Job Market Paper

Patient Welfare Implications of Innovation in the U.S. Antidepressant Market[1]

Paris Cleanthous[2]
paris.cleanthous@yale.edu

Department of Economics
Yale University

This Version: October 2002

Abstract

Returns to innovation constitute a major component of social welfare. In this study, I formulate an empirical methodology that quantifies patient welfare benefits from pharmaceutical innovation in the U.S. antidepressant market. The antidepressant market has experienced an impressive stream of innovations over the last three decades and available data on antidepressants are exceptionally rich and accurate. The study 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 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 caused by the existence of pharmaceutical insurance coverage. The study estimates large patient welfare gains from innovation and helps explain 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.

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

[1]I am particularly grateful to my advisers Steve Berry, Penny Goldberg and Fiona Scott Morton for their invaluable guidance and continuous support. I also thank Christos Cabolis, Sofronis Clerides, C. Scott Grove, Alvin Klevorick, Sangin Park, Subrata Sen, Andreas Theotokis, Christopher Timmins, Dick Wittink and seminar participants at Yale University and at the Pharmaceutical Economics Ad Hoc Group meetings for helpful comments and suggestions. I am greatly indebted to Merck & Co., Inc. for graciously providing me with the IMS dataset and the people at Merck, especially Steven Propper and John Kennedy. I thank Susan Schmitter Jack of the CDC (NCHS) for helping me obtain an important portion of the data. The usual disclaimer applies.

[2]Address: 104 Lake Place, Apt. 203, New Haven, CT 06511. Website: http://www.econ.yale.edu/~paris

"No one has a big new answer on what to do about health care costs. And it's all made worse because health costs are rising in bad economic times," Drew Alman, president of a health research group. (The New York Times, August 11, 2002)

1 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 study, 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 three decades and available data on antidepressants are exceptionally rich and accurate. The study 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 from the IMS HEALTH Inc. dataset, graciously furnished by Merck & Co.; 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 study utilizes a structural discrete choice model of hedonic demand 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. Random 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 different 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 study 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. 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 study, especially after controlling for prescription insurance, can help evaluate the merits of competing claims. The study finds positive 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 2 reviews the relevant literature. Section 3 analyzes the characteristics of the market for antidepressants and the pertinent characteristics of the pharmaceutical industry. Section 4 presents the adaptation of previous theory and the methodology in estimating demand for antidepressants using various functional forms of demand. 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. Section 5 describes the dataset and the estimation procedure. Section 6 presents and discusses the results. Section 7 uses the demand estimation results to infer welfare implications of innovation in antidepressants. Section 8 provides a summary of contributions and key findings and discusses possible research extensions.

2 Existing 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 other 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. In this latter study, 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. The current study, instead, recomputes 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 another study, Goolsbee and Petrin (2002) estimate the gains to consumer welfare generated by the introduction of direct broadcast satellites.

In the pharmaceutical industry, work that attempts to quantify the economic value of innovation has been scarce. Lichtenberg (1996a, 1996b, 1998a, 2001) 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 (2001) 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 (1999) and Lichtenberg (1998b) 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.

An essential component of studies that try to evaluate the introduction of new products in differentiated product industries is demand estimation. Lancaster (1966) and (1977) introduced the characteristics approach to demand, and with the logit demand model, McFadden (1973) projected the products onto a space of characteristics. The utility of a choice is then determined by a parametric form for the interaction between consumer and product characteristics. Aggregating over the choices of consumers with different characteristics produces the demand function. Anderson, De Palma and Thisse (1992) analyze in detail the use of discrete choice models in the theory of product differentiation. In the absence of individual-level data that directly matches consumers preferences to observed purchases, models were developed that explicitly model consumer heterogeneity by employing data on the distribution of consumer characteristics. These models estimate the unknown parameters governing the distribution of consumer heterogeneity.

Notable empirical work that employ market-level data includes Berry, Levinsohn and Pakes[3] (1995, 1999) who examine equilibrium in the automobile industry and its implications for voluntary trade restrictions. They allow for the existence of unobserved characteristics which helps eliminate price endogeneity from the model. 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. Among other papers, Berry, Carnall and Spiller (1996) study hubs in the airline industry; Nevo (2000b, 2001) examines price competition and mergers in the ready-to-eat cereal industry; and Davis (2000) studies spatial competition in movie theaters. The use of individual-level data greatly complemented the results of studies that use market-level data. Goldberg (1995) estimates the demand for automobiles to investigate trade policy issues; BLP (1998) show how use of micro-level data augments the importance of the use of market level data; Sudhir (2001) estimates consumer pricing behavior in the automobile industry.

A substantial economic literature on pharmaceuticals, and antidepressants especially, is carried out by Ernst R. Berndt and co-authors. 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. Regarding research on other pharmaceutical markets, Berndt et al. (1990, 1993) and Griliches and Cockburn (1994) focus on the construction of price indices for a subset of the anti-infective class of drugs. Suslow (1992) constructs a hedonic pricing function for anti-ulcer drugs. Berndt et al. (1994) 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.

[3] Henceforth, BLP.

Other notable work in pharmaceuticals include Scott Morton (1997, 1999) and Ellison et al. (1997). The first paper examines the properties of price and market share, in particular, how they respond to a policy change (the MFN Clause of the Medicaid Rebate Law of 1990). The second paper focuses on how advertising affects the incidence of generic entry in a market of active ingredients. Ellison et al. examine price competition among four Cephalosporin drugs. They model demand for products rather than characteristics and treat it as a two-stage budgeting problem.

3 Market Background

Before estimating the demand for antidepressant drugs and welfare implications thereof, it is necessary to understand this pharmaceutical market. First, this section briefly outlines the institutional structure of the pharmaceutical industry and describes its features pertinent to this study. Then, it offers the reasons behind the choice of this specific pharmaceutical market. It continues with a discussion about clinical depression, analyzing its causes, the diagnostic process and its treatment, pharmacological and other. Then, this section presents a patient's choice set of antidepressants and how it has changed over time, the dates each drug entered and possibly withdrew from the market, and the drug characteristics. I conclude the section by discussing how a patient's choice of drugs is made.

3.1 Pharmaceutical Industry

An impressive stream of new products is observed, especially over the latter half of the twentieth century, due to rigorous research and development. 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. 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.

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

[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.

acute the illness the higher the insensitivity. The insensitivity is exacerbated by higher income and by insurance coverage.

Another unique industry characteristic is the unusually vigorous advertising and other promotional activity, such as detailing.[5] Spurred by product novelty, trademarking, and the difficulty consumers and prescribing physicians have in becoming informed about the efficacy of drug products, expenditure on promotional activities for pharmaceutical products ranks high among industries for both prescription and over-the-counter drugs. Detailing is reportedly the primary information source to 57 percent of physicians; 85 percent give the process a "strong vote of confidence," because of the valuable information it provides.[6] Though promotional activities are not considered in the current study, their effect on demand and welfare are the focus of a companion paper.[7]

3.2 Pertinent Characteristics of the Therapeutic Category

The question then becomes one of choosing the therapeutic area that would best address these questions. In order to perform the research required to answer the previous questions, the therapeutic area has to satisfy certain essential characteristics. First, the category has to include a considerable number of drugs. Second, it should have existed for a long time, so there exists a benchmark against which to compare innovation. Third, there must have been substantial qualitative innovation in the market over the years so that new generation drugs are much better than the original drugs.

This paper concentrates on the market for antidepressants because it satisfies best all of these conditions. The data come from the IMS America Ltd. Dataset (via Merck & Company, Inc.), which are exceptionally detailed for antidepressants. A therapeutic category or area includes drugs that are FDA-approved to be used in the pharmacological treatment of a specific disease. Drugs, therefore, that are therapeutically bioequivalent but are FDA approved for the treatment of different diseases will fall into different categories. I take the antidepressant class as described by the IMS America Ltd. Dataset, USC codes 64300-64399. The importance of this step is that I use only the drugs in the IMS antidepressant class.[8] Other useful, but not necessary, characteristics of the antidepressant category are as follows. Interaction across therapy areas and diseases is not a feature of the market. That is, depression does not require combinations of drugs from different categories. Finally, prescription drugs have better data than over-the-counter (OTC) drugs and antidepressants do not have major OTC sales.

3.3 The Market for Antidepressant Drugs

3.3.1 Depression

Antidepressant drugs are used in the pharmacological treatment of depression. The word *depression* tends to be associated with the condition of feeling sad or despondent. However, one must distinguish this colloquial usage from the clinical definition of depression, which describes a serious psychotic and incapacitating condition. The market of treating clinical depression is potentially quite lucrative.

[5] When a representative of a manufacturer of drugs calls on doctors, pharmacists, and other professional distributors to promote new drugs.

[6] Scherer, F. M. (1996).

[7] Cleanthous (2002).

[8] Suslow (1992) 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, a New Generation 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.

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 17.1 percent and 10.3 percent respectively, second only to hypertension.[9] 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. A person with untreated depression cannot function efficiently in her business or social environment, overuses the health care system, takes longer to recover from other illnesses and may commit suicide.

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.[10] Under-diagnosis and under-treatment may be due to various causes: Patients may not link their symptoms to a disease; public comprehension of mental diseases is generally poor; depression still constitutes a social stigma and primary care physicians miss diagnosing depression half of the time.[11]

Causes, Diagnosis & Pharmacological Treatment[12] Depression has been *associated* with a chemical (neurotransmitter) imbalance in the brain. Serotonin is the principal neurotransmitter associated with depression but the neurotransmitters dopamine, norepinephrine and adrenaline may also be important. The cause of the chemical imbalance could be biological and/or psychological. Antidepressant drugs are prescribed to fight directly the cause of the chemical imbalance. However, the current range of antidepressants can only treat some of the causes of the chemical imbalance.

There currently exists no definitive biological test for the diagnosis of depression. Consequently, the psychiatrist diagnoses depression with only the external characteristics (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.[13] 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 five main types of antidepressant drugs (according to IMS America Ltd.):

Type of Antidepressants	**Examples**
Monoamine Oxidase Inhibitors (MAOI)	Nardil, Marplan
Tricyclic and Tetracyclic Antidepressants (TCA)	Elavil, Endep
Selective Serotonin Reuptake Inhibitors (SSRI)	Prozac, Paxil
New Generation Antidepressants (NEWGEN)	Wellbutrin
Combinations in Antidepressants (COMBO)	Limbitrol

Monoamine Oxidase Inhibitors (MAOI) remedy the chemical imbalance by killing the enzymes that destroy the neurotransmitters after they are released by the secreting cells. All other types of antidepressants work directly in correcting the chemical imbalance in a neurotransmitter. Tricyclic

[9] Kessler et al. (1994) and Journal of the American Medical Association (2002).

[10] Miranda (1994).

[11] Depression Guideline Panel (2002) and Salmans (1995).

[12] A more detailed discussion on the causes, diagnosis and effects of depression are in an online appendix.

[13] The criteria for diagnosis of depression are listed in the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. The list of symptoms is as follows: depressed mood; diminished interest or pleasure in activities; weight loss/gain or decrease/increase in appetite; insomnia or hyper-insomnia; abnormal speeding up or slowing down of one's activities or mental processes; fatigue; feeling of worthlessness or excessive guilt; lack of concentration or indecisiveness; recurrent thoughts of death or suicide, or suicide attempts. To be diagnosed with depression, a patient must exhibit at least five of the symptoms on a daily basis for a two-week period. The five symptoms must include one of the first or second symptoms, or both.

and Tetracyclic Antidepressants (TCA) inhibit the reuptake of serotonin and norepinephrine into the central nervous system; Selective Serotonin Reuptake Inhibitors (SSRI) prevent the reuptake of serotonin by the secreting cells; New Generation Antidepressants (NewGen) consist of reuptake inhibitors of other neurotransmitters and of antagonists of other enzymes; Combinations of Antidepressants (COMBO) are combinations of TCA and have the combined effect of their constituent drugs on the serotonin and norepinephrine. Types are subdivided even further into collections of drugs with the same molecule (active ingredient), for example, amitriptyline (generic), Elavil and Endep.

History of Antidepressants[14] The first two antidepressants, iproniazid and imipramine, were introduced in the late 1950s and were of two different types: MAOI and TCA. Though MAOIs were more successful in the very beginning with two more molecules having been introduced, the TCAs soon overtook their success since MAOIs were found to have more severe overall side effects. Expansion in the market continued steadily with the introduction of more TCAs and drug combinations of TCAs and 1981 marked the introduction of trazodone, the first new generation antidepressant. Finally, the first SSRI, fluoxetine hydrochloride (Prozac), entered the market in 1988 and swept the market. This was due to unprecedented media attention that proclaimed Prozac "a wonder drug," the marketing efforts of Lilly and its less severe side effects, especially, the fact that an overdose of the drug would not be lethal. Other than that, the reported therapeutic advantage of Prozac was not any different to the existing drugs.[15] This disproportionate success of Prozac led to introductions of more SSRIs in the 1990s and more new generation antidepressants. Tables 2 and 3 report in detail the historic entry and withdrawal of antidepressants into the market from 1958 to the present. Whereas entry represents the year in which the drug is first sold in the market, withdrawal years represent collectively those years in the IMS data with no sales. The data are chronologically ranked within molecules, molecules within types and types within the antidepressant market.

Other Treatments 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. Patients may either disregard depressive symptoms altogether, or accept depression as a way of life. Patients may also try different treatments for depression. There are two main alternative treatments. Psychotherapy and electroconvulsive therapy. The former is attempts to treat depression through the use of psychological techniques designed to encourage communication of conflicts and insight into problems. Though it is generally advised that psychotherapy is attended in conjunction with taking antidepressants, it is the case that some patients rely only on psychotherapy for the treatment of depression.

3.3.2 Choice in Antidepressants

The Choice Set 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 4 lists a combination of all the possible choices in antidepressants that appeared at least once over the twenty-two year period of the dataset used. The table divides the antidepressant medications into their different types and molecules. For instance, a choice of a specific drug among Prozac, Sarafem, Prozac Weekly or the generic alternative presupposes a choice of molecule, in this case Fluoxetine Hydrochloride, which in turn presupposes a choice of type of antidepressant medication, here SSRI. Note that the choice in antidepressants should be viewed as simultaneous rather than

[14] See Currie (1998) for a more detailed discusiion.

[15] In fact, standard medical reference books (Physician's Desk Reference, Drug Facts & Comparison, Clinical Psychopharmachology) still record TCAs as the suggested first-line therapy.

hierarchical. The divisions into groups are market segmentation characteristics and help the choice maker in matching tastes and preferences to drug characteristics.

