utility levels. Given a market share function that can be inverted, the means of patient utility are uniquely determined, $\delta_j = s_j^{-1}\left(\widetilde{s}_j\right)$. Traditional IV techniques to estimate the unknown parameters β and α are, therefore, feasible.

iii The Random Coefficients Logit Model

I employ a random coefficients specification for utility. The full random coefficient logit (RCL) model allows for patient heterogeneity using both a multivariate normal distribution and demographic data on prescription drug insurance and income. Moreover, it retains the dimensionality advantage, the flexibility in substitution patterns and, in addition, allows interaction between patient and drug characteristics. In particular, it addresses the need to capture patient heterogeneity in the valuation of the different drug characteristics as both depressive symptoms and side effects vary widely across individuals. More importantly, it models for patient heterogeneity in income and prescription drug insurance schemes. The RCL model incorporates patient heterogeneity by assuming a distribution for any unobserved heterogeneity and by using time-varying demographic data to model the distribution of patient income and prescription drug coverage. The latter approximates patient heterogeneity in price-sensitivity and preference for branded drugs.

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. Petrin (2002), however, explains that market-level data may or may not provide good information on how patients substitute between drugs. As remains to be seen in the estimation results, the use of demographics in this paper provide ample information on patients' substitution patterns and, hence, estimates of the parameters of the distribution of patient preferences are precise.

iv Including Information on Patient Heterogeneity

In the absence of patient-level data, I use aggregate-level information that relate average patient demographics to some of the drug characteristics (observed patient characteristics). As in Nevo (2000) I allow patient preferences to vary with the ν_i and incorporate the distribution of patient characteristics in the model. The ν_i are modeled as a combination of an observed component (patient demographics), D_i, and an unobserved component, τ_i. This allows the inclusion of infor-

mation about the distribution of the marginal disutilities of price and the preference for branded drugs obtained from demographic data. Though side effects also vary among individuals, this paper observes no patient-level data on side effects.

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:

$$\begin{pmatrix}\alpha_i \\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \Pi D_i + \Lambda\tau_i, \;\; D_i \tilde{} P_D^*(D), \;\; \tau_i \tilde{} P_\tau^*(\tau) \tag{3}$$

where D_i is a $d \times 1$ vector of demographic variables, τ_i is the unobserved component of patient characteristics, $P_\tau^*(.)$ is a parametric distribution, and $P_D^*(.)$ a non-parametric distribution derived from the data. Π is a $(K+1) \times d$ matrix of coefficients that measure the relation between demographics and patient preferences. K is the number of characteristics that enter in the model. Finally Λ is a $(K+1) \times (K+1)$ matrix of parameters.[17] To complete the model I assume that τ_i and D_i are independent and that $P_\tau^*(.)$ follows a standard multivariate normal distribution.

Given the latter distributional assumption, Λ allows a different variance for each component of τ_i and a correlation among these patient preferences. Let $\theta_d = [\theta_1, \theta_2]$[18] and rewrite the utility model in equation (1) using equations (2) and (3). Then,

$$\begin{aligned} u_{ijt}(\nu_{it}, x_{jt}, \xi_{jt}, p_{jt}; \theta_d) &= \delta_{jt}(x_{jt}, \xi_{jt}, p_{jt}; \theta_1) + \mu_{ijt}(\tau_{it}, D_{it}, x_{jt}, \xi_{jt}, p_{jt}; \theta_2) + \epsilon_{ijt} \\ \text{where} \;\; \delta_{jt} &= \alpha p_{jt} + x'_{jt}\beta + \xi_{jt}, \;\; \mu_{ijt} = [p_{jt}, x_{jt}] * (\Pi D_{it} + \Lambda \boldsymbol{\tau}_{it}) \end{aligned} \tag{4}$$

and the probability that a patient chooses drug j at time t can now be written for the random coefficients model 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:

$$s_{ijt} = \frac{\exp\left(\delta_{jt} + \mu_{ijt}\right)}{1 + \sum_{k \in J_t} \exp\left(\delta_{kt} + \mu_{ikt}\right)}. \tag{5}$$

[17] I allow all characteristics to have random coefficients but only price and preference for branded drugs are allowed to vary with demographics. The rest of the entries in Π are zeros forcing the distribution of patient characteristics for those other characteristics to only draw from their unobserved component.

[18] θ_1 enters the estimation linearly and θ_2 enters non-linearly.

where s_{ijt} represents the probability that patient type i will purchase drug j at time t. The own- and cross-price elasticities of demand are as follows:

$$\eta_{s_{jt},p_{kt}} = \frac{\partial s_{jt}}{\partial p_{kt}} \cdot \frac{p_{kt}}{s_{jt}} = \left\{ \begin{array}{l} \frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) dP^*_\tau(\tau) dP^*_D(D), \ j = k \\ -\frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dP^*_\tau(\tau) dP^*_D(D), \ j \neq k \end{array} \right\} \tag{6}$$

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.

V DATA

Data include market shares and prices, drug characteristics (physical or otherwise) and distribution on patient demographics, $P^*_D(D)$. Antidepressant sales data come from IMS Health Inc. and are complete unlike data on other therapeutic areas. These 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; values are deflated using the Consumer Price Index of the Bureau of Labor Statistics with 1980 as the base year.

Data run for 22 years from 1980 to 2001. In this period the market for antidepressants includes 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. IMS data on antidepressant market entry and exit and market segmentation were verified against data from the FDA[19]. The FDA also provided the data on therapeutic equivalence evaluations, on patents and on approvals. In the case of all antidepressant medications, the FDA qualifies all medications within the same molecule as therapeutically equivalent.

Two types of antidepressants were simultaneously introduced in 1959, MAOI and TCA, with the respective entry of Marplan and Tofranil in the market. 5HT2 Receptor Antagonists[20] appeared

[19] Food and Drug Administration (2009)

[20] Also referred to as New Generation (New Gen) antidepressants in the medical literature.

in 1982 with the advent of Desyrel and Prozac's entry in 1988 created the SSRIs. NDRI, SNRI and NaSSA were introduced in 1989, 1994 and 1996, respectively. 28 antidepressant innovations took place from 1980 to 2001, 12 of which were molecular innovations of different types, both branded and generic. These include side effects, half-life, dosage frequency, 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).

The physical drug characteristics are eight side effects, a drug's half-life and the average dosing frequency. Side effects included in the estimation are the drug's fatality rate, anticholinergic effects (dry mouth, blurred vision, urinary hesitancy, constipation), drowsiness, insomnia and/or agitation, orthostatic hypotension (abnormally low blood pressure), cardiac arrhythmias, gastrointestinal distress and weight gain of more than 6 kilograms. 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.

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,[21] 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. However, 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 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

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

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. 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.

VI ESTIMATION

I estimate demand parameters following closely the Generalized Method of Moments (GMM) approach presented in BLP (1995) and applying some of the suggestions in Nevo (2000). The general idea was, given the parameters, to draw ϵ_{ijt} from the assumed distribution for each patient in a sample and use these draws and parameters to construct simulated patient choices in antidepressants. Averaging these draws across simulated patients constructed a sample-average choice probability (market share). I then compared these simulated probabilities to the true probabilities in the data. At the market level the simulated probabilities from ns simulation draws, $\widehat{s}_j^{ns}(\theta)$, are unbiased estimates of the true market shares, s_j, that is $E\left[\widehat{s}_j^{ns}(\theta)\right] = s_j$. This implies $\widehat{s}_j^{ns}(\theta) = s_j + e_j$. The simulation error term, e_j, has zero covariance with all the data and as the number of draws becomes larger this error term tends to zero.