Characteristics of Antidepressants Historical evidence indicates that no one antidepressant is clearly more effective than another in achieving the desired health outcome.[16] A major source of differentiation, therefore, is the mechanism of action of an antidepressant as this is identified by a drug's type. Another major source of differentiation is an antidepressant's side effect profile that is common to drugs of the same active ingredient (molecule). Examples of side effects include a drug's fatality, dry mouth, blurred vision and drowsiness. Table 1 provides detail on the occurrence of these side effects and a summary of other antidepressant characteristics.

The table also includes information on dosage requirements (initial dosage, daily dosage requirement in milligram, average frequency of daily administration), available forms of administration (tablet, liquid, injectable, capsule) and the drug's half-life or bioavailability. The latter characteristic is the rate at which the drug becomes available at the site of physiological activity after administration. When modeling a hedonic pharmaceutical demand one may also include as characteristics dummy variables of whether a drug is a brand or a generic and whether generic competition exists. Looking at Table 1 it can be seen that characteristics are quite similar within types. Moreover, it can be concluded that SSRIs tend to have a less adverse side effect profile than other drugs and most importantly they are less likely to be fatal. The favorable characteristics of SSRIs will prove important in analyzing the primary demand effects of product introduction.

Choice Maker[17] In most industries consumers chose the product, the quantity and the method of payment. In the case of prescription drugs the decision is shared by the patient, the physician and sometimes the prescription drug coverage provider. If a patient were left alone to make a decision, she would base that decision on the expected health outcome of a treatment and the cost of the treatment net of any insurance co-payment. A patient's expectation on a health outcome depends on her information about the treatment, which in turn depends on factors like health awareness, direct-to-consumer advertising, word-of-mouth, personal experience with antidepressants or medication for symptomatically similar diseases. However, legislation prevents and protects the patient from making an uninformed decision by requiring that a prescribing physician makes the treatment choice. The patient, therefore, can only participate in the optimization of her utility by trying to affect the physician's preferences. It is reasonable to assume that drug-prescribing physicians care about their patients and, thus, try to maximize their patients' utility.

In the case of depression, patients are highly heterogeneous in their response to treatment, hence, experience with other patients should only influence a physician's decision initially. For the same reason, existing protocols and guidelines for the treatment of depression are merely suggestive in nature.[18] What is more, existing formularies[19] only make a distinction between branded and generic antidepressants and not across types and molecules. The initial choice of an antidepressant type and molecule is based on the patient's own or her family's medical history. In the absence of a medical history, physicians start an experimentation phase; often, a physician will begin with antidepressants with the least overall side effects: some SSRI, TCA. Therapeutic effects appear within two to six weeks. Treatment of depression typically takes much longer. It may vary from a few years in the cases of mild depression to a person's life span. This implies that a patient's initial experimentation

[16] Depression Guideline Panel (2002).

[17] Wosińska (2002) has a nice discussion about the interaction of the various agents when making a choice in pharmaceuticals, in general. The discussion, here, focuses on antidepressants.

[18] An example of suggested guidelines from the Depression Guideline Panel is available on my website.

[19] The Lewin Group (2000).

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 the verbal statements of 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*, which is very common in the pharmaceutical industry. Spurious product differentiation arises when consumers perceive physically identical products to differ in quality.[20] The decision to buy brand over generic is influenced by the patient's perception of quality and the price difference between two drugs (depends on prescription drug insurance). 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.[21] In an extension of the current study data broken down by different distribution channels will help decipher the extent of the agency problem in the choice of antidepressants.

4 Modeling Demand for Antidepressants

4.1 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,[22] 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.

[20] A wide literature exists on an additional step in a patient's choice: whether to comply with a prescription or not. In the current version of the paper I am only dealing with a patient's decision to purchase the antidepressant and will deal with the issue non-compliance in a future version.

[21] State legislation exists that forces the use of generic medication when that exists, for example in Maine, Ellison and Snyder (2001).

[22] 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.

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 hypothesis is that each patient buys at most one unit of the good. 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 (more accurately) by the variety in individual patient preferences for drug characteristics. Market-level demand is then obtained by aggregating individual demands.

Given that the data available are aggregated across patients, I initially assume that extra data on patient characteristics are not available.[23] Therefore, all the parameter estimates of the demand system use only drug level data on prices, quantities and drug characteristics. The methodology that follows allows the different demand models to be estimated with only market-level price and quantity data. Eventually, I incorporate information on the distribution of pertinent individual patient characteristics, specifically, income and prescription drug insurance coverage.[24]

4.2 General Model Overview[25]

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, $j = 0$. 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. Observed characteristics are described in Section *3.3.2*. Unobserved product characteristics include variables with unavailable data such as promotional activity[26] and additional drug side effects. Patients are assumed to observe all the drug characteristics.

Following the suggestion in BLP (1995) and Petrin (2002) I rewrite the utility function at time period t as:

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

The utility level is, therefore, broken into three components. First, δ_{jt} is the mean utility level and 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_{ij}, is a random utility component across drugs and patients and is assumed to be independent and identically distributed (i.i.d.) across both drugs and patients. The sum of the latter two components represents the deviation from the mean utility level for each patient i and is a measure of the idiosyncratic valuation of drug j's characteristics.

Each patient type will purchase one unit of that 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,

[23] BLP (1995) approach this in a similar fashion. The use of aggregate data is common in the literature.

[24] In an extension of this paper, data that directly match patient choices to patient characteristics are utilized for the latter six years of the dataset.

[25] This discussion follows closely the discussion in Berry (1994) and BLP (1995).

[26] Work in progress extends the model described here and includes data on promotional activity directed towards patients and towards physicians.

therefore, choose to purchase one unit of drug j, at time t, **if and only if**

$$u_{ijt}(\theta_d) - u_{ikt}(\theta_d) > 0, \, \forall k \geq 0, \, k \neq j, \, \forall t. \tag{2}$$

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 $A_j(\delta) = \{\mu_i \,|\, \delta_j + \mu_{ij} > \delta_k + \mu_{ik}, \, \forall k \neq j\}$. 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 $A_j(\delta)$:

$$s_j\left(\nu_i, x_j, \xi_j, p_j; \theta_d\right) = \int_{A_j(\delta)} f(\mu; \sigma) d\mu \tag{3}$$

A closed form solution may exist depending on the density function chosen. Otherwise, simulation methods can be used to estimate the drug market shares.

At every time period t, each patient maximizes her level of utility as follows:

$$\max_{j \in \{0,1,...J_t\}} u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt}, \tag{4}$$

where the mean utility level is as in Berry (1994): $\delta_j \equiv \alpha p_j + x_j' \beta + \xi_j$. The β are the marginal utilities (or disutilities, accordingly) of the drug's observed characteristics and the α is the marginal disutility associated with price. Note that this formulation of utility specifies that the unobserved characteristic is identical for all patients. By letting the price coefficient vary across patients in the full random coefficients model the ξ_j captures the elements of vertical product differentiation in the antidepressant market.

4.2.1 Estimating from the mean utility levels

To be able to estimate demand for antidepressants, one needs to know the size of a market, M. 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. In this study M is directly observed[27] so shares can be calculated from the data: $\widetilde{s}_j = q_j / M$. To solve for the parameters that enter the market share function, set:

$$\widetilde{s}_j = s_j(\delta(x, p, \xi); \theta_d) \; for \; j = 0, 1, ..., J_t, \tag{5}$$

which should hold exact at the true values of δ.

However, a problem arises due to the presence of the mean of patients' valuations of unobserved drug characteristics, ξ_j. 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, $\delta_j = s_j^{-1}(\widetilde{s}_j)$. Traditional IV techniques to estimate the unknown parameters β and α are now feasible.

[27] M could also be unobserved and left as a parameter to be estimated.

4.2.2 General Model Assumptions

I concentrate here on the main assumptions that form building blocks for the various models of demand proposed below:

Assumption, A1: (DCM). Each patient purchases at most *one* unit of the drug.

Assumption, A2: There exists an outside option.

Assumption, A3: The market size is directly observed and is equal to the portion of the U.S. population that is estimated to be clinically depressed, $\widetilde{M}$.

Assumption, A4: The random component of the utility specification across drugs and patients, ϵ_{ijt}, are independent and identically distributed (i.i.d.) across both drugs and patients.

Assumption, A5: (Logit) The distribution of consumer preferences over the unobserved product characteristics, ϵ_{ijt}, is the extreme value distribution (EVD): $\epsilon \sim exp(-exp(-\epsilon)$.

Assumption, A6: Extra data that match patient characteristics to the drugs each patient purchased are not available. That is, the term that captures the heterogeneity in patient preferences for observed (by the econometrician) drug characteristics is zero: $\mu_{ijt}(\theta_d) = 0$.

Assumption, A7: All patients are identical in their valuation of the observed drug characteristics (including price), for all drugs j. That is, assume invariant coefficients of observed demand characteristics across patients: $\alpha_i = \alpha$ and $\beta_i = \beta$, $\forall i$.

Assumptions A1-A5 hold throughout. Assumptions A6-A7 hold for the simple logit and nested multinomial logit models discussed below. The full random coefficient logit model allows for patient heterogeneity using both a multivariate normal distribution and demographic data on prescription drug insurance and income.

4.3 The Multinomial Logit Model

Let assumptions A1 to A7 hold. The latter two assumptions make the model simple and thus the patient utility becomes: $u_{ijt} = \delta_{jt} + \epsilon_{ijt}$. Then the familiar logit probability that any patient i purchases drug j at period t is:

$$\Pr(\textit{patient } i \textit{ buys drug } j \textit{ at time } t) = \frac{e^{\delta_{jt}}}{\sum_{t=1}^{J_t} e^{\delta_{jt}}}. \tag{6}$$

This is equivalent to the market share for drug j at time t, s_{jt}. Normalizing the mean utility level of the outside option, δ_0, to zero this market share becomes:

$$s_{jt}(\delta) = \frac{e^{\delta_{jt}}}{1 + \sum_{k=1}^{J_t} e^{\delta_{jt}}} \; for \; j = 0, ..., J_t. \tag{7}$$

The δ_j is uniquely identified by taking logarithms and subtracting the outside option share from each market share. Then δ_j can be regressed on observed product characteristics using a standard IV regression:

$$\ln(s_{jt}) - \ln(s_{0t}) = \delta_{jt} \equiv \alpha p_{jt} + x'_{jt}\beta + \xi_{jt} \tag{8}$$

Once the estimates are available, own-price and cross-price elasticities of demand can be computed:

$$\eta_{s_{jt},p_{kt}} = \frac{\partial s_{jt}}{\partial p_{kt}} \cdot \frac{p_{kt}}{s_{jt}} = \left\{ \begin{array}{l} \alpha p_{jt}(1 - s_{jt}),\ j = k \\ -\alpha p_{kt} s_{kt},\ j \neq k \end{array} \right\}. \tag{9}$$

This simple logit formulation gives a convenient formula with a closed form solution and few parameters to estimate. Assumptions A6 and A7 imply that the variation in patient preferences enters the model only through the additive term ϵ_{ijt}. This places strong restrictions on the pattern of cross-price elasticities from the estimated model. All properties of market demand (including market shares, elasticities and, thus, substitution patterns) are determined solely by the δ_{jt} and not by similarities in product characteristics. In particular, products with same market shares will have the same cross-price elasticities with any other product. For example, if Prozac and generic Amitriptyline are found to have the same market shares and there is a decrease in the price of Zoloft (same antidepressant type as Prozac, SSRI), then the simple logit model would suggest that its effect on Prozac's demand would be the same as the effect on the demand for generic Amitriptyline, an not an intuitive implication. Substitution patterns between antidepressant drugs may become unreasonable.

Own price elasticities of this simple multinomial logit also present a problem in the antidepressant market. In most cases the market shares are very small, sometimes negligible. This is mainly due to the large share of the outside option (especially in the beginning of the period) and the fact that some of the innovative antidepressants have swept the market. This implies that an antidepressant's own-price elasticity of demand is almost constant, which makes the elasticity correlated with that product's price. As a result, the higher the price the higher the elasticity and, hence, the lower the markup, which will not always be true.

4.4 The Two-Level Nested Multinomial Logit Model

In contrast to the simple logit model, the nested multinomial logit model addresses the problem of unreasonable substitution patterns by interacting individual patient attributes with an a priori grouping of drug characteristics. The imposed correlation between attributes is still restricted by the choice of groupings. When the latter is normative, grouping is considered a shortcoming of the nested logit model. In the antidepressant market, I hypothesize that three such groupings may exist: types, molecules and 'brandness'.[28] Accordingly, a one-level (NML1), two-level (NML2) or a three-level (NML3) multinomial logit models can be used. The NML1 only groups drugs into types. The NML2 is the focus of this section, though estimates of all three will be provided.

Let assumptions A1 to A7 hold. Assume that a patient chooses among J_t different drugs at every time period t. This time, however, her choice is threefold. The patient makes a choice among five options: the outside option and the four exhaustive and mutually exclusive types that comprise of the four different types of antidepressants: MAOI, TCA, SSRI and New Generation antidepressants. This is first nested choice. Let c denote these **types** within the therapeutic area of antidepressants, where $c = 0, 1, ..., J_{ct}$. It is important to allow J_c vary across time since two of the sub-categories did not exist over the whole span of the dataset.

Within each type of antidepressants, a patient still has to choose a **molecule**. This is the second nested choice. Let m_c denote a molecular group within type c, where $m_c = 1, ..., J_{m_c t}$. Again I allow

[28] 'Brandness' refers to a choice between a branded and a generic drug

J_{m_c} to vary across time. Finally, the patient reaches a decision on which drug j to purchase. I include the qualification of a drug as branded or generic in the form of a dummy variable.

Following the analysis in McFadden (1978) and Cardell (1997) and extending the example in Berry (1994) to a NML2, the utility level in equation 4 becomes:

$$\begin{aligned} u_{ijt} &= \delta_{jt} + \psi_{ijt} \\ \text{where } \psi_{ijt} &= \zeta_{ict} + (1-\rho_c)\,\zeta_{im_ct} + (1-\rho_c)\left(1-\rho_{m_c}\right)\epsilon_{ijt}; \quad \rho_c, \rho_{m_c} \in [0,1)\ . \end{aligned} \tag{10}$$

The idiosyncratic component of the utility function has the variance components structure as laid out in Cardell (1997). Since the ϵ_{ij}[29] are i.i.d. extreme value, Cardell proves that so is $\zeta_{im} + \left(1-\rho_{m_c}\right)\epsilon_{ij}$. Similarly, given that $\zeta_{im} + \left(1-\rho_{m_c}\right)\epsilon_{ij}$ is i.i.d. extreme value, so is $\zeta_{ic} + (1-\rho_c)\left\{\zeta_{im} + \left(1-\rho_{m_c}\right)\epsilon_{ij}\right\}$, which is equivalent to ψ_{ij}. For patient i, the variable ζ_{ict} is common to all drugs of the same type c and has a distribution function that depends upon ρ_c with $0 \leq \rho_c < 1$. As ρ_c approaches one, the within-type correlation of utility levels goes to one, hence patients are substituting towards another product of the same type. In a similar fashion, ζ_{im} is common to all drugs that belong to molecule m, and ρ_m represents the within-molecule correlation.