More precisely, to correct for price endogeneity I let $Z = [z_1, ..., z_N]$ be a set of instruments satisfying $E[Z'\omega(\theta^*)] = 0$, where ω is an error term defined as $\omega_{jt} = \delta_j(x, p, \widetilde{s}; \theta_2) - (x_j'\beta + \alpha p_{jt})$. As above, θ^* represents the true value of the model parameters. The GMM estimate, $\widehat{\theta} = \arg\min \omega(\theta)' Z A^{-1} Z' \omega(\theta)$, where A is a consistent estimate of $E[Z'\omega\omega'Z]$. The unobserved product characteristics are computed by equating the estimated shares, s, to the shares observed in the data, $\widetilde{s}$, and solving for mean utility levels. No closed form solution exists; the estimation of the market shares and their inversion to get the mean utility levels require numerical methods. Given a successful inversion, the ω_{jt} can be computed. In addition to observed side-effects, drug characteristics, x_j, include dummy variables for type, molecule and taste for branded drugs over generics to account for the observed segmentation in the antidepressant market. I, finally, solve for $\widehat{\theta}$ using a non-linear search.[22]

[22] See Nevo (2000) for a detailed technical description of the computation algorithm and discussion of alternative

i Instrumental Variables

To correct for the potential 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, will be 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 and/or type, the existence of me-too and generic competition. These instrumental variables approximate competition in the market. Additionally, I use advertising expenditure that varies both by drug j and annually to approximate firm costs. These variables are correlated with drug prices but are uncorrelated with the ξ_j.[23]

ii Patient Characteristics

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

methodologies.

[23] Following the suggestion in Nevo (2000b) drug dummy variables were also tested as instruments. The fact that unobserved drug characteristics include variables such as direct-to-consumer advertising, word-of-mouth and social trend effects implies that they might be influencing patient utility. Their inclusion, though, does not alter substitution patterns still driven by side effects and market segmentation. The main benefit of using drug dummy variables is that they account for characteristics that do not vary annually. Their inclusion as instruments allows the model to use all fixed information contained in the characteristics. It allows the separation between the exogenous variation in prices (due to competition) and endogenous variation (due to average unobserved valuation). Inclusion of drug dummies as instruments did not alter the results significantly.

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:

$$\sum_{i=1}^{ns} \frac{\exp\left(\delta_{jt} + \mu_{ijt}\right)}{1 + \sum_{k \in J_t} \exp\left(\delta_{kt} + \mu_{ikt}\right)} \tag{7}$$

Simulation draws are held constant as the parameters change otherwise changes in the objective function would be due to simulation changes.[24]

iii Parameters

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. However, some people view some side effects as positive. These positive valuations could either lead to positive coefficients or reduce the extend of the average negative valuation; for example, drowsiness, insomnia and agitation could be viewed as the target effects when the drug is taken for entertainment purposes. The frequency coefficient is expected to be negative since a higher per day dosage adversely affects patient compliance and therefore the drug's effectiveness. A 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 and prefer to feel the action of the drug sooner. It is ambiguous, therefore, what sign to expect for the coefficient of half-life.

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 (3)), by a minimum-distance procedure as in Nevo (2000). It also estimates standard deviations (λ's in equation (3)) which are estimates of the unobserved heterogeneity

[24] Note that I also have to account for simulation variance. This is nonlinear and increases, in a relative sense, as shares decrease.

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's in equation (3)). 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 (3). This restricts the overall price sensitivity to non-positive values.

VII RESULTS

i Descriptive Statistics

Summary statistics on entry are presented in Table III as well as summary data on observed drug characteristics. 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. However, gastrointestinal distress is still a very common side effect in SSRIs. 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 every patient. In other words, a side effect profile is used by individual patients to come up with individual decisions. Additionally, SSRIs seem to have a closer correlation of side effect occurrence among them than do drugs in other types.

INSERT TABLE III ABOUT HERE

TCAs have the highest average occurrence in five of the adverse side effects (anticholinergic effects, drowsiness, orthostatic hypotension, cardiac arrhythmias and weight gain) and the lowest in two adverse effects (insomnia/agitation and gastrointestinal distress). SSRIs rank exactly the opposite to TCAs on average in the same side effects. For instance, SSRIs have the highest occurrence of insomnia and agitation. In addition, they have the lowest average dosing frequency and the highest average half-life but that is mainly the effect of the high half-life of fluoxetine (Prozac). In contrast, 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. MAOIs rank somewhere in the middle for all side effects but the most important one, fatality rates.

They have, by far, the highest rate of occurrence. Since this characteristic is expected to have the highest marginal disutility, it is not surprising to observe that MAOIs occupy a tiny part of the market.

Before comparing estimated choice probabilities to observed market shares, it would be instructive to look at the trends of these shares over the time period of the available dataset. 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. It seems that they serve a quite steady portion of the depressed population that seeks medical treatment. MAOI revenue shares are also steady which means that their prices must have been rising in the same fashion as the prices of other antidepressants. One would think that prices should have fallen as a result of the decreasing quantity shares. However, if these drugs are geared towards a selected few patients that will always buy them, higher prices can be sustained.

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. The drop in shares is more pronounced for revenues than for quantities. The newer drugs come in with higher prices and reap the benefits innovation has on sustaining high prices. Both revenue and quantity shares for new generations keep rising until 1988 when Prozac and 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 both in revenues and quantities. 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 FIGURE I ABOUT HERE

Figure 1 shows the evolution of the market shares of the seven types of antidepressants over the twenty-two year period. One can see how new generations first and SSRIs after displaced TCAs from the market. Comparing these to revenues, TCA revenues were passed to these new types quicker than patients. SSRIs also decreased the market share of new generations as they came into the market. New generation revenue and quantity shares then remained relatively stable until 1995 when NDRI, SNRI and NasSSA started stealing both from TCAs and SSRIs.

INSERT FIGURE 2 ABOUT HERE

To examine the average trends in observed patient characteristics before the estimation results are presented, Figure 2 depicts per capita real GDP and compares its rising trend to the rising trend of prescription drug costs per patient and prescription drug costs per patient covered by insurance. Whereas costs covered by prescription drug plans follow a similar trend 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. Figure 3 shows how antidepressant prescription costs per patient overtake total per capita prescription costs in 1995 and the rate of growth of antidepressant prescription costs rises in 1997, when direct-to-consumer pharmaceutical advertising laws were relaxed in the U.S.

INSERT FIGURE 3 ABOUT HERE

ii Demand Estimation

Tables IV and V report the demand estimation results. Table IV reports the results for two different models: the simple logit model and the full RCL. The results of the former are given for an OLS version and an IV version correcting for price endogeneity in the first two columns of Table IV. The results of the latter are given both when no demographics are used and when demographic

information is incorporated. The third and fourth columns of Table IV show the mean effects, α's and β's, of the RCL models. The standard deviations from the addition of the unobserved characteristics and estimated parameters from the addition of demographics for the RCL model with demographics are then presented in Table V.

INSERT TABLE IV ABOUT HERE

In the simple logit model, the disutility for price was correctly detected and is high. I observe, however, that when moving from the OLS model to its IV counterpart, estimated price sensitivity about triples from -0.59 to -1.56. This implies that correction for price endogeneity is important and necessary. The signs of five out of eight estimates of side effect coefficients are negative as expected and statistically significant for both versions of the model. The magnitude of the fatality rate coefficient is largest in the IV version. Similar results are obtained for four other adverse side-effect coefficients: drowsiness, insomnia and agitation, cardiac arrhythmias and weight gain. Looking at the IV-Logit model, the disutility is highest for cardiac arrhythmias and lowest for weight gain. Two side effects, anti-cholinergic effects and orthostatic hypotension, have small positive signs, 75% of which are significant. The coefficient on half-life is only significant for the IV logit and both are small and positive and the average dosing frequency coefficient is obtained negative as expected and significant. The statistical and economic significance of this results is analyzed 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. This says that people are still swayed by perceptions that favor branded over generic drugs even when the two are therapeutically equivalent. It is possible that people have insurance plans that cover for branded drugs as well but more likely is the explanation that 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 brandness coefficient, similarly to the price coefficient, is allowed to vary with demographics in the RCL model. 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.