The market share of drug j as a fraction of the total second-nesting group is equal to the probability that a patient would purchase drug j over all drugs in molecule m_c:

$$s_{j/m}\left(\delta_j, \rho_c, \rho_{m_c}\right) = \frac{e^{\delta_j/(1-\gamma)}}{\underbrace{\sum_{k \in J_{m_c}} e^{\delta_k/(1-\gamma)}}_{D_m}} \ \forall j \in \{1, 2, ..., J_{m_c}\}\,. \tag{11}$$

where $\gamma = [1-(1-\rho_{m_c})(1-\rho_c)] \in [0,1)$. Similarly, the market share of the second-nesting group as a fraction of the total first-nesting group is equivalent to the probability that a patient would purchase molecule m over all drugs of type c:

$$s_{m/c}\left(\delta_j, \rho_c, \rho_{m_c}\right) = \frac{D_m^{\left(1-\rho_{m_c}\right)}}{\underbrace{\sum_{k \in J_m} D_m^{\left(1-\rho_{m_c}\right)}}_{D_c}} \ \forall j \in \{1, 2, ..., J_c\}\,. \tag{12}$$

Therefore, the probability of choosing a type c drug is::

$$s_{j/c}\left(\delta_j, \rho_c, \rho_{m_c}\right) = \frac{D_c^{(1-\rho_c)}}{\underbrace{\sum_{k \in J_c} D_c^{(1-\rho_c)}}_{D}}. \tag{13}$$

and the overall market share for drug j is

$$\begin{aligned} s_j\left(\delta_j, \rho_c, \rho_{m_c}\right) &= s_{j/m_c}\left(\delta_j, \rho_c, \rho_{m_c}\right) \cdot s_{m_c/c}\left(\delta_j, \rho_c, \rho_{m_c}\right) \cdot s_{j/c}\left(\delta_j, \rho_c, \rho_{m_c}\right) \\ &= e^{\delta_j/(1-\gamma)} \cdot D_m^{-\rho_{m_c}} \cdot D_c^{-\rho_c} \cdot D^{-1}. \end{aligned} \tag{14}$$

[29] For expositional simplicity, the t's are dropped for the remainder of this section.

With $\delta_0 = 0$, the share of the outside option $s_0\left(\delta_0, \rho_c, \rho_{m_c}\right) = \frac{1}{D}$.Taking logs of market shares and finding the difference between them now gives:

$$\ln\left(s_j\right) - \ln\left(s_0\right) = \delta_j/(1-\gamma) - \rho_{m_c}\ln(D_m) - \rho_c \ln(D_c). \tag{15}$$

Taking logarithms of equations (11) and (12) leads to an analytic expression for s_j^{-1}:

$$\delta_j = \ln\left(s_j\right) - \ln\left(s_0\right) - \rho_{m_c}\ln(\widetilde{s}_{m_c/c}) - \gamma \ln(\widetilde{s}_{j/m_c}), \tag{16}$$

where the $\widetilde{s}_{j/m_c}$ indicate observed shares of a drug within a molecule and the $\widetilde{s}_{m_c/c}$ of a molecule within a type of antidepressants. Inserting equation (**??**) into the equation above I get the equation to be estimated:

$$\ln\left(s_j\right) - \ln\left(s_0\right) = \delta_j + \rho_{m_c}\ln(\widetilde{s}_{m_c/c}) + \gamma \ln(\widetilde{s}_{j/m_c}), \tag{17}$$

Notice that if the within-molecule drug correlation is set to zero, $\rho_{m_c} = 0$, then the above equation is the same as the NML1 equation with an imposed correlation among drugs of the same type but not molecule. Similarly, if the within-type molecule correlation is set to zero, $\rho_c = 0$, then the above equation is again the same as the one-level nested multinomial logit formula with an imposed correlation among drugs of the same molecule but not type. If both the within-molecule drug correlation and the within-type molecule correlation are set to zero then equation (17) reduces to the estimation equation for the simple multinomial logit, presented above.

Estimates of β, α, ρ_{m_c} and ρ_c can be obtained from a linear instrumental variables regression of differencs in logarithmic market shares on drug characteristics, prices and the logarithms of within-molecule drug shares and of within-type molecule shares. As already explained, instrumental variables are required to correct for the endogeneity of prices and the logarithms of group shares. With these estimates, own-price and cross-price elasticities of demand, $\eta_{s_j,p_k} = \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j}$, can be estimated:

$$\eta_{s_j,p_k} = \left\{ \begin{array}{l} \frac{\alpha p_j}{(1-\gamma)}[1 - \rho_{m_c} s_{j/m_c} - \rho_c(1-\rho_{m_c}) s_{j/m_c} s_{m_c/c} - (1-\gamma)s_j],\ j = k \\ \frac{\alpha p_k}{(1-\gamma)}[1 - \rho_{m_c} s_{k/m_c} - \rho_c(1-\rho_{m_c}) s_{k/m_c} s_{m_c/c} - (1-\gamma)s_k],\ j \neq k, \text{same } m \\ \frac{\alpha p_k}{(1-\gamma)}[-\rho_c(1-\rho_{m_c}) s_{k/m} s_{m_c/c} - (1-\gamma)s_k],\ j \neq k, \text{same } c, \text{ different } m \\ -\alpha p_k s_k,\ j \neq k,\ j \text{ of different } c \text{ and } m. \end{array} \right\} \tag{18}$$

To summarize, the NML2 model has allowed for more flexible substitution patterns between antidepressants by letting patient preferences to be somewhat correlated across drugs j. Like the multinomial logit model, the NML2 model has a closed-form solution for the market share equation (3) and is, therefore, computationally simpler than a full random coefficients model. However, the division into the a priori groups might be more technical than realistic and might only approximate substitution patterns across antidepressant drugs. The same advantages and disadvantages hold true for the cases of the one-level and three-level multinomial nested logit models. The more restrictive the structure placed on patient choice the less representative the results, when the structure does not accurately describe the market.

4.5 The Random Coefficients Logit Model

This section focuses on a random coefficients specification for utility. The full random coefficients logit (RCL) model 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.

Let assumptions A1-A5 hold and relax assumptions A6-A7. By relaxing A7, that is $\mu_{ij}(\theta_d) \neq 0$, a model with random coefficients on the drug characteristics avoids the problem of a priori unreasonable substitution patterns. The utility specification for patient i choosing drug j returns now to its original form:

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt} \tag{19}$$

where, as above, the mean utility level $\delta_{jt} = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt}$. In this way the patient heterogeneity in preferences will be captured. A patient's substitution to a new drug due to an increase in the price of the initial drug chosen will depend on the attributes of her initial choice, her income and her prescription drug coverage. As already mentioned, the model will be estimated, first, by assuming a multivariate normal distribution of patient tastes. Then, more precise estimates will be obtained by using information on the distribution of patient preferences as it relates to some of the drug characteristics. The benefit of this methodology is that it does not require observations on patient purchase decisions to estimate the demand parameters. Petrin (2002), however, explains that market-level data may or may not provide good information on how patients substitute between drugs.[30] As remains to be seen in the estimation results, the use of demographics in this study provide ample information on patients' substitution patterns and, hence, estimates of the parameters of the distribution of patient preferences are precise.

4.5.1 Including Information on Patient Heterogeneity[31]

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). This section discusses how patient preferences vary with the ν_i and incorporates in the model the distribution of patient characteristics. 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 information 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 study 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:

$$\begin{pmatrix}\alpha_i \\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \Pi D_i + \Lambda\tau_i,\ D_i \tilde{} P_D^*(D),\ \tau_i \tilde{} P_\tau^*(\tau) \tag{20}$$

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

[30] An improvement in the richness of the results would be to use patient-level data, like survey data, similarly to Goldberg (1995) and BLP (1998). This is the focus of an extension to this paper.

[31] The discussion here borrows from Nevo (2000a).

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.[32] Some additional assumptions are required to complete the model formulation.

Assumption, A8: $P^*_\tau(.)$ has a standard multivariate normal distribution.

Assumption, A9: τ_i and D_i are independent.

Under assumption A8, Λ allows a different variance for each component of τ_i and a correlation among these patient preferences. Let $\theta_d = [\theta_1, \theta_2]$[33] and rewrite the utility model in equation (1) using equations (19) and (20). Then,

$$\begin{aligned} u_{ijt}(\nu_{it}, x_{jt}, \xi_{jt}, p_{jt}; \theta_d) &= \delta_{jt}(x_{jt}, \xi_{jt}, p_{jt}; \theta_1) + \mu_{ijt}(\tau_{it}, D_{it}, x_{jt}, \xi_{jt}, p_{jt}; \theta_2) + \epsilon_{ijt} \\ \text{where } \delta_{jt} &= \alpha p_{jt} + x'_{jt}\beta + \xi_{jt}, \;\; \mu_{ijt} = [p_{jt}, x_{jt}] * (\Pi D_{it} + \Lambda \boldsymbol{\tau}_{it}) \end{aligned} \tag{21}$$

and following a similar procedure as in the model overview the probability that a patient chooses drug j at time t as expressed in equation (3) can now be written for the random coefficients model using Bayes rule and under the distributional assumptions[34] as:

$$\begin{aligned} s_{jt}\left(x_{jt}, \xi_{jt}, p_{jt}; \theta_2\right) &= \int_{A_{jt}(\delta)} dP^*(D, \tau, \epsilon) \\ &= \int_{A_{jt}(\delta)} dP^*(\epsilon|D, \tau) dP^*(\tau|D) dP^*(D) \\ &= \int_{A_{jt}(\delta)} dP^*_\epsilon(\epsilon) dP^*_\tau(\tau) dP^*_D(D) \end{aligned} \tag{22}$$

Under assumption A5 on the distribution of the random utility component, the RCL market shares become:

$$s_{ijt} = \frac{\exp\left(\delta_{jt} + \mu_{ijt}\right)}{1 + \sum_{k\in J_t} \exp\left(\delta_{kt} + \mu_{ikt}\right)}. \tag{23}$$

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 can be derived similarly as in the previous models.

$$\eta_{s_{jt}, p_{kt}} = \frac{\partial s_{jt}}{\partial p_{kt}} \cdot \frac{p_{kt}}{s_{jt}} = \left\{ \begin{array}{l} \frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) dP^*_\tau(\tau) dP^*_D(D), \;\; j = k \\ -\frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dP^*_\tau(\tau) dP^*_D(D), \;\; j \neq k \end{array} \right\} \tag{24}$$

Unlike in the previous model specifications, each patient type now 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 no longer the result of just the chosen functional form and cross-elasticities are larger for products that are closer in terms of their characteristics. The a priori grouping restrictions of the nested multinomial logit models are now relaxed but I do include dummy variables for type, molecule and taste for branded drugs as product characteristics.

[32] Two points are of notice here. First, I allow all characteristics to have random coefficients. This is necessary as was explained in Section 3. Second, only price and preference for branded drugs are allowed to vary with demographics. This means that 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.

[33] The distinction between θ_1 and θ_2 is necessary for estimation purposes. Whereas θ_1 enters the estimation linearly, θ_2 enters non-linearly.

[34] $P^*(.)$ denote population distribution functions.

5 Data & Estimation

5.1 Data Description

The purpose of this section is to briefly describe the data employed in this study. The data required are divided into three categories: market shares and prices, drug characteristics (physical or not), distribution on patient demographics, $P_D^*(D)$. Sales data come from the IMS America Inc. dataset and have been graciously furnished for this project by Merck and Company, Inc. IMS data on antidepressants are both detailed and 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.

I take the antidepressant class as described by IMS, USC codes 64300-64399. Data in this study run for twenty-two years spanning the time period from 1980 to 2001. Over the twenty-two year span the market for antidepressants includes a total of forty-seven drugs, which are sub-divided into twenty-four molecules. These are in turn grouped into four types of antidepressants.[35] These constitute market-segmentation characteristics, as does the distinction between branded and generic drugs. IMS data on drug entry (both branded and generic[36]) into and withdrawal[37] from the antidepressant market are also used in conjunction to similar data from the FDA's 22^{nd} edition of the *Approved Drug Products* catalog (the 'Orange Book'). The latter catalog also provided the data on therapeutic equivalence evaluations between branded antidepressant drugs and their generics, patent and exclusivity rights enforcement and expiration dates. In the case of all antidepressant medications, the FDA qualifies all medications within the same molecule as therapeutically equivalent.

Table 2a presents the year of entry for each antidepressant medication beginning at the introduction of the therapeutic area in 1959. Two types of antidepressants were simultaneously introduced in 1959, MAOI and TCA, with the respective entry of Marplan and Tofranil in the market. New Generation antidepressants appeared in 1982 with the advent of Desyrel and, finally, Prozac's entry in 1988 created the SSRIs. Twenty-eight antidepressant innovations took place from 1980 to 2001, twelve of which were molecule innovations. Ten of these innovations were in TCA, eight in SSRI and ten in new generation antidepressants. There was no entry in MAOI. Both branded and generic entry took place. Seventeen branded antidepressant drugs were introduced and eleven generics. Generic entry was mostly in TCAs. Summary statistics on entry are presented in Table 2b. Table 3 lists the years of no IMS-reported data, the 'withdrawal' years. These events could represent withdrawal of a drug by the producing firm and thus no sales, missing data, or insufficient sales that do not reach the IMS minimum reporting threshold.

Data on physical antidepressant drug characteristics including side effects, active ingredients, half-life, dosage information and types of administration come from various sources: the *Drug Information Handbook*, *Drug Facts and Comparisons*, *Prescription Drug Reference*, *Depression Guideline Panel* and the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. The latter two sources also provided the information on symptomatology and suggestive guidelines for the treatment of depression. Data on the segmentation of the antidepressant market

[35] Until 1995, data on drugs that derive from combinations of antidepressants were not included in the 64300-64399 range. For this reason data on this fifth type of antidepressants are incomplete and will be excluded.

[36] Though sales for generic drugs are broken down by distributor, I aggregate data over all distributors for each generic drug.

[37] Withdrawal of a drug is equivalent to zero quantity reported at a specific year. This means that production of the drug in question was suspended, or quantities sold were below the minimum threshold reported by IMS, or even that data might be missing, which is hardly the case for antidepressants.

into types and molecules came from IMS and was verified using FDA publications.

Table 1 shows the observed characteristics for all antidepressant molecules. 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. The table also includes information on dosage requirements (initial dosage, daily dosage requirement in milligrams, average frequency of daily administration), available forms of administration (tablet, liquid, injectable, capsule) and the drug's half-life or bioavailability.

The reported characteristics show that MAOIs tend to have the most adverse side effect profile including high fatality rates, whereas some new generation antidepressants have the least adverse side effect profile. 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 just like the one in Table 1 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. New generation antidepressants have the least correlation of side effect occurrence. In fact, the characteristics of new generation antidepressants are quite variable so grouping them together may be unwise.

Finally, the table includes information on market segmentation characteristics. Model estimation uses dummy variables to account for these market segmentation characteristics, specifically, a drug's type, molecule and whether the drug is branded or generic. Market segmentation information is also available in Table 4. The main aim of this table is to provide a union of all available choice sets in the twenty-two year period, which includes all forty-seven drugs.[38] The table provides a quick reference to market segmentation for antidepressants and is useful in the analysis of the estimation results. Patients use these virtual characteristics in addition to the drug's physical characteristics when making their decision. Table 4 provides both the drug's trade name as well as the drug's generic name.

U.S. demographic data are available for the twenty-two 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). With some help from the center it was feasible to come up with the distribution of out-of pocket payments for prescription drug expenditures, that is, prescription drug expenditures net of any prescription drug insurance coverage (private or public). NCHS also provided combined data on income and out-of-pocket prescription drug expenditures, which helped construct the joint distribution of income and prescription drug insurance used in this study. These data are important for allowing patient heterogeneity in prices and preference for branded drugs.[39] Finally, demographic data about the prevalence of depression are also available and are salient in estimating the correct welfare effects.

[38] Note that no year in the dataset includes all forty-seven antidepressants. This is due to the 'withdrawal' years.

[39] Some extentions to the data are being used in the subsequent chapters of my dissertation to improve on the results of this paper. These include sales data segmented into the various distribution channels (Health Maintenance Organizations (HMO), federal and non-federal hospitals, clinics, pharmacies, mail order, and long-term care) from 1995 onwards and data on advertising and detailing. The latter data are also available by drug, but only for the last six years of the dataset.

5.1.1 Market Size and the Outside Option

One needs to know the size of a market, M, in order to be able to estimate demand in that market. Given M, the observed output quantity of the firm producing product j is: $q_j = M \cdot s_j(x, \xi, p; \theta_d)$. In estimating demand, M could either be observed, for example, in BLP (1995) it is assumed to be the number of U.S. households, or unobserved and left as a parameter to be estimated. This study supposes that M is observed.

The question then becomes what is a valid M? By definition, M is the number of possible consumers in the antidepressant market. This depends on how one defines the market for antidepressant drugs. In the case of antidepressants, a good choice might be the portion of the population that is estimated to be clinically depressed, $\widetilde{M}$. However, as explained in Section 3.3.1, 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.

Finally, since doctors depend on their patients to confirm the existence of symptoms, people who are not clinically depressed may seek to purchase antidepressants for various other reasons like the treatment of mild anxiety disorders or entertainment. Information on the latter two cases is not widely available but assume their net effect to be small. Note the upward trend in the percentage of those clinically depressed that seek treatment between 1980 and 2002 in Figure 9b. $\widetilde{M}$ is thus better as it includes all those people that could be taking antidepressants but choose not to. Robustness tests evaluate the validity of this choice.

The existence of the outside option is also 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. Though data on the outside good would be welcome, Berry (1994) and BLP (1995) show that it is not necessary. They show that, given certain conditions (such as normalizing the mean utility level of the outside good, δ_0, to zero), one can determine the utility of the outside good.

5.2 Estimation[40]

In the multinomial logit model, the parameters are first estimated using an ordinary least-squares regression, assuming away price endogeneity. Then a two-stage least squares regression instrumenting for prices corrects for the correlation between prices and the unobserved drug characteristics:

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt}. \tag{25}$$

The nested multinomial logit models are also estimated with and without instruments. The estimation equations for the three models (NML1, NML2, NML3) are, respectively:

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt} + \rho_c \ln(\widetilde{s}_{j/c,t}) \tag{26}$$

[40] More detail on estimation and computation is available on my website.

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

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt} + \rho_{m_c} \ln(\widetilde{s}_{m_c/c,t}) + \gamma \ln(\widetilde{s}_{j/m_c,t}) \tag{27}$$

where the two naturally-occurring nested groups are assumed to be the 'type' and 'molecule' of the antidepressant and ρ_c is the within-molecule correlation coefficient of utility;

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x'_{jt}\beta + \xi_{jt} + \rho_b \ln(\widetilde{s}_{m_c/c,t}) + \gamma_1 \ln(\widetilde{s}_{b/m_c,t}) + \gamma_2 \ln(\widetilde{s}_{j/b,t}) \tag{28}$$

where the additionally naturally-occurring nesting is assumed to be the distinction between branded and generic drugs and ρ_b is the correlation coefficient of utility for branded or generic drugs within a molecule.

In the random coefficient model 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 takes the form:

$$\widehat{\theta} = \arg\min_{\theta} \omega(\theta)' Z A^{-1} Z' \omega(\theta), \tag{29}$$

where A is a consistent estimate of $E[Z'\omega\omega'Z]$. The unobserved product characteristics, defined as in Berry (1994), are computed by equating the estimated shares in equation (22), s, to the shares observed in the data, $\widetilde{s}$, and solving for mean utility levels. Unlike in the logit and nested logit models, no closed form solution exists. The estimation of the market shares as well as 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. Finally, equation (29) is solved using a non-linear search.[41]

5.2.1 Instrumental Variables

To correct for the above-mentioned price endogeneity, I need to specify variables that can act as instruments for price in the demand equations. Variables that are correlated with specific functions of the observed drug prices, but are not correlated with the unobserved demand disturbances, ξ_j, 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};

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

the time passed since generic entry took place in the same molecule. 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.[42]

5.2.2 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 the inclusion of insurance data, the estimated effect on utility of the various drug characteristics will be closer to their true values.[43] Given the ns draws of the observed and unobserved characteristics I average over the implied logit shares:

$$\sum_{i=1}^{ns} \frac{\exp\left(\delta_{jt} + \mu_{ijt}\right)}{1 + \sum_{k \in J_t} \exp\left(\delta_{kt} + \mu_{ikt}\right)} \tag{30}$$

Simulation draws are held constant as the parameters change otherwise changes in the objective function would be due to simulation changes.[44]

5.2.3 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. For reasons explained in section 5.1.1 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

[42] 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.

[43] In an extension of this paper, I use demographic variables, D_{ij}, that directly match patient i's characteristics to her choices which improves the model considerably. τ_{ij} can also be retrieved from survey data on observed patient side effects and their subsequent decisions.

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

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. It is ambiguous, therefore, what sign to expect for the coefficient of half-life.

In the nested logit models, ρ_c, ρ_{m_c}, ρ_b, γ, γ_1, γ_2 are correlation coefficients and have to be between zero and one. In NML1 ρ_c is the within-type utility correlation parameter and is the coefficient of the market share of drug j within type c. In NML2, ρ_{m_c} is the within-molecule utility correlation parameter and is the coefficient of the market share of drug j within type m_c; ρ_c is obtained by inserting the estimates for ρ_{m_c} and γ in $\gamma = [1 - (1 - \rho_{m_c})(1 - \rho_c)]$. In NML3, ρ_b is the correlation parameter of branded or generic drugs within the same molecule; it is the coefficient of the market share of drug j within a set of branded or generic drugs of the same molecule; ρ_c and ρ_{m_c} are obtained by solving the equations of γ_1, γ_2 simultaneously given the estimates for ρ_b, γ_1 and γ_2. The closer these correlation coefficients are to one the more valid the assumption that the nesting groups are naturally-occurring.

Finally, in the random coefficient model, additional information is obtained from the inclusion of unobserved and observed patient characteristics. This model estimates mean effects, the means of the distribution of marginal utilities (α's and β's in equation (20)), by a minimum-distance procedure.[45] It also estimates standard deviations (λ's in equation (20)) which are estimates of the unobserved heterogeneity about the mean effects. Finally, it estimates coefficients of demographic interactions with price and preference for 'brandness', that is, estimates of the observed heterogeneity about the mean effects (π's in equation (20)). 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 (20). This restricts the overall price sensitivity to non-positive values.

6 Results

6.1 Descriptive Statistics[46]

6.1.1 Drug Characteristics

Table 5a reports drug characteristics for selected antidepressants. Though physical characteristics do not change over time, the existence of a generic counterpart does. In fact, the time since generic entry is used as an instrument to correct for price endogeneity. As a generic enters and infiltrates the market, the price of the branded version of the drug is usually observed to increase at an increasingly decreasing fashion and at some point starts decreasing. Average antidepressant characteristics are reported in Table 5. The table provides means and standard deviations across the four types of antidepressants.

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

[45] See Nevo (2000a) for details.

[46] A continuously-updated online appendix on my website contains a more complete set of descriptive statistics and results, and includes a number of robustness checks.

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.

It is evident from the reported standard deviations that SSRIs and MAOIs have the lowest variance in side effects, if one excludes half-life. On the other hand, new generations have the highest variance and TCAs the second highest. This is expected since these types are more general groupings than the other two. As shown in Table 1, new generation antidepressants include drugs with varied antidepressant agents and TCAs include both tricyclic and tetracyclic molecules. This makes the case against using types as 'naturally-occurring' nests of the data as will also become evident from the estimation results on the nested logit correlation coefficients.

6.1.2 Market Shares

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. Table 6 reports annual market shares over time for the four antidepressant types. Table 6a reports revenue shares whereas Table 6b reports quantity shares. 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 14.51 percent of antidepressant sales, yet they only amass 1.19 percent 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.[47] 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 start rising again in the latter half of the 1990s when some of the newer drugs in the type are introduced. 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 went by 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.

Figures 3a and 3b show the diffusion of the different antidepressants into the market over the twenty-two year period. One can see how new generations first and SSRIs after displaced TCAs from

[47] This is partially explained by the fact that MAOIs have substantially different side effects, especially high fatality rates. This implies that depressed people who take MAOIs value these drugs for some reason other than their side effects, particularly, the way they fight against depression in the brain. It appears that their ability to inhibit monoamine oxidase enzymes from destroying the neurotransmitters that create emotions is valued highly by those patients that use them, high enough to cancel the disutility form the high fatality rates.

the market in Figure 3b and how TCA revenues were passed to these new types quicker than patients were passed in Figure 3a. 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 they started expanding again stealing both from TCAs and SSRIs.

Figure 4 presents the same picture as Figure 3 but divides the four antidepressant types into their respective molecules. While amitriptyline and doxepin are obviously the two dominant TCA molecules, the figure shows that the relative molecule shares within the type do not change by much. It is apparent that the shares of all TCAs decrease comparatively as the total share of TCAs, both for revenues and quantities, drops. In new generations, trazodone drives the initial increase in market shares but is overtaken by bupropion (generic Wellbutrin) as it enters and captures part of the new generation shares initially but shares most of the new generation shares eventually (latter half of the 1990s) with venlafaxine (generic Effexor). The introduction and diffusion of these molecules helped capture even part of the market share held by SSRIs. The utilization of SSRIs is largely driven by fluoxetine (generic Prozac), especially when it is the only molecule in the category, but also by other SSRI molecules such as sertraline (generic Zoloft) and paroxetine (generic Paxil) that capture a sizeable portion of the market share.[48][49]

6.1.3 Revenues, Quantities, Prices

To make more sense of the trends in market shares, it would also be useful to briefly look at trends in average revenues and quantities of the four antidepressant types and trends in total revenues, quantities and average prices for all antidepressants. Figure 1 shows the trend in total antidepressant quantities and revenues (both in current and constant 1980 dollars). Real revenues overtake the trend in quantities sold in the mid-1990s. This suggests escalating prices which are observed (on average) in Figure 2. These escalating antidepressant prices are not unlike the rising prices in all prescription drugs. Figure 11a shows the overall trend in real prescription costs in the U.S. and compares it to the trend in antidepressant costs.[50] The two trends seem to overlap. Figure 11b shows the overall prescription costs in per capita terms (over all U.S. population) and compares it to the per capita antidepressant costs (over the size of the antidepressant market). Both costs are increasing, however, antidepressant costs have been increasing in a quicker fashion. Though antidepressant costs per patient were lower than total prescription costs in the beginning of the period, they caught up total costs in 1995 and finally overtook them. This observation makes demand analysis in antidepressants even more important.

Figures 5, 6, 7, and 8 show trends in prices (Figures 5a, 6a, 7a, and 8a) and in quantities sold

[48] Figures 3, 4, and 5 show only relative shares of antidepressant drugs and exlude the outside option. It is true that the outside option has been decreasing exponentially over the years. This might mean that depressed people who switch from the outside option to pharmaceuticals might choose only among the newer drugs, hence the surge in market shares of SSRIs and new geneneration antidepressants.

[49] An online appendix figure presents shares for individual drugs grouped by their molecules as in Figure 4, which in turn are grouped by their type as in Figure 3. The importance of this figure is two-fold. On one hand, it shows the relative importance of individual drugs within their types and how these helped drive the market shares up or down. For instance, Prozac clearly determines the trend visible in SSRIs both in quantities but more so in revenues. On the other hand, it shows the relative importance of drugs within their respective molecules. It is apparent that lower-priced generic drugs captured a great proportion of their branded counterparts. This explains how TCA quantity shares decreased by much less than did TCA revenue shares. Also, the effect of the secondary branded drugs in a molecule becomes apparent. While this is only a recent fad, one can see that the quantity shares of drugs like Prozac Weekly, Sarafem, Wellbutrin SR, Remeron Soltab, Effexor XR largely displace the shares of the originating branded drugs. The effect is less pronounced in revenue shares immediately after introduction of the secondary brand takes place. Diffusion of the secondary brands into the market, though, boosts their revenue shares.

[50] Note that the trends are normalized so that it is easier to compare them.

(Figure 5b, 6b, 7b, and 8b) in TCAs, New Generations, SSRIs and MAOIs respectively. Figure 5 is devoted to TCAs where most generic competition took place. In general, one observes the falling or stable trends in the prices of generic drugs and the simultaneous rising trends in their quantities. At the same time as generic introduction takes place in a molecule and the surge in generic quantities begins, a falling and eventually dying trend is observed in the quantities of the branded counterparts is observed. For example, generic amitriptyline that has gradually swept the TCA market share, has been charging the lowest prices. Overall revenue shares of amitriptyline, though much smaller than quantity shares are still much larger than the revenue shares of other TCAs. At the same time as the quantity of generic amitriptyline increases after its 1977 introduction, the quantity of Elavil, the innovator brand, is dropping at an opposing rate. The secondary branded drug in this molecule, Endep, introduced only two years before the generic, suffers irreparable damage and withdraws from the market in 1989. Quantities for Endep thereafter stay lower than the minimum threshold reported by IMS. Trends in other TCAs are similar to amitriptyline's with generic competition driving out their respective brands. However, Figure 5a shows that prices for branded drugs are still rising slowly. This allows these drugs to maintain high revenues and when the fall in quantity is smaller than the fall in price rising revenues are observed. The higher prices are sustained by people who have a strong preference for branded over generic drugs and are willing to pay the premium.