Three of the mean coefficients of the RCL model without demographics (third column of Table IV) are statistically insignificant. These include the coefficient on the preference for brandness which is later allowed to vary with observed patient demographics. Nevertheless, the coefficients for both price sensitivity and preference for brandness are economically significant. A price disutility close to unity and a positive preference for brandness were obtained. Estimated standard deviations from the addition of the unobserved characteristics were mostly significant which means that the normality assumption for the distribution of the unobserved characteristics is valid. However, the standard deviation for price sensitivity was reported low. These results reinforce the need for the use of demographics to correctly model patient decisions when it comes to the choice of characteristics such as price and brandness.[25]

INSERT TABLE V ABOUT HERE

Table V 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. 30 of the 32 parameters are statistically significant. 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 GI distress. 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 mostly statistically and economically significant. The last three columns present the effects of demographics (observed patient characteristics) on the mean coefficients. These estimates are all statistically and economically significant.

Apart from the anticholinergic effects, all adverse side effects have negative mean coefficients and relatively large and significant standard deviations. The negative coefficients suggest that the average patient gets more disutility the more these side-effects occur. There is no immediate

[25] Cleanthous (2003) estimates five different models (the simple logit model, three nested multinomial logit models and the random coefficient model) with and without instrumental variables and contrasts their results. 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.

explanation behind the anticholinergic effects estimate, which means that on average patients prefer drugs with this characteristic because they dislike it the least. A possible explanation for this result is that most depressed people place high priority on other side effects and ignore side effects such as dry mouth so that on average their choices predict a small positive utility derived from the occurrence of these side effects. The estimated standard deviations are estimates of the random patient heterogeneity around these means. Since many 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.

Half-life has a negative coefficient, though it is small and has a relatively high variance. 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 fatality rate, orthostatic hypotension, cardiac arrhythmias, weight gain and administration frequency all have large negative values with relatively small standard deviations. The big negative value signifies the high disutility obtained from occurrence of these side effects; the relatively low standard deviations suggest that occurrence of these effects does not offer positive utility to any patient. The case is different for drowsiness, insomnia and agitation. Though large negative coefficients are obtained implying high disutility for the average consumer, it is interesting to see that relatively large estimated standard deviations imply that some patients obtain positive utility from occurrence of these characteristics. The positive and statistically significant coefficient of anticholinergic effects, as explained before, may be due to the fact that patients are not really taking this less adverse side-effect into consideration when choosing the best medication for them. The relatively large standard deviation, though statistically insignificant, is economically significant and shows that for many patients anticholinergic effects do not provide positive utility.

The parameters of most importance in this final model are the coefficients on the preference of brandness and price sensitivity. As presented in equation (3), 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 just above unity, -1.108. This is the disutility obtained by the average patient. The relatively small estimated standard deviation suggests that most of the heterogeneity (85%) 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.[26] 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.[27]

INSERT TABLE VI ABOUT HERE

The estimates of the coefficients on the preference for brandness 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 (80%) in patient preference for brandness is explained by the observed demographics. Tables VI and VII show the combined effect of demographics on price and brand sensitivity, respectively. 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. Poorer

[26] 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.

[27] This is discussed in the next section.

patients without prescription drug coverage are the most sensitive to price changes and have the lowest preference for brandness. Moving from the latter extreme case to the former, the off diagonal results show that poorer but insured patients are less price sensitive and have a higher preference for brandness than do wealthier, uninsured patients.

INSERT TABLE VII ABOUT HERE

Table VIII presents own- and cross-price elasticities of demand for selected antidepressants. These are weighted averages for the years 2000 and 2001.[28] The reported cross-price elasticities are averaged over drugs of the same type. The selection includes drugs of two TCA molecules and one NDRI that have experienced generic introduction, all SSRIs and two MAOIs. The selected antidepressants help show the superiority of the RCL model over the other estimated models in describing substitution patterns. 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.

INSERT TABLE VIII ABOUT HERE

To show the superiority of the full model to the other estimated models I, first, test and reject the joint hypothesis that all the non-linear parameters are zero.[29] I then follow the suggestion by Nevo (2000) and compute the variation of cross-price elasticities in the various columns by dividing the maximum elasticity in a column by the minimum. The ratio varies from 5 to 25 in the sample Table VIII and from 4 to 45 in the complete table of all cross-price elasticities. The lower the ratio, the more characteristics need to be added in the analysis to help overcome the logit-imposed restrictions on substitution patterns.

[28] Simple averages when the drug exists in both years, weighted averages when the drug exists only one of the years.
[29] Hausman and McFadden (1984)

VIII WELFARE IMPLICATIONS

i Evaluating Patient Welfare

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.[30] 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.

The antidepressant market is a differentiated products market. As a result, when calculating patient welfare due to antidepressant innovation a need arises for estimating the degree to which the new antidepressant replaces older antidepressants, the incremental value gained by patients who switch and the competitive impact of innovation on the market prices of existing antidepressants. For example, in 1988, Prozac entered the market. This requires estimation of a demand system that will allow me to calculate consumer surplus both when Prozac was in and out of the market. With such a demand system I will be able to capture the impact of Prozac as it diffused into the market as well as its immediate entry effect.

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 (5) 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.

[30] Similar analyses in Trajtenberg (1990) and Ellickson, Stern and Trajtenberg (2001).

Summation of all these shares obtains the welfare at every time period, t, W_t:

$$W_t = \sum_{j=0}^{J_t} W_{jt}. \tag{9}$$

Therefore, incremental patient welfare due to a innovation at time t, call it INN_t, is the difference in patient welfare before and after the innovation, that is $\Delta W = W_t - W_{t-1}$, ceteris paribus. However, at the time of a specific drug introduction, welfare gains might arise due to concurrent introduction of other innovations, changes in the prices or other observed drug characteristics of existing drugs, and because of market withdrawal of existing drugs. INN_t captures all these gains. In other words, it implies that the entire change in patient welfare at time t is attributed to the innovation in question. In the extreme case that nothing else is happening in the economy apart from the innovation of drug j, INN_t is the incremental patient welfare due to that innovation, that is, $INN_t = INN_{jt}$. In all other cases, INN_t is not accurate and should just represent an upper bound for the innovation's contribution to patient welfare, call it $\overline{INN}_{jt}$. More precisely, $\overline{INN}_{jt}$ represents the maximum gain in patient welfare that could have arisen from the innovation of drug j at time t. To have a more accurate measure of INN_{jt}, I recalculate the incremental patient welfare due to innovation after removing drug j from the existing choice set at the time of innovation, keeping prices and observed drug characteristics for all other drugs the same. This latter measure allows calculation of a lower bound for the innovation's contribution to patient welfare, call it $\underline{INN}_{jt}$. More precisely, $\underline{INN}_{jt}$ represents the minimum gain in patient welfare caused by the innovation of drug j at time t. This, in fact, is a better measure of patient welfare as it represents an increment in welfare due only to the advent of drug j. It, therefore, avoids having to distinguish between multiple innovations occurring at the same time period does not attribute welfare gains due to other reasons to the innovation of drug j.[31]

ii Moral Hazard

As already explained, the demand estimation results have uncovered an important issue in this literature, the moral hazard issue. Patients insured against prescription drug expenditures are willing to pay higher prices for their medications than they would be willing to pay when uninsured.

[31] The implicit assumption from this formulation is that production of other drugs might not have been the same had the investigated innovation not taken place.

This is reflected in the very low estimated marginal disutility of price that results from the presence of prescription drug coverage. 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 private willingness-to-pay. The difference in the two is attributed to moral hazard.

INSERT TABLE IX ABOUT HERE

Tables IX and X present the welfare estimates ($\underline{INN}_{jt}\ \forall j$) of all 28 antidepressant innovations that took place between 1980 and 2001.[32] Drugs in bold in the table represent introduction of an antidepressant type and drugs in italics represent molecular introductions. Estimated patient welfare for patients with and without insurance is presented in Table IX both in total constant 1980 dollars and in per unit 1980 dollars. The latter represents the 'true' patient's willingness-to-pay over the price charged. Looking at the first column of welfare estimates, the magnitude of the surplus is enormous, even though I report the lower bound of the gains from innovation, $\underline{INN}$, described above. When calculating the patient surplus per unit in the next column I get values that are enormously bigger than the actual price paid. This is a problem that arises when estimating welfare gains based on a discrete-choice model of demand with errors distributed extreme value. An additional explanation offered by this paper is 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 in the next two columns that exclude prescription drug insurance.

With the exclusion of insurance, estimated patient gains are more insightful.[33] Finding surplus per unit (average daily dosage) in the next column, shows a patient's willingness-to-pay above the price of the product. The last column of Table IX shows the excess willingness-to-pay accrued annually. 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. Comparing this to the average annual cost of

[32] Note that values for $\overline{INN}_{jt}$ are huge as they incorporate all welfare gains at time t, and proved uninformative. They are, therefore, not provided in this paper.

[33] The total patient surplus from 19 of the 28 innovations is of the same or lower order of magnitude as the drug revenues.

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, 1.7 times more a year for a Remeron treatment and so on. Moving to Table X, the last column depicts the ratio of this per unit patient surplus on price.[35] A value of 0.50 in this column, for instance, says that a patient would be willing to pay one and a half times as much for a drug than she is currently paying. For Prozac, a patient is, therefore, willing to pay almost 31 times as much, whereas for 1.5 times.

INSERT TABLE X ABOUT HERE

Relative gains help evaluate the importance and success of different innovations in the antidepressant market. To help analyze relative gains, the first four columns of Table X present rankings of the four columns of patient welfare gains in Table IX. The first column shows that the innovation of Prozac [fluoxetine], which was also the first drug in a new type of antidepressants (SSRI), offered the highest gains in patient surplus. Prozac ranks first in all the four different estimates. Similarly, the innovation of Remeron [mirtazapine], which marked the introduction of the newst type of antidepressants (NaSSA), offered the second highest gains in patient surplus. The introduction of NDRI with the innovation of Wellbutrin [bupropion] aslo fairs well whereas the introduction of new generation antidepressants with the advent of Desyrel [trazodone] and SNRI with Effexor [venlafaxine] were not as successful. Their overall patient surplus rank in the middle of all antidepressant innovations. Desyrel's and Effexor's per unit valuations and when taking insurance into account diminish the relative importance of their innovation. Moving from the valuation of patient surplus with insurance (first two columns) to its valuation without insurance (last two columns), a dramatic increase in the relative importance of generic innovations is observed. This is not surprising. Demand estimation has shown that insured patients tend to be less price sensitive and have a higher preference for brandness than uninsured patients. Observe that the relative importance of branded drugs between themselves stays mostly the same as does the relative importance of generic drugs when only compared to other generics.

These results on welfare gains and patient willingness-to-pay are useful to pharmaceutical companies. On one hand, there is now evidence of relative patient preference for different drug characteristics. Research and development departments of pharmaceutical companies can use this

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

[35] This is the per unit surplus divided by the drug's price for that unit (average daily dosage).

information to adapt the characteristics of new innovations, to try and meet consumer needs. On the other hand, once a new drug is developed, a pharmaceutical company can use the calculated willingness-to-pay to price it. Moreover, this can be done even before developing a drug. The possible benefits of a hypothetical drug can be evaluated using the welfare gains presented in this paper. The results of this section may also be used in conjunction with other economic studies to help solve important public policy questions and address pharmaceutical industry concerns. For instance, comparing these results to research and development costs provide cost-benefit analysis of new drug introduction. This is useful both for the government to evaluate the fairness of pharmaceutical pricing practices, but also for pharmaceutical companies to evaluate the effectiveness of their existing, upcoming and hypothetical innovations.

IX CONCLUSION

In this paper, I formulate an empirical methodology that quantifies patient welfare benefits from pharmaceutical innovation in the U.S. antidepressant market. The paper employs an original dataset that consists of annual observations on prices, quantities 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. While evaluating pharmaceutical innovation in antidepressants, I uncover and address the moral hazard issue that arises due to the existence of pharmaceutical insurance coverage. 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. 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 and when patients are insured against prescription drug expenditures. Moreover, patients demonstrate a high preference for branded drugs. The wealthier the patients and the more insurance coverage they have, the higher the preference.

Welfare estimation involves the calculation of an upper bound for incremental patient surplus when all the gains obtained are attributed to the innovation in question and a lower bound when

the innovative drug is excluded from the choice set at the time of innovation. The latter provides those gains to innovation attributed solely to a new product introduction. I obtain large gains for patients, particularly when insurance coverage exists. Relative gains help evaluate the importance of different innovations in the antidepressant market; the innovation of Prozac, which was also the first drug in a new category of antidepressants, offered the highest gains in patient surplus.

An important extension this paper uses monthly pharmaceutical data from 1996 to 2008 to incorporate patient-level information on prescription drug insurance. This is a major improvement in the model as inclusion of disaggregate data will more accurately address the moral hazard issue. I break down the IMS data by the various distribution channels (for example, non-federal hospitals, private pharmacies and health maintenance organizations) which will aid in explaining the role of institutions in the choice of pharmacological treatment. Survey data that match patient choices directly to patient income and insurance information will then provide even more precise estimates on patient willingness-to-pay and welfare.

Another extension[36] incorporates a more detailed analysis of marketing efforts in the antidepressant market and informative advertising, specifically. When adding informative advertising to demand estimation there is an upset in the estimated parameters for drug characteristics as advertising sways the choices of both physicians and patients towards the more advertised drugs. More importantly, price becomes more inelastic as patients become more certain for their choices through direct-to-consumer advertising and preference for 'brandness' varies in direction and significance. Finally, this paper has incited work[37] that combines concurrent demand and supply evaluations of innovation to identify strategic decisions in the pharmaceutical industry regarding the degree of innovation. In particular, pharmaceutical firms are modelled to take into account the large innovation uncertainty that characterizes the industry, employ promotional activity to sway the decisions of consumers and physicians, and the huge fixed costs associated with pharmaceutical research and development activity that vary with innovativeness.

X REFERENCES

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

[36] Cleanthous (2009)

[37] Cleanthous et al (2010)

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., Iain M. Cockburn, and Zvi Griliches. 1996. "Pharmaceutical Innovations and Market Dynamics: Tracking the Effects on Price indices for Antidepressant Drugs," *Brookings Papers on Economic Activity. Microeconomics*: 133-188.

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.

Berndt, Ernst R. et al. 2002. "The Medical Treatment of Depression, 1991-1996: Productive Inefficiency, Expected Outcome Variations, and Price indices," *Journal of Health Economics*, 21(3): 373-96.

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.

Bresnahan, Timothy F. and Robert J. Gordon (eds), 1997. *The Economics of New Goods*, University of Chicago Press, Chicago.

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

Cleanthous, Paris, 2009. "Pharmaceutical Informative Advertising," mimeo University of Cyprus.

Cleanthous, Paris, Wonjoon Kim and Dick Wittink, 2009. "Innovate or Imitate? Firms' Strategic Innovation Decision in the Pharmaceutical Industry," mimeo KAIST and 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.