Figure 6 is devoted to New Generations. With less drugs in this type it becomes easier to see the underlying trends. Figure 6b shows that quantities in all drugs are rising. The upward trend in quantities is slowed down when other molecules are introduced and quantities start decreasing when generic introduction takes place in the molecule. For example, the introduction of Effexor (venlafaxine) led to a decrease in the upward trend in the sales of Wellbutrin (bupropion) while the introduction of generic trazodone led to a decrease in the sales of Desyrel (trazodone). Also, the introduction of two secondary brands Wellbutrin SR and Effexor XR led to a dramatic fall in the quantities of their innovating brands. More important is the fact that when generic bupropion enters the market, the sales of Wellbutrin (bupropion-innovator) go down even further, whereas the sales of Wellbutrin SR remain steady. Figure 6a depicts the trends in the prices of these drugs and shows that, similarly to TCAs, prices of branded drugs do not fall when generic competition takes place, even when generic prices fall. Instead they still increase but a at a decreasing rate. Moreover, secondary brands are able to support high the high prices with which they entered the market.

Figure 7 shows prices and quantities of SSRIs. Figure 7b depicts the rising trend in the sales of Prozac which is affected by a couple of shocks from the introduction of other SSRIs, mainly Zoloft, Paxil and Celexa, and the introduction of its secondary brands in 2000 and 2001 (Sarafem and Prozac Weekly). Zoloft sales are unharmed by other branded introduction. However, all branded sales seem to suffer a decrease in their upward trend right after the introduction of fluvoxamine (generic Luvox), the first generic SSRI. The second generic in the market, fluoxetine (generic Prozac) was only introduced in 2002. Figure 7a shows that Zoloft's rising sales can be partially explained by its relatively low prices. Prozac's price has been increasing the most but it is only average compared to the price of high price of Luvox that also has the lowest SSRI sales. Generic fluvoxamine first enters the market with a price as high as Luvox but immediately this price takes a dive. Note that a drop in the rate of increase of sales of all SSRIs is already observed even with the introduction of generic fluvoxamine at such a high price. The explanation lies in the fact that a generic SSRI can be paid for even by strict prescription drug coverage plans that favor generics over branded drugs. The subsequent drop in SSRI sales' rates is explained by the dive in the price of the generic.

Figure 8, finally, depicts the prices and quantities of MAOIs. As already predicted by Table 6, Figure 8b shows that MAOI sales are stable on average, hence the relatively stable overall quantity shares for this type. At the same time, Figure 8a shows that MAOI prices have been rising proportionally to the rise in overall revenues in antidepressants, hence, the relatively stable overall revenue

shares for this type.

6.1.4 Patient Characteristics

Finally, it would be beneficial to examine the average trends in observed patient characteristics before the estimation results are presented. Figure 10a depicts normalized per capita real GDP and compares its rising trend to the rising trend of prescription drug costs per patient and covered prescription drug cost per patient. 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 $12,276 in 1980 and $17044 in 2001 and prescription drug expenditures per patient in 1980 dollars were $52.81 in 1980 and rose to $239.27 in 2001. Respectively, antidepressant expenditures per patient were $11.29 in 1980 and $385.53 in 2001.[51] Out of these expenditures, 30.6 percent was covered by a prescription plan in 1980 and 66.1 percent in 2001.

6.2 Demand Estimation

Tables 7, 8, and 9 report the demand estimation results.[52] Table 7 reports the results for four different models: the simple logit model and three nested multinomial logit models described in section 4, NML1, NML2, NML3. Two versions of each model were estimated: an OLS version and an IV version correcting for price endogeneity. Looking at the OLS results, estimated price sensitivity increases when moving from the logit model to NML1 and decreases from NML1 to NML3. When moving from each OLS model to its IV counterpart, estimated price sensitivity more than triples. In fact it is six times as big in the NML2 model. This implies that correction for price endogeneity is important and necessary. The disutility for price was correctly detected and is high; it ranges from -1.6 to -2.6.

In general, statistical significance in coefficients increases when moving from the OLS logit models rightward and up to the NML2 models, and decreases when moving to the NML3 models. This means that market segmentation, and the additional structure it imposes in NML1 and NML2, improves on the model's prediction. However, the further restriction imposed by NML3 worsens the model. The signs of five out of eight estimates of side effect coefficients are negative as expected and statistically significant for all eight models. The magnitude of the fatality rate coefficient is largest in the NML2 model. Also, moving from OLS to IV estimation the coefficient increases. Similar results are obtained for four other adverse side-effect coefficients: drowsiness, insomnia and agitation, cardiac arrhythmias and weight gain. Looking at the IV-NML2 model, the disutility is highest for cardiac arrhythmias and lowest for weight gain.

Two side effects, anti-cholinergic side effects and orthostatic hypotension, have small positive signs, 80 percent of which are significant. There is no immediate explanation behind these estimates, which means that on average patients prefer drugs with these characteristics because they dislike them 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 and constipation so that on average their choices predict a small positive utility derived from the occurrence of these side effects. Similarly for orthostatic hypotension. The latter, however, is a serious adverse effect. It is possible though to take additional medication to treat hypotension and patients are not link its occurrence to antidepressant medication. Gastrointestinal (GI) distress is sometimes obtained as positive and significant and other times as negative and significant.

[51] Overall prescription costs are divided by the total population to get per capita values. Antidepressant costs are divided by the market size.

[52] Additional estimation results with alterations to the main model and robustness checks are available online.

The coefficients on half-life are only significant for NML2 and IV logit and are all small and positive. 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 results show that the hassle effect won over the severe side-effects effect. However, as will be discussed subsequently, the random coefficient logit results find a negative coefficient. The average dosing frequency coefficient is obtained always negative as expected and mostly significant. The high disutility obtained from having to take a drug in frequent repetitions supports the positive result in the coefficient of half-life. That is, people place great value on frequency that they prefer drugs with lower half-lives as well.

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 mostly 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.

To evaluate the different nested multinomial logit models, one has to look at the estimated correlation coefficients for the different groupings, the ρ's. All correlation coefficients are between zero and one as expected. Type correlation is small for all three models and insignificant for the NML1 (both OLS and IV) model which in fact uses type as the only naturally-occurring nesting. The validity of this grouping is shaken. In fact, the analysis of the different characteristics of antidepressants within the same type showed that for TCAs and especially new generations were highly variable. Note, though, that the coefficients are larger and significant in the NML2 case. Here, the molecule correlation coefficients are both significant and bigger than the type coefficients, yet quite small. Correlation coefficients in NML3 are much larger especially for the third nesting group of branded drugs within a molecule and generic drugs within a molecule. In fact, this would be expected to be even closer to one since in the majority of the cases a single drug it compared to itself.

The low and sometimes insignificant coefficients are reason to favor the full random coefficient model which places no restriction on the correlation between antidepressants. These eight models all restrict random coefficients (reported in Table 9) to be zero. The observed high significance of most of the estimated random coefficients leads Wald tests to favor the random coefficient models (both with and without demographics) over the more restrictive specifications of the models in Table 7. Dummy variables for type, molecule and 'brandness' are included as drug characteristics in the RCL model to capture any correlation that may exist. Table 8 presents the results of the IV logit, IV-NML2 and RCL models together. The results of the latter are given both when no demographics are used and when demographic information is incorporated. The third and fourth columns of Table 8 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 9.

Three of the mean coefficients of the RCL model without demographics (third column of Table 8) 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.

Table 9 presents results for the full random coefficients model using demographic income and prescription drug insurance described above. 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. 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, for example, now has a negative coefficient, though it is small and has a relatively high variance. This means that, when variability in patient preferences was allowed, 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 (20), given assumption A9 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 percent) 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.[53] 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.[54]

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 is 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 percent) in patient preference for brandness is explained by the observed demographics. Table 10 shows the combined effect of demographics on price and brand sensitivity. Wealthier patients who also have prescription coverage are almost insensitive to changes in the price of a drug and have a higher preference for branded drugs. Poorer 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.

Table 11 presents own- and cross-price elasticities of demand for selected antidepressants.[55] These are weighted averages for the years 2000 and 2001.[56] The reported cross-price elasticities are averaged over drugs of the same type. The selection includes drug of two TCA molecules and one new generation antidepressant 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 new generations and MAOIs and much less to TCAs.

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.[57] I then follow the suggestion by Nevo (2000a) 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 11 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.

[53] 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.

[54] This is discussed in Section 6.3.

[55] Tables including all own- and cross-price elasticities of demand for more years are included online.

[56] Simple averages when the drug exists in both years, weighted averages when the drug exists only one of the years.

[57] Hausman and McFadden (1984)

6.3 Welfare Implications[58]

The underlying assumption for a demand-based assessment of patient welfare is that consumer surplus can be measured by the revealed preferences of consumers through their observed choices. 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:

$$W_{jt} = \int_{i=1}^{ns} -\frac{1}{\alpha_i} \int_{p_j}^{\infty} s_{ijt}\left(q_j | q_k = p_k \forall k < j, q_k = \infty \forall k > j\right) dq_j dF(\alpha_i, \sigma_\alpha) \tag{31}$$

where each s_{ijt} is computed using the estimated parameters as in equation (23) 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 (31) 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.

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

$$W_t = \sum_{j=0}^{J_t} W_{jt}. \tag{32}$$

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,

[58] This section borrows from the analysis in Trajtenberg (1990) and Ellickson, Stern and Trajtenberg (1999).

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.[59]

6.3.1 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. 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.

Table 12 presents the welfare estimates ($\underline{INN}_{jt}\ \forall j$) of all 28 antidepressant innovations that took place between 1980 and 2001.[60] Shaded rows in the table represent introduction of an antidepressant type and rows in italics represent molecule introductions. Estimated patient welfare for patients with and without insurance is presented in Table 12a 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 study 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.[61] 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 12a shows the excess willingness-to-pay accrued annually. In other words, an individual patient would be willing to pay \$8928.90 in a year over the amount already spent to be able to use Prozac. Moving to Table 12b, the last column depicts the ratio of this per unit patient surplus on price.[62] 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.

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 12b present rankings of the four columns of patient welfare gains in Table 12a. 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. The introduction of new generation antidepressants with the advent of Desyrel [trazodone] was not as successful.

[59] 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.

[60] 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.

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

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

The overall patient surplus ranks in the middle of all antidepressant innovations. Desyrel's per unit valuation and when taking insurance into account diminishes the relative importance of its 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 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 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 study. 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.

7 Concluding Remarks

7.1 Summary & Discussion

In this study, I formulate an empirical methodology that quantifies patient welfare benefits from pharmaceutical innovation in the U.S. antidepressant market. The study 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 study 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 of my job market paper uses monthly pharmaceutical data from 1996

to 2001 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 incorporates a more detailed analysis of the workings of the antidepressant market, offers estimation of welfare implications of the different functional forms of demand for antidepressants presented in this paper (multinomial logit and different levels of nested logit) and compares their performance. One more functional form of demand will be considered: the no-epsilon model. The problem with ϵ_{ij} is that a new characteristic is added every time there is introduction in the market. Any product has positive sales and all products are substitutes. Extension to the no ϵ_{ij} model eliminates the random utility component and product substitution is towards any product in the choice set.

8 References

Anderson, Simon-P, Andre de Palma and Jacques-Francois Thisse, 1992. *Discrete Choice Theory of Product Differentiation,* Cambridge, Massachusetts: Harvard University Press.

Approved Drug Products with Therapeutic Equivalence Evaluations. (*The Orange Book*), 22^{nd} Edition, 2002. 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.

Berndt, Ernst R. Zvi Griliches and Joshua Rosett, 1993. "On the Accuracy of Producer Price indices for Pharmaceutical Preparations," NBER Working Paper: 3490.

Berndt, Ernst R., Zvi Griliches and Joshua Rosett, 1993. "Auditing the Producer Price Index: Micro Evidence from Prescription Pharmaceutical Preparations," Journal of Business and Economic Statistics, 11(3): 251-64.

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).

Berry, Steven T., Michael Carnall and Pablo Spiller, 1996. "Airline Hubs: Costs, Markups and the Implications of Customer Heterogeneity," NBER Working Paper: 5561.

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

Berry, Steven T., James Levinsohn and Ariel Pakes, 1998. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," NBER Working Paper: 6481.

Berry, Steven T., James Levinsohn and Ariel Pakes, 1999. "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy," American Economic Review, 89(3): 400-430.

Bresnahan, Timothy F., 1987. "Competition and Collusion in the American Automobile Industry: The 1955 Price War," *The Journal of Industrial Economics*, Vol. XXXV, June.

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

Cardell, Scott, 1997. "Variance Components Structures for the Extreme Value and Logistic Distributions with Applications to Models of Heterogeneity," *Econometric Theory*, 13: 185-213.

Caves, Richard E., Michael D. Whinston and Mark A. Hurwitz, 1991. "Patent Expiration, Entry and Competition in the U.S. Pharmaceutical Industry," Brookings Papers on Economic Activity. Microeconomics: 1-48.

Cleanthous, Paris, 2002. "Analyzing the Effects of Marketing Efforts by Pharmaceuticals on Patient Welfare," Yale University.

Currie, Gillian, 1998. "Pricing, Advertising and Entry in the Market for Antidepressant Drugs," Yale University Ph.D.

Davis, Peter, 2000. "Empirical Models of Demand for Differentiated Products," European Economic Review, 44(4-6): 993-1005.

Depression Guideline Panel, 2002. Web Publication, U.S. Department of Health and Human Services, Public Health Service.

Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), 2002. American Psychiatric Association.

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

Lacy, Charles F., Lora L. Armstrong and Morton P. Goldman, 2002. *Drug Information Handbook,* Web Publication, Lexi-Comp.

Eaton, B. Curtis, and Richard G. Lipsey, 1989. "Product Differentiation," in Handbook of Industrial Organization, 1(10), ed. Richard Schmalensee and Robert D. Willig: 723-768. Amsterdam: North Holland.

Ellickson, Paul, Scott Stern and Manuel Trajtenberg, 1999. "Patient Welfare and Patient Compliance: An Empirical Framework for Measuring the Benefits from Pharmaceutical Innovation," NBER Working Paper: 6890.

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, 2001. "Countervailing Power in Wholesale Pharmaceuticals," manuscript M.I.T.

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

Goolsbee, Austan and Amil Petrin., 2002. "The Consumer Gains from Direct Broadcast Satellite and the Competition with Cable TV," manuscript University of Chicago, GSB.

Griliches, Zvi and Iain Cockburn, 1994. "Generics and New Goods in Pharmaceutical Price Indexes," *American Economic Review*, 84(5): 1213-32.