Dixit, A., and J.E. Stiglitz, 1997. "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 70: 312-326.

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.

Ellison, Sara Fisher, Iain Cockburn, Zvi Griliches and Jerry Hausman, 1997. "Characteristics of Demand for Pharmaceutical Products: an Examination of Four Cephalosporins," *Rand Journal of Economics*, 28(3): 426-446.

Ellison, Sara Fisher and Christopher M. Snyder, 2010. "Countervailing Power in Wholesale Pharmaceuticals," Journal of Industrial Economics, 58(1): 32-53.

Goldberg, Pinelopi Koujianou, 1995. "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica*, 63(4): 891-951.

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, 1996. "Valuation of New Goods Under Perfect and Imperfect Competition," in Timothy Bresnahan and Robert J. Gordon (eds), 1997. *The Economics of New Goods*, University of Chicago Press, Chicago.

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.

Lichtenberg, Frank R., 1996. "Do (More and Better) Drugs Keep People Out of Hospitals?," American Economic Review, 86(2): 384-8.

Lichtenberg, Frank R., 2000. "The Effect of Pharmaceutical Utilization and Innovation On Hospitalization and Mortality," in Bart van Ark, Simon Kuipers, and Gerard Kuper (eds), 2000. *Productivity, Technology, and Economic Growth*, Ch. 11: 317-344. Kluwer Academic Publishers, Boston.

Lichtenberg, Frank R., 2001a. "Are the Benefits of Newer Drugs Worth Their Cost? Evidence from the 1996 MEPS," Health Affairs 20(5): 241-51.

Lichtenberg, Frank R., 2001b. "The Allocation of Publicly Funded Biomedical Research," in Ernst Berndt and David Cutler (eds.), 2001. *Medical Care Output and Productivity, Studies in Income and Wealth,* Vol. 62: 565-89. University of Chicago Press.

Lichtenberg, Frank R., 2003. "Pharmaceutical Innovation, Mortality Reduction and Economic Growth," in K. M. Murphy and R. H. Topel (eds), 2003. *Measuring the Gains from Medical Research. An Economic Approach*, Ch. 3: 74-109, University of Chicago Press, Chicago & London.

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

Murphy, Kevin M. and Robert H. Topel, 2003. *Measuring the Gains from Medical Research. An Economic Approach*, University of Chicago Press, Chicago & London.

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

Nevo, Aviv, 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2): 307-42.

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.

Sudhir, K., 2001. "Competitive Pricing Behavior in the Auto Market: A Structural Analysis," *Marketing Science,* 20(1): 42-60.

Suslow, Valerie. 1996. "Measuring Quality Change in Pharmaceutical Markets: Hedonic Price Indexes for Anti-Ulcer Drugs," in R. Helms (ed.), 1996. *Competitive Strategies in the Pharmaceutical Industry*, American Enterprise Institute, Washington, D.C.: 49-72.

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.

Wosińska, Marta, 2005. "Direct-to-Consumer Advertising and Drug Therapy Compliance," *Journal of Marketing Research*, 42(3) :323–332

TABLE I
ENTRY IN THE ANTIDEPRESSANT MARKET

	Originator Brand		Secondary Brand		Generic
Generic Name	Name	Entry	Name	Entry	Entry
MONOAMINE OXIDASE INHIBITORS (MAOI)					
Isocarboxazid	Marplan®	1959	-	-	None
Phenelzine	Nardil®	1959	-	-	None
Tranylcypromine	Parnate®	1961	-	-	None
TRICYCLIC ANTIDEPRESSANTS & RELATED COMPOUNDS (TCA)					
Imipramine	Tofranil®	1958	Janimine®	1975	1975
Amitriptyline	Elavil®	1961	Endep®	1975	1977
Nortriptyline	Aventyl®	1963	Pamelor®	1977	1992
Protriptyline	Vivactil®	1967	-	-	1996
Doxepin	Sinequan®	1969	Adapin®	1973	1986
Trimipramine	Surmontil®	1969	-	-	1988
Desipramine	Pertofrane®	1971	Norpramin®	1975	1987
Imipramine Pamoate	Tofranil PM®	1973	-	-	2010
Amoxapine	Asendin®	1980	-	-	1989
Maprotiline	Ludiomil®	1981	-	-	1988
Clomipramine	Anafranil®	1990	-	-	1996
5HT2-RECEPTOR ANTAGONISTS (NewGen)					
Trazodone	Desyrel®	1981	-	-	1986
Nefazodone	Serzone®	1995	-	-	2003
Buspirone	Buspar®	1987	-	-	2000
SELECTIVE SEROTONIN REUPTAKE INHIBITORS (SSRI)					
Fluoxetine	Prozac®	1988	Sarafem™ Prozac Weekly®	2000 2001	2002
Sertraline	Zoloft™	1992	-	-	2006
Paroxetine	Paxil™	1993	Paxil CR™	2003	2003
Fluvoxamine	Luvox®	1994	-	-	2000
Citalopram	Celexa™	1998	-	-	2004
Escitalopram	Lexapro®	2003	-	-	-
NOREPINEPHRINE AND DOPAMINE REUPTAKE INHIBITORS (NDRI)					
Bupropion	Wellbutrin®	1989	Wellbutrin SR® Zyban® Wellbutrin XL®	1996 1997 2003	1999
SEROTONIN and NORADRENALINE REUPTAKE INHIBITORS (SNRI)					
Venlafaxine	Effexor®	1994	Effexor-XR®	1997	2006
Duloxetine	Cymbalta®	2004	-	-	-
Desvenlafaxine	Pristiq®	2009	-	-	-
Milnacipran	Savella™	2010	-	-	-
NORADRENERGIC AND SPECIFIC SEROTONERGIC ANTIDEPRESSANTS (NaSSA)					
Mirtazapine	Remeron®	1996	Remeron Soltab®	2001	2003

Notes: Pharmaceutical data for entry into the American market come from IMS Health, Inc. and the Food & Drug Administration (Annual). Data run upto the beginning of 2010.

TABLE II
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 III

ENTRY & AVERAGE CHARACTERISTICS BY TYPE OF ANTIDEPRESSANTS

Drug Type	No. Molecules	No. Brands	No. Generics	Moleculed Entry	Branded Entry	Generic Entry	Year of First Entry in the Market	Average Annual Revenue Shares (%)			Average Annual Quantity Shares (%)				Average Dosing	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 IV
PARAMETERS FROM DEMAND ESTIMATION ACROSS MODELS

Variable	OLS Logit	IV Logit	Random Coefficients (no demographics)	Random Coefficients (demographics)
	Price Coefficients (α's)			
Price	-0.591**	-1.563**	-1.095*	-1.108**
	(0.113)	(0.286)	(1.045)	(0.253)
	Coefficients of Characteristics (β's)			
Half-Life	0.002	0.038*	-0.673*	-0.343*
	(0.030)	(0.034)	(0.505)	(0.289)
Fatality	-1.672**	-2.244**	-5.078**	-5.159**
	(0.189)	(0.252)	(1.719)	(1.693)
Anti-Cholinergic	0.578*	0.449**	0.021*	0.144**
	(0.462)	(0.175)	(0.019)	(0.013)
Drowsiness	-1.017**	-0.899**	-3.402**	-2.769**
	(0.159)	(0.171)	(0.360)	(1.021)
Insomnia/ Agitation	-0.819**	-0.647**	-7.701*	-7.768**
	(0.143)	(0.158)	(4.118)	(1.374)
Orthostatic Hypotension	0.490	0.590**	-2.421*	-1.266**
	(1.119)	(0.129)	(2.322)	(0.502)
Cardiac Arrhythmias	-1.385**	-1.726**	-3.607**	-1.942**
	(0.162)	(0.194)	(1.094)	(0.184)
GI Distress	0.181*	-0.077	1.796	0.131
	(0.098)	(0.125)	(2.785)	(1.004)
Weight Gain	-0.427**	-0.312**	-2.925	-4.528**
	(0.126)	(0.137)	(3.844)	(1.241)
Average Dosing Frequency	-0.439**	-0.557**	-5.436*	-3.183**
	(0.129)	(0.140)	(3.221)	(1.287)
Brand Dummy	1.360**	0.614**	1.529	3.038**
	(0.203)	(0.294)	(2.747)	(1.540)
Constant	1.589**	3.382**	10.332**	60.790**
	(0.684)	(0.868)	(1.505)	(3.893)

Notes: Standard errors are in parentheses. * indicates t-statistic > 1 and ** indicates t-statistic > 2. Regressions included type dummies. Number of observations = 656.

TABLE V
PARAMETERS FROM RANDOM COEFFICIENT MODEL

	Means	Standard Deviations	Interactions with Demographics		
Variable	(α & β's)	(λ's)	Income	Income Sqrd	Prescription Insurance
Price	-1.108**	0.143**	0.103**	-0.064**	0.811**
	(0.025)	(0.004)	(0.014)	(0.013)	(0.076)
Half-Life	-0.343*	1.095**	-	-	-
	(0.289)	(0.001)			
Fatality	-5.159**	0.735**	-	-	-
	(1.693)	(0.098)			
Anti-Cholinergic	0.144**	0.283	-	-	-
	(0.013)	(0.292)			
Drowsiness	-2.769**	1.584**	-	-	-
	(1.021)	(0.220)			
Insomnia/ Agitation	-7.768**	4.896**	-	-	-
	(1.374)	(0.998)			
Orthostatic Hypotension	-1.266**	0.350*	-	-	-
	(0.502)	(0.295)			
Cardiac Arrhythmias	-1.942**	0.335**	-	-	-
	(0.184)	(0.090)			
GI Distress	0.131	0.547**	-	-	-
	(1.004)	(0.000)			
Weight Gain	-4.528**	0.683**	-	-	-
	(1.241)	(0.004)			
Average Dosing Frequency	-3.183**	1.424**	-	-	-
	(1.287)	(0.010)			
Brand Dummy	2.038**	0.766**	1.110**	0.247**	1.832*
	(0.540)	(0.296)	(0.099)	(0.005)	(1.141)
Constant	60.790**	2.229*	-	-	-
	(3.893)	(1.692)			

Notes: Standard errors are in parentheses. * indicates t-statistic > 1 and ** indicates t-statistic > 2.
Regressions included type dummies. Number of observations = 656.

TABLE VI
RANDOM COEFFICIENT LOGIT ELASTICITIES: PRICE

	Price Sensitivities	
	High Income	Low Income
With Full Insurance	-0.165*	-0.297**
	(0.138)	(0.101)
Without Any Insurance	-0.889**	-1.108**
	(0.012)	(0.025)

Notes: Standard errors are in parentheses. * indicates t-statistic > 1 and ** indicates t-statistic > 2.

TABLE VII
RANDOM COEFFICIENT LOGIT ELASTICITIES: 'BRANDNESS'

	Brand Sensitivities	
	High Income	Low Income
With Full Insurance	4.759**	3.870**
	(1.860)	(1.681)
Without Any Insurance	3.082**	2.038**
	(0.284)	(0.540)

Notes: Standard errors are in parentheses. * indicates t-statistic > 1 and ** indicates t-statistic > 2.

TABLE VIII
OWN AND CROSS-PRICE ELASTICITIES OF DEMAND FOR SELECTED ANTIDEPRESSANTS (2000-2001)

Type	Molecule	Drug	Drug	Cross-Price Elasticities					
				OPE	TCA	NewGen	SSRI	MAOI	ALL
					26	10	8	3	47
1	2	2	Nardil	-0.313	0.552	0.447	0.308	0.752	0.501
1	3	3	Parnate	-0.436	0.616	0.494	0.343	0.780	0.554
2	4	4	Elavil	-0.209	0.355	0.437	0.302	0.402	0.366
2	4	5	Amitriptyline [G]	-0.456	0.381	0.179	0.116	0.102	0.275
2	5	6	Amoxapine [G]	-0.220	0.275	0.222	0.166	0.137	0.236
2	5	7	Asendin	-0.089	0.707	0.543	0.393	0.496	0.605
4	17	33	Celexa	-0.116	0.133	0.329	0.515	0.137	0.240
4	18	34	Prozac	-0.188	0.151	0.301	0.550	0.375	0.265
4	18	35	Prozac Weekly	-0.024	0.149	0.308	0.514	0.310	0.255
4	18	36	Sarafem	-0.158	0.156	0.298	0.573	0.384	0.272
4	19	37	Fluvoxamine [G]	-0.536	0.028	0.113	0.200	0.025	0.075
4	19	38	Luvox	-0.094	0.276	0.450	0.496	0.188	0.345
4	20	39	Paxil	-0.045	0.309	0.456	0.330	0.200	0.337
4	21	40	Zoloft	-0.065	0.407	0.415	0.312	0.218	0.380
5	22	41	Bupropion [G]	-0.192	0.049	0.087	0.131	0.039	0.070
5	22	42	Wellbutrin	-0.352	0.223	0.405	0.329	0.165	0.276
5	22	43	Wellbutrin SR	-0.157	0.254	0.405	0.308	0.181	0.291

Notes: Average of elasticities for the last two years of the dataset: 2000 and 2001. 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. Types 1-5 stand for MAOI , TCA, NewGen, SSRI and NDRI, respectively. The number of drugs in each type of antidepressants is displayed under each type heading. [G] indicates a generic drug.

TABLE IX

PATIENT WELFARE DUE TO INNOVATION IN ANTIDEPRESSANTS 1981 - 2001

Type	Molecule	Drug No.	Drug Name	Generic Name	Entry	Patient Surplus				
						With Insurance		No Insurance		
						Total (million $)	Per Unit ($)	Total (thousand $)	Per Unit ($)	Annual Prescription ($)
2	*11*	*22*	*Ludiomil*	*Maprotiline*	*1981*	943	49	77	0.00	1
3	***16***	***31***	***Desyrel***	***Trazodone***	***1982***	7,391	661	352	0.03	12
2	8	15	Generic	Doxepin	1986	902	72	172	0.01	5
3	16	32	Generic	Trazodone	1986	23	30	5	0.01	3
2	7	11	Generic	Desipramine	1987	42	22	30	0.02	6
2	11	21	Generic	Maprotiline	1988	1,360	1,569	348	0.40	147
2	14	28	Generic	Trimipramine	1988	820	2,284	166	0.46	169
4	***18***	***34***	***Prozac***	***Fluoxetine***	***1988***	4,477,592	54,894	1,995,377	24.46	8,929
2	5	8	Generic	Amoxapine	1989	6,396	27,930	1,353	5.91	2,157
5	***22***	***42***	***Wellbutrin***	***Bupropion***	***1989***	19,073	4,288	1,430	0.32	117
2	*6*	*9*	*Anafranil*	*Clomipramine*	*1990*	5,127	394	702	0.05	20
2	12	24	Generic	Nortriptyline	1992	8,271	424	4,768	0.24	89
4	*21*	*40*	*Zoloft*	*Sertraline*	*1992*	773	7	36	0.00	0
4	*20*	*39*	*Paxil*	*Paroxetine*	*1993*	7,952	151	734	0.01	5
6	***23***	***44***	***Effexor***	***Venlafaxine***	***1994***	5,579	139	297	0.01	3
4	*19*	*38*	*Luvox*	*Fluvoxamine*	*1994*	4,100	16,120	461	1.81	661
3	*15*	*30*	*Serzone*	*Nefazodone*	*1995*	1,068	37	212	0.01	3
2	6	10	Generic	Clomipramine	1996	7,480	16,082	2,086	4.48	1,637
2	13	26	Generic	Protriptyline	1996	5,452	22,016	1,082	4.37	1,595
5	22	43	Wellbutrin SR	Bupropion	1996	11,757	12,881	1,362	1.49	545
7	***24***	***46***	***Remeron***	***Mirtazapine***	***1996***	58,491	47,182	7,264	5.86	2,139
6	23	45	Effexor-XR	Venlafaxine	1997	20,861	4,990	1,436	0.34	125
4	*17*	*33*	*Celexa*	*Citalopram*	*1998*	9,982	774	1,356	0.11	38
5	22	41	Generic	Bupropion	1999	2,123	2,282	459	0.49	180
4	18	36	Sarafem	Fluoxetine	2000	37,998	9,481	3,353	0.84	305
4	19	37	Generic	Fluvoxamine	2000	45,325	40,902	8,669	7.82	2,856
4	18	35	Prozac Weekly	Fluoxetine	2001	52,657	32,936	6,297	3.94	1,438
7	24	47	Remeron Soltab	Mirtazapine	2001	26,072	11,115	3,002	1.28	467