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.

Kessler, D. A. et al., 1994a. "Therapeutic-Class Wars: Drug Promotion in a Competitive Market Place," *New England Journal of Medicine,* 331(20): 1350-1353.

Kessler, D. A. et al., 1994b. "Lifetime and 12-Month Prevalence of DSM-III-R psychiatric orders in the U.S.," *Arch Gen Psychiatry,* 51:8-19.

Lancaster, Kelvin, 1966. "A New Approach to Consumer Theory," *Journal of Political Economy*, 74: 132-157.

Lancaster, Kelvin, 1977. *Consumer Demand: A New Approach*, New York: Columbia University Press.

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

Lichtenberg, Frank R., 1996b. "The Effect of Pharmaceutical Utilization and Innovation On Hospitalization and Mortality," NBER Working Paper: 5418.

Lichtenberg, Frank R., 1998a. "Pharmaceutical Innovation, Mortality Reduction and Economic Growth," NBER Working Paper: 6569.

Lichtenberg, Frank R., 1998b. "The Allocation of Publicly-Funded Biomedical Research," NBER Working Paper: 6601.

Lichtenberg, Frank R., 2001. "The Benefits and Costs of Newer Drugs: Evidence from the 1996 Medical Expenditure Panel Survey," NBER Working Paper: 8147.

McFadden, Daniel, 1973. "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, eds., *Frontiers of Econometrics*, New York, Academic Press.

McFadden, Daniel, 1978. "Modeling the Choice of Residential Location," in A. Karlgvist, et al., eds., *Spatial Interaction Theory and Planning Models*. Amsterdam: North-Holland.

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

Murphy, Kevin M. and Robert Topel, 1999. "The Economic Value of Medical Research," Working Paper, University of Chicago, Graduate School of Business.

Nevo, Aviv, 2000a. "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, 2000b. "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *Rand Journal of Economics*, 31(3): 395-421.

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, 1995. Depression: Questions you have. Answers you need. People's Medical Society Publications.

Scherer, F. M., 1996. Industry Structure, Strategy, and Public Policy. Ch. 9: 336-390. United States: HarperCollins.

Scott Morton, Fiona, 1997. "The Strategic Response by Pharmaceutical Firms to the Medicaid Most-Favored-Customer Rules," *Rand Journal of Economics*, 28(2): 269-90.

Scott Morton, Fiona, 1999. "Entry Decisions in the Generic Pharmaceutical Industry," *Rand Journal of Economics*, 30(3): 421-40.

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

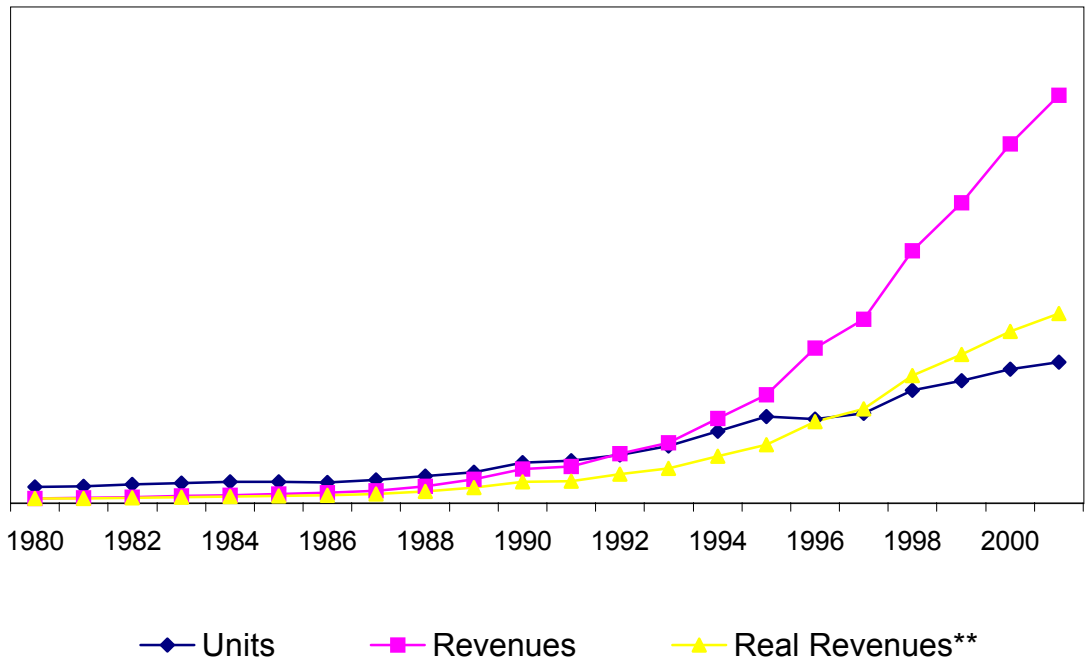
Suslow, Valerie. 1992. "Are There Better Ways to Spell Relief: A Hedonic Pricing Analysis of Ulcer Drugs," School of Business Administration, University of Michigan, Working Paper: 696.

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.

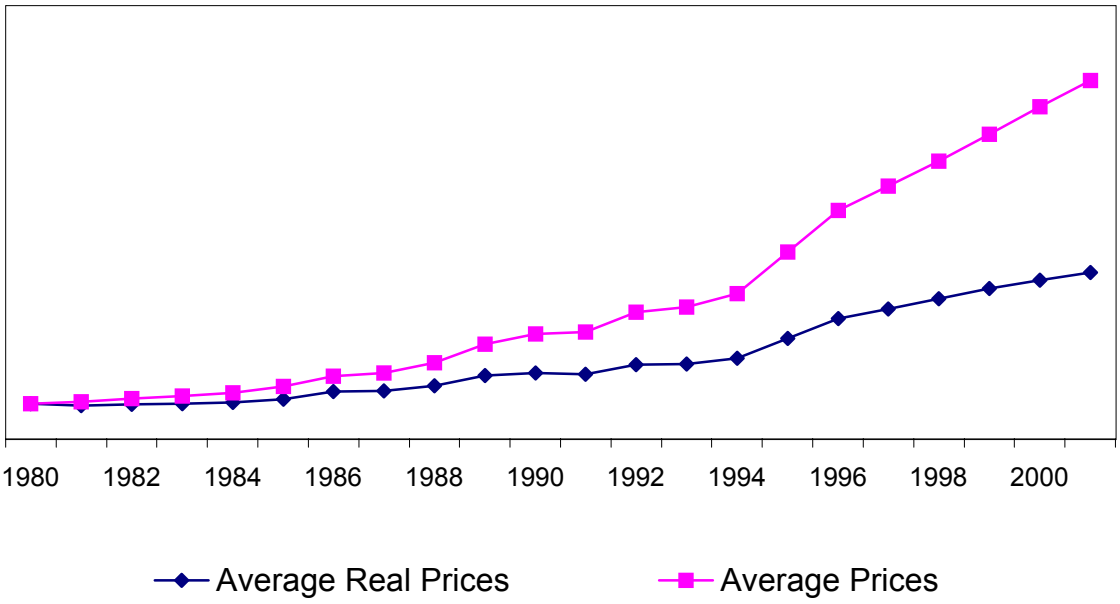
Wosińska, Marta, 2002. "Promoting to Multiple Agents: The Case of Direct-to-Consumer Drug Advertising," manuscript Harvard Business School.

Figure 1: Antidepressant Market Units & Revenues (1980-2001)*



Source: IMS Health, Inc.
* Daily dosage units
** Deflated to 1980 ('000s) dollars using GDP deflators (BLS)

Figure 2: Antidepressant Market Average Nominal & Real* Prices (1980-2001)



Source: IMS Health, Inc.
* Deflated to 1980 ('000s) dollars using GDP deflators (BLS)

Figure 3a: Annual Revenue Shares of Types (%)

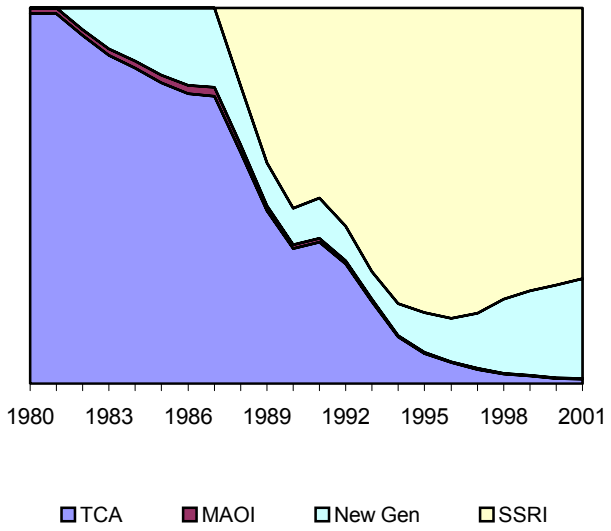


Figure 3b: Annual Quantity Shares of Types (%)

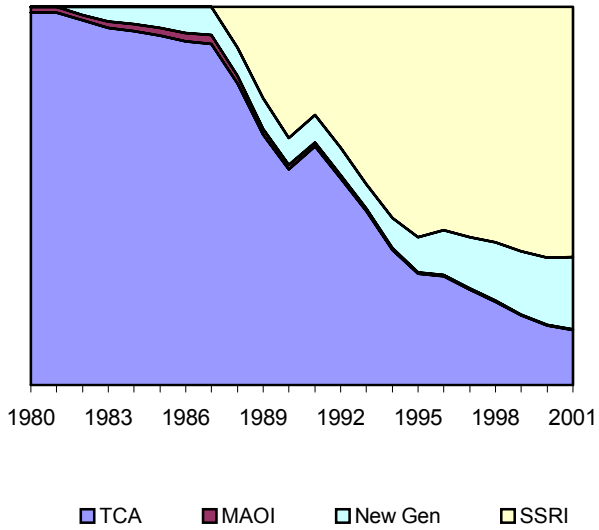


Figure 4a: Annual Overall Revenue Shares of Molecules (%)

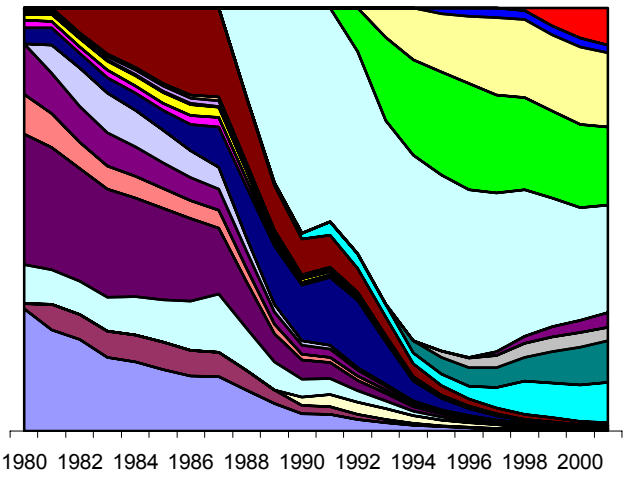
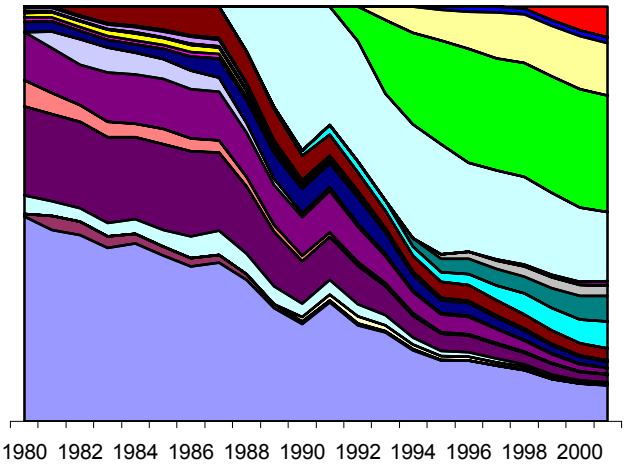
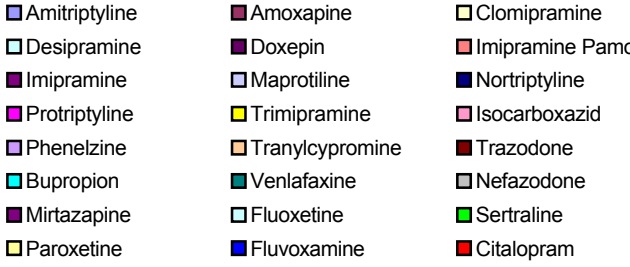


Figure 4b: Annual Overall Quantity Shares of Molecules (%)





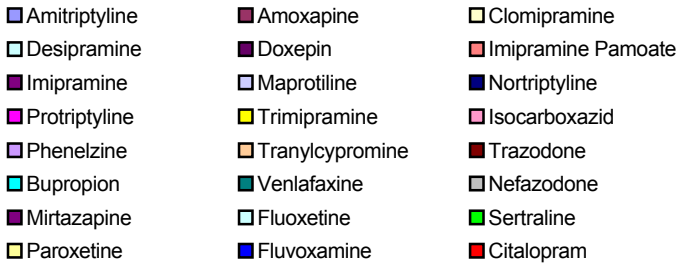


Figure 5a: TCA Prices (1980-2001)

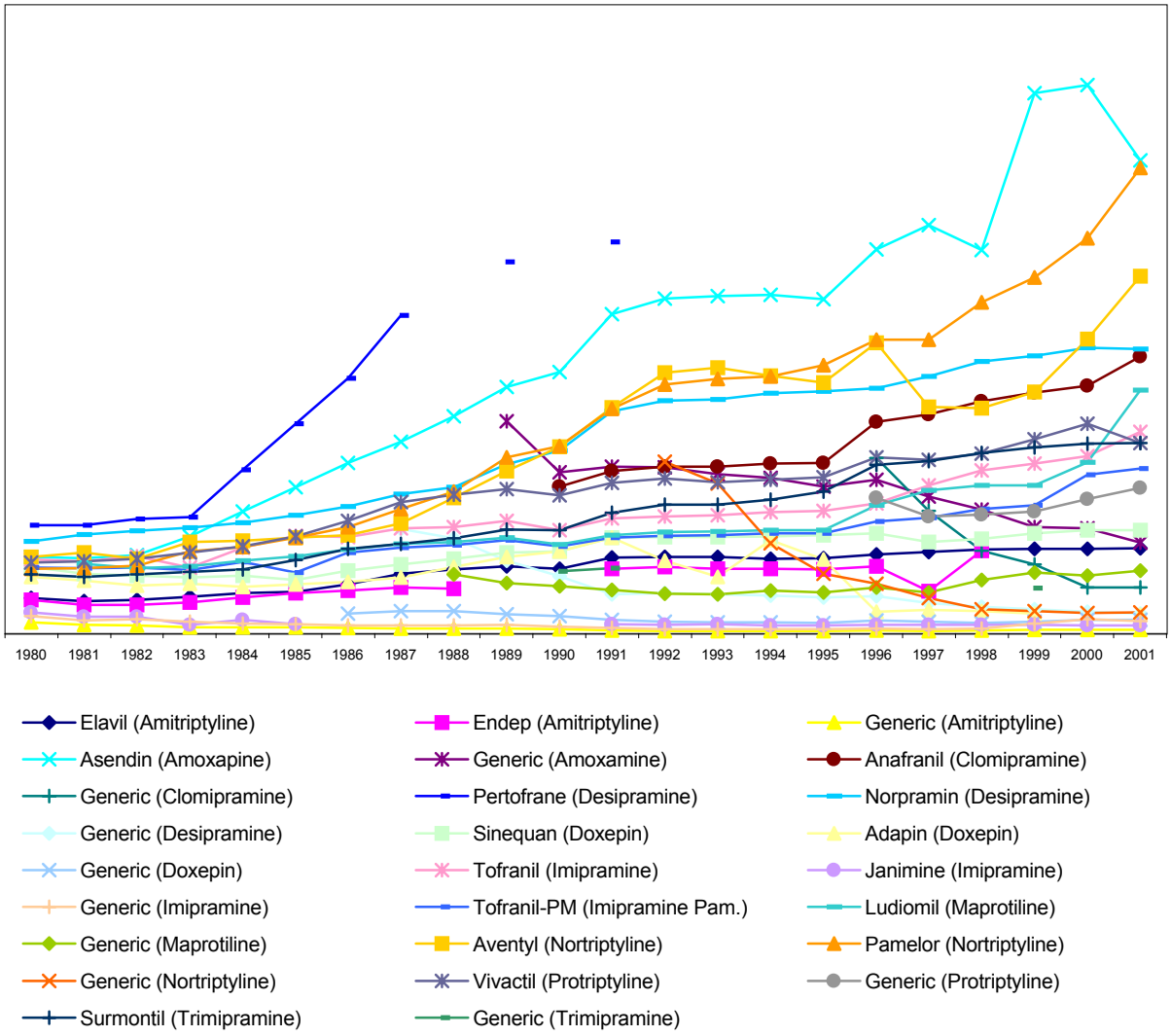


Figure 5b: TCA Quantities (1980-2001)

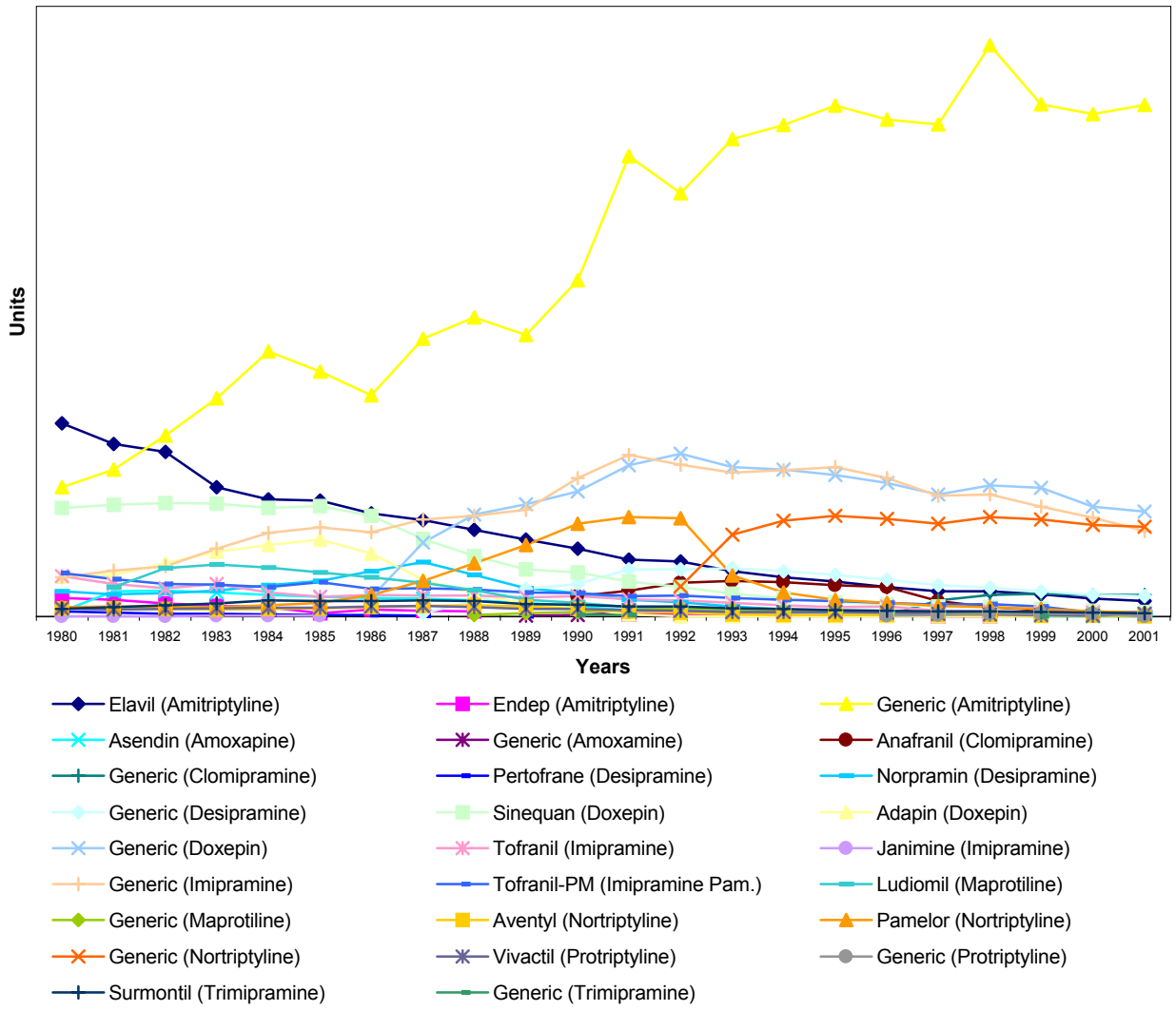


Figure 6a: New Generation Prices (1980-2001)

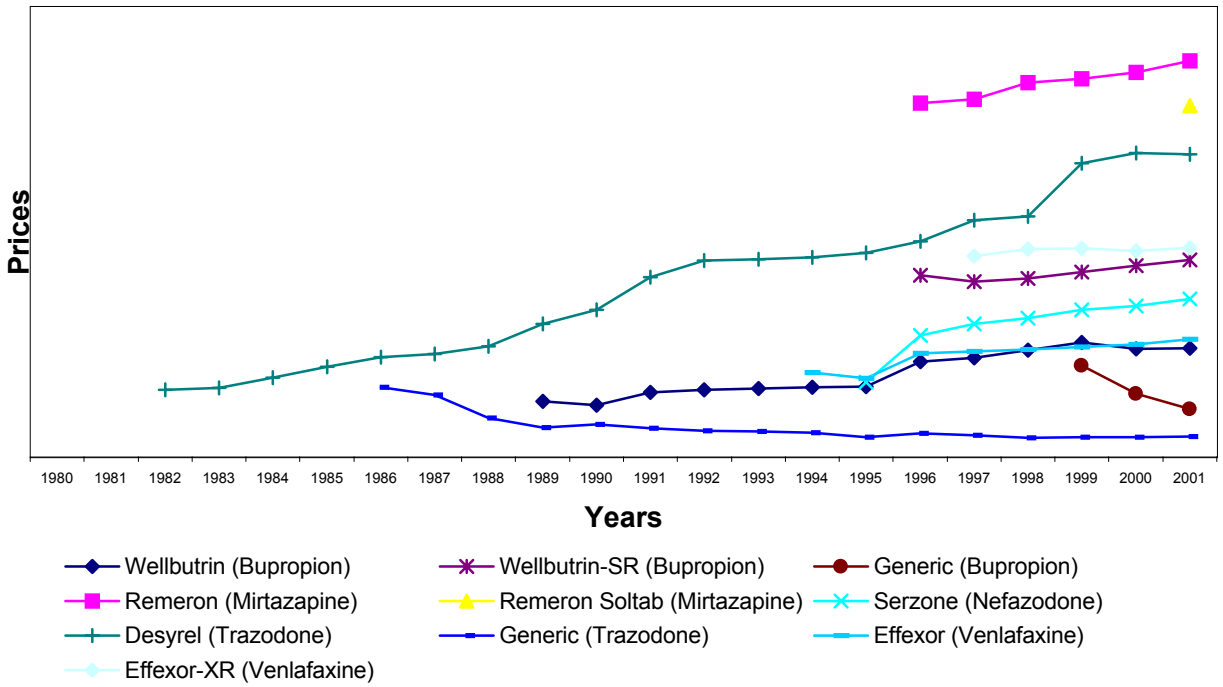


Figure 6b: New Generation Quantities (1980-2001)

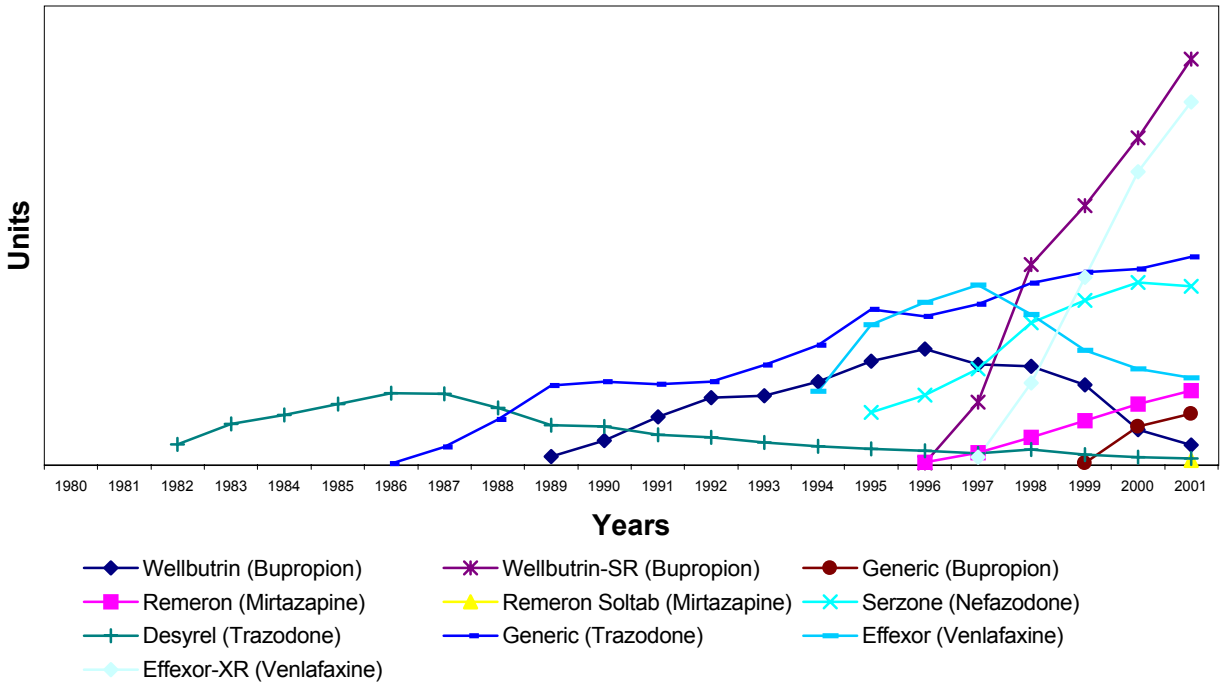


Figure 7a: SSRI Prices (1980-2001)

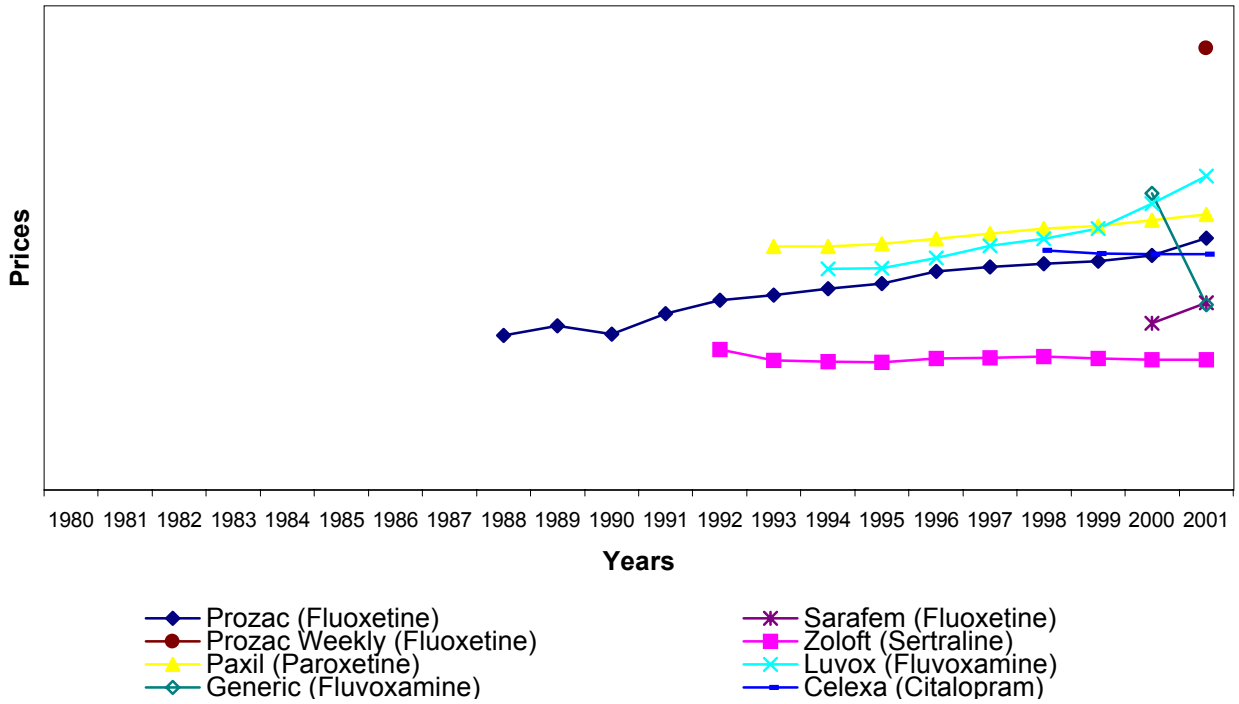


Figure 7b: SSRI Quantities (1980-2001)