Notes: Tables shows value of innovation for all 28 innovations in the data. Types 1-7 stand for MAOI, TCA, NewGen, SSRI, NDRIs, SNRI and NaSSA respectively. Molecular (active ingredient) innovation is shown in italics. 'Type' (mechanism of action) innovation is shown in bold.

TABLE X

WELFARE IMPLICATIONS OF ANTIDEPRESSANT INNOVATION 1981 - 2001

Type	Molecule	Drug No.	Drug Name	Generic Name	Entry	Patient Surplus Rankings				Per Unit Patient Surplus to Price Ratio
						With Insurance		No Insurance		
						Total	Per Unit	Total	Per Unit	
4	***18***	***34***	***Prozac***	***Fluoxetine***	***1988***	1	1	1	1	30.68
7	***24***	***46***	***Remeron***	***Mirtazapine***	***1996***	2	2	3	4	1.49
4	18	35	Prozac Weekly	Fluoxetine	2001	3	4	4	7	0.40
4	19	37	Generic	Fluvoxamine	2000	4	3	2	2	3.89
4	18	36	Sarafem	Fluoxetine	2000	5	11	6	11	0.61
7	24	47	Remeron Soltab	Mirtazapine	2001	6	10	7	10	0.33
6	23	45	Effexor-XR	Venlafaxine	1997	7	12	9	15	0.15
5	***22***	***42***	***Wellbutrin***	***Bupropion***	***1989***	8	13	10	16	0.52
5	22	43	Wellbutrin SR	Bupropion	1996	9	9	11	9	0.74
4	*17*	*33*	*Celexa*	*Citalopram*	*1998*	10	17	12	18	0.06
2	12	24	Generic	Nortriptyline	1992	11	19	5	17	0.22
4	*20*	*39*	*Paxil*	*Paroxetine*	*1993*	12	21	15	22	0.01
2	6	10	Generic	Clomipramine	1996	13	8	8	5	4.02
3	***16***	***31***	***Desyrel***	***Trazodone***	***1982***	14	18	19	20	0.04
2	5	8	Generic	Amoxapine	1989	15	5	13	3	4.37
6	***23***	***44***	***Effexor***	***Venlafaxine***	***1994***	16	22	21	25	0.01
2	13	26	Generic	Protriptyline	1996	17	6	14	6	5.05
2	*6*	*9*	*Anafranil*	*Clomipramine*	*1990*	18	20	16	19	0.06
4	*19*	*38*	*Luvox*	*Fluvoxamine*	*1994*	19	7	17	8	1.59
5	22	41	Generic	Bupropion	1999	20	15	18	12	0.49
2	11	21	Generic	Maprotiline	1988	21	16	20	14	1.06
3	*15*	*30*	*Serzone*	*Nefazodone*	*1995*	22	25	22	24	0.01
2	*11*	*22*	*Ludiomil*	*Maprotiline*	*1981*	23	24	25	27	0.01
2	8	15	Generic	Doxepin	1986	24	23	23	23	0.11
2	14	28	Generic	Trimipramine	1988	25	14	24	13	1.20
4	*21*	*40*	*Zoloft*	*Sertraline*	*1992*	26	28	26	28	0.00
2	7	11	Generic	Desipramine	1987	27	27	27	21	0.02
3	16	32	Generic	Trazodone	1986	28	26	28	26	0.01

Notes: 'Per Unit Patient Surplus to Price Ratio' is the ratio of the private marginal willingness-to-pay over the price of the drug (Table VIII). Types 1-7 stand for MAOI, TCA, NewGen, SSRI, NDRIs, SNRI and NaSSA respectively. Molecular (active ingredient) innovation is shown in italics. 'Type' (mechanism of action) innovation is shown in bold.

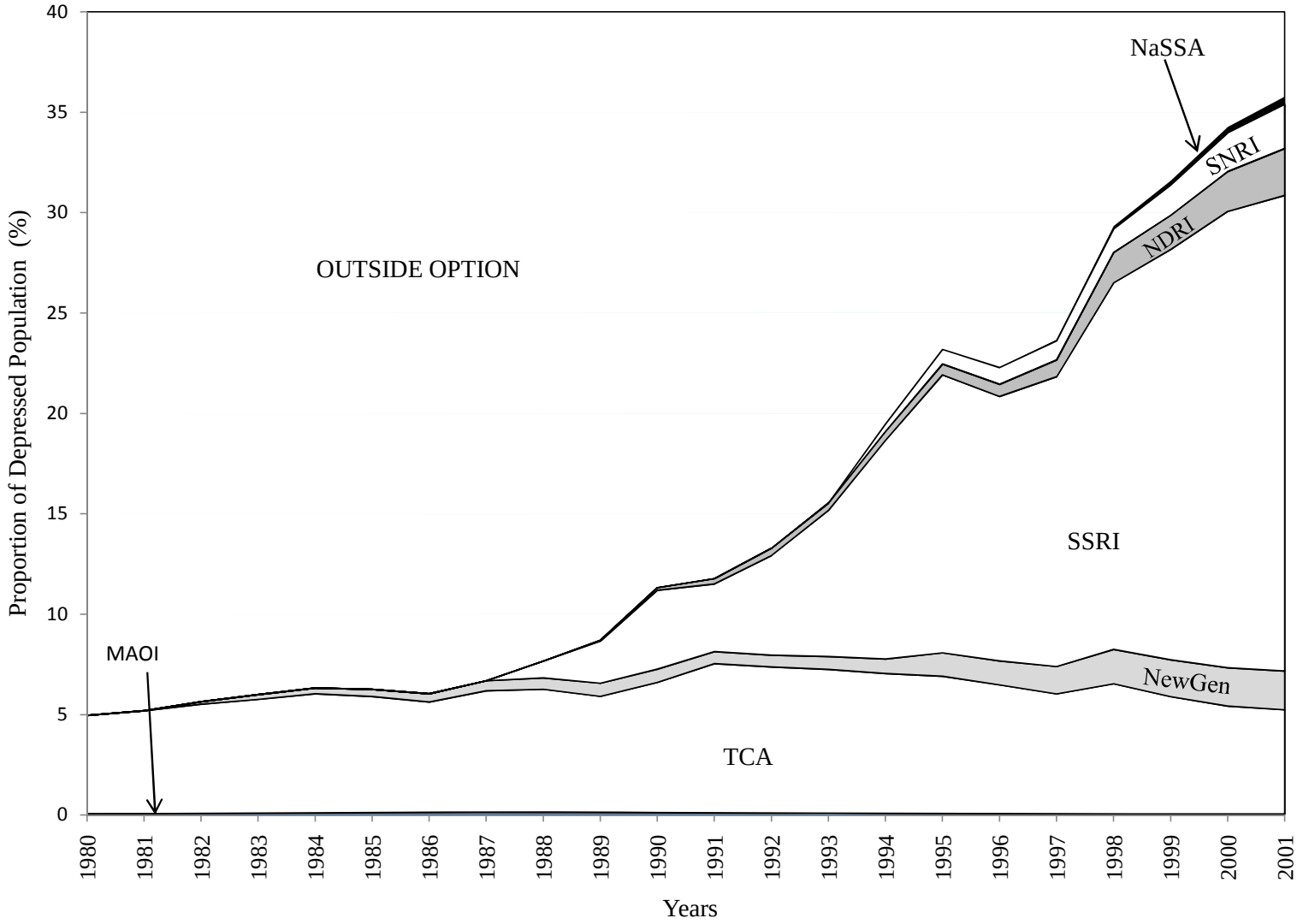

Figure 1
Market Share Evolution by Type

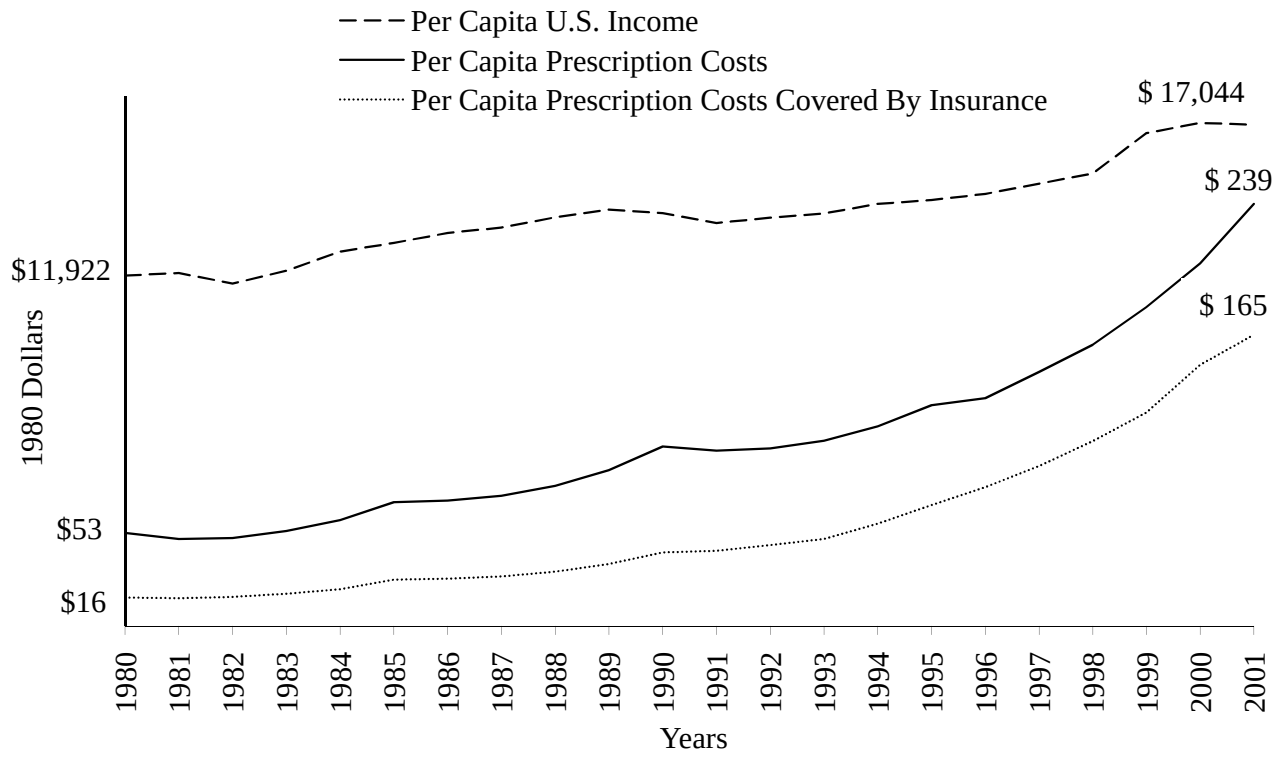

Figure 2
Per Capita U.S. Income & Prescription Costs

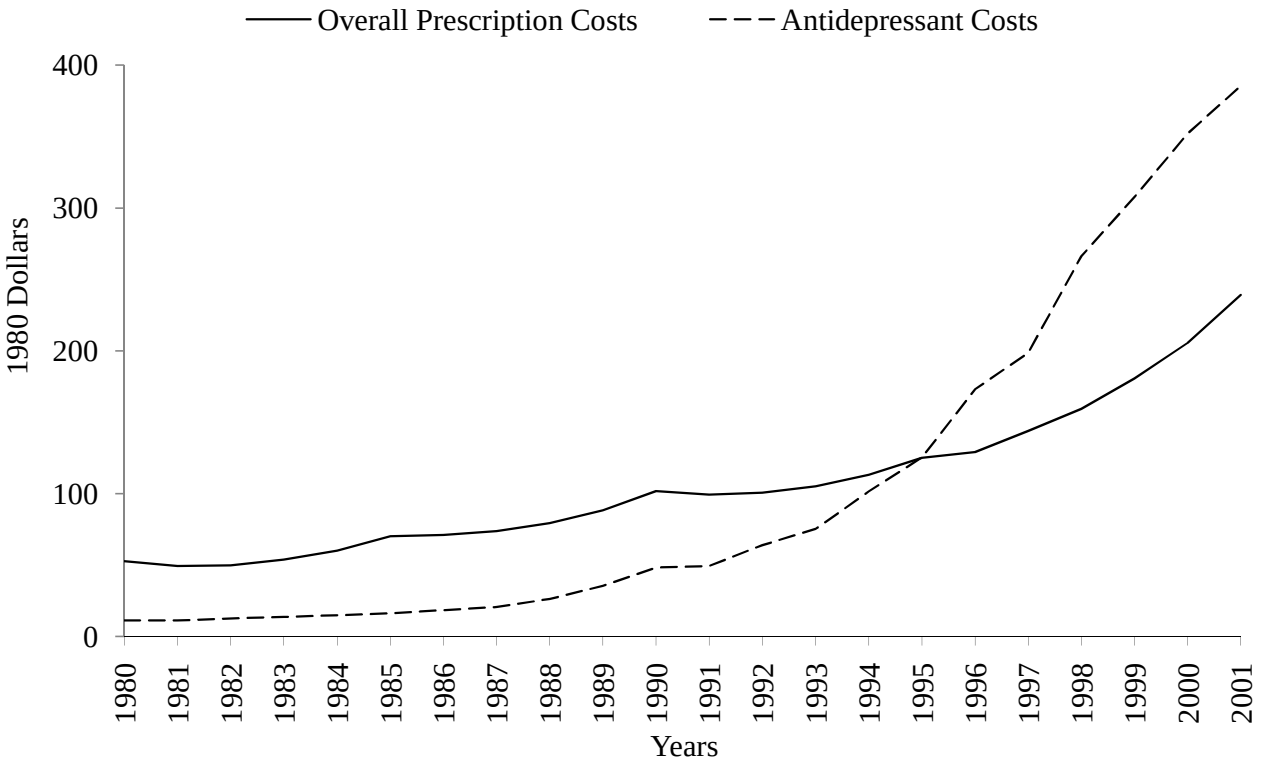

Figure 3
Per Capita Total & Antidepressant Prescription Costs