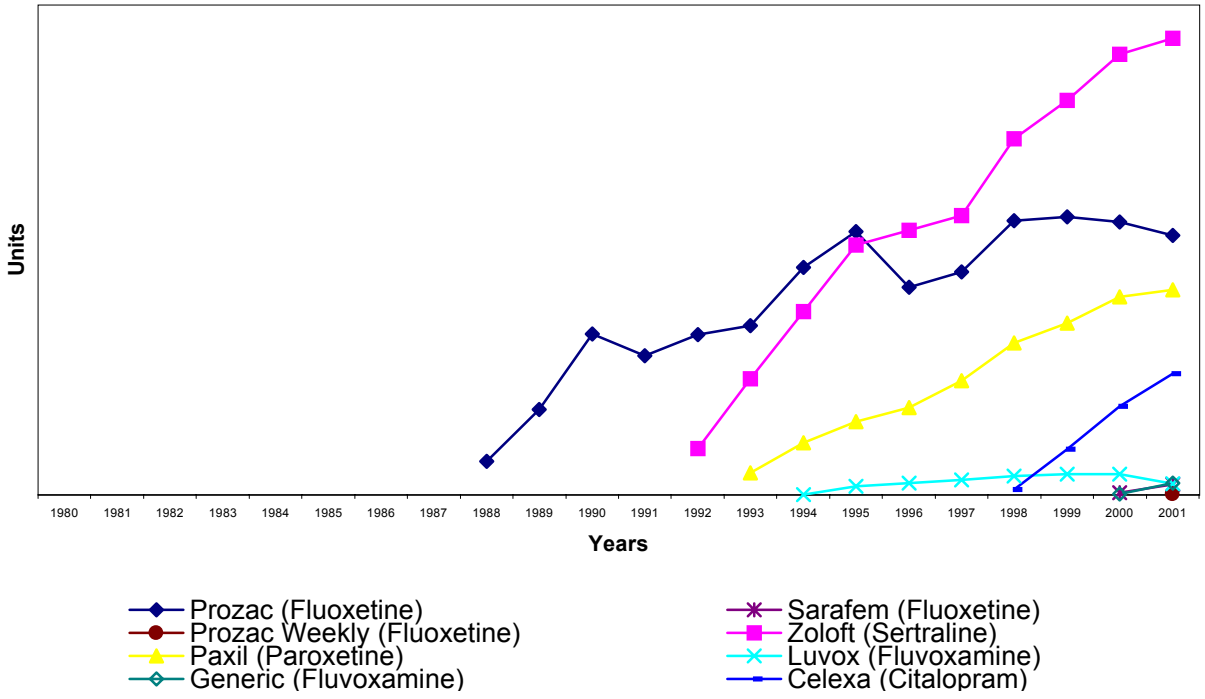


Figure 8a: MAOI Prices (1980-2001)

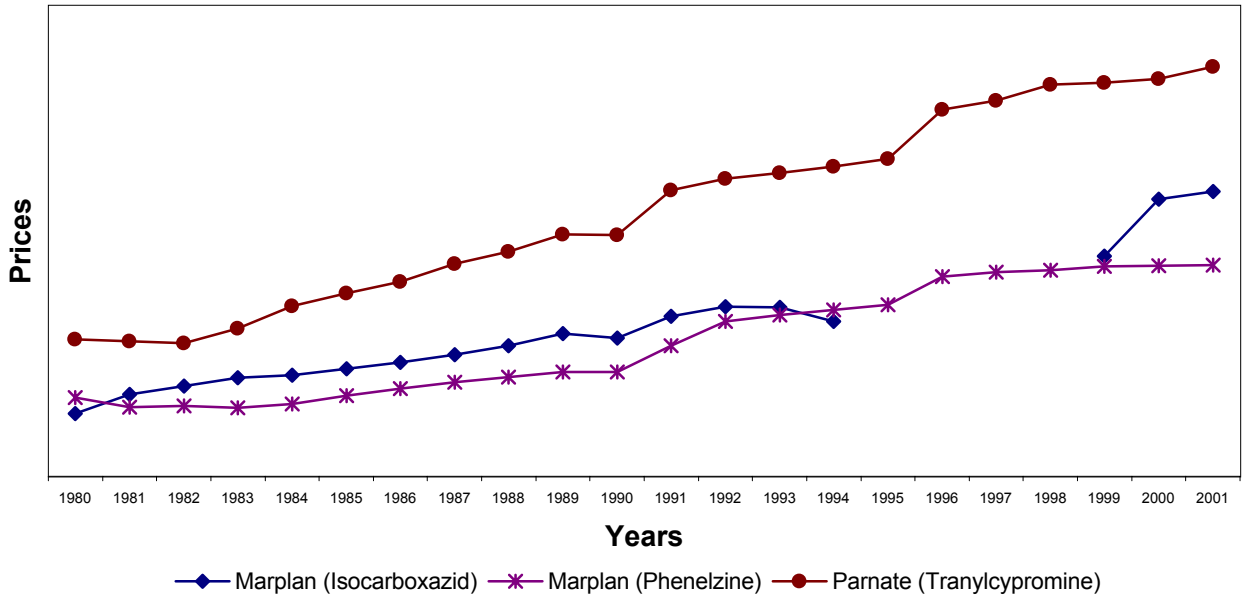


Figure 8b: MAOI Quantities (1980-2001)

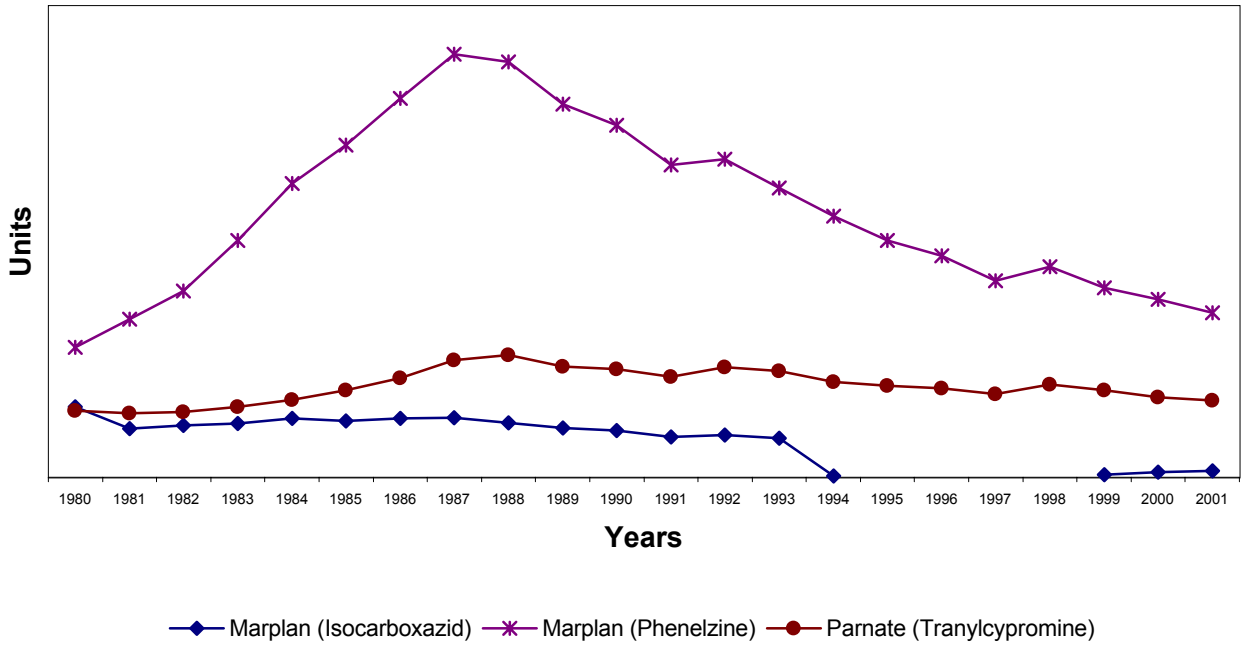


Figure 9a: Market Size

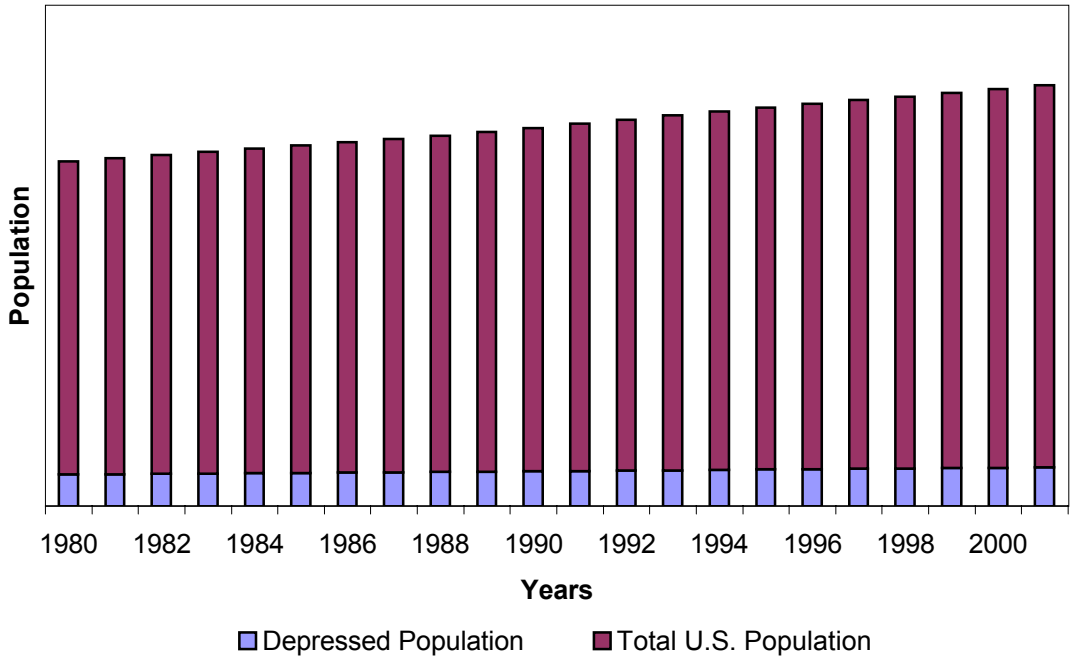


Figure 9b: Clinically depressed population that seeks pharmacological treatment

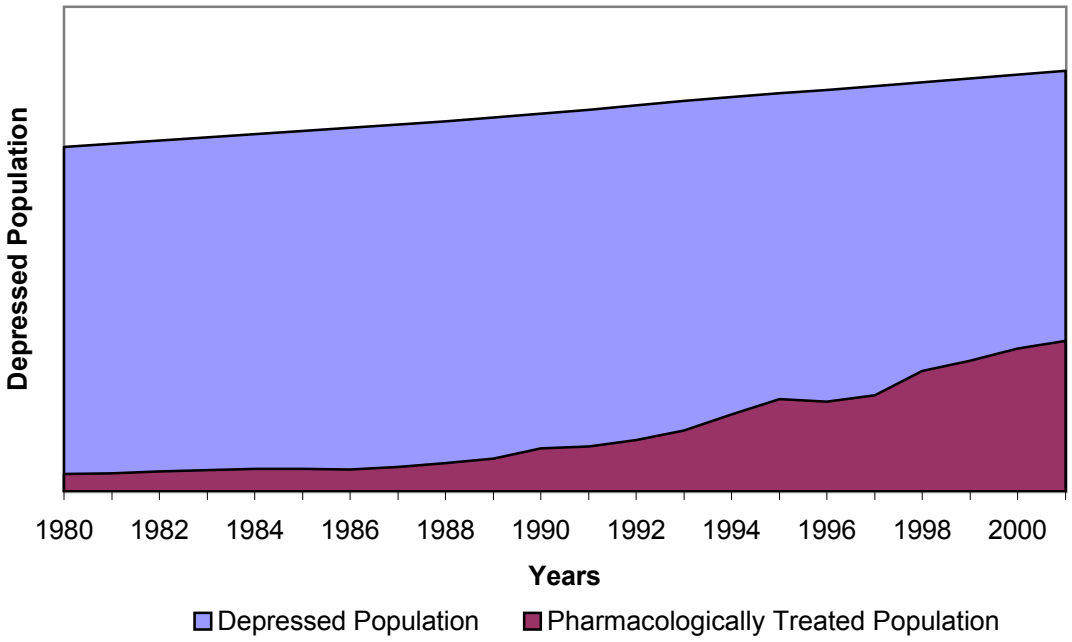


Figure 10a: Per Capita U.S. Income & Prescription Costs

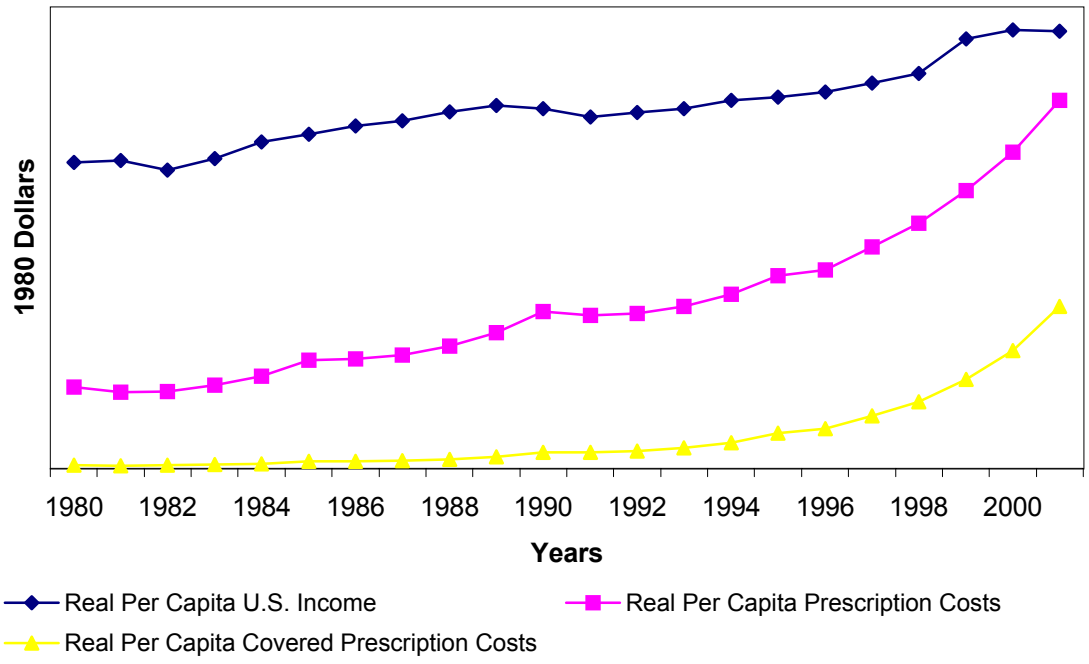


Figure 10b: Prescription Drug Insurance (All Patients)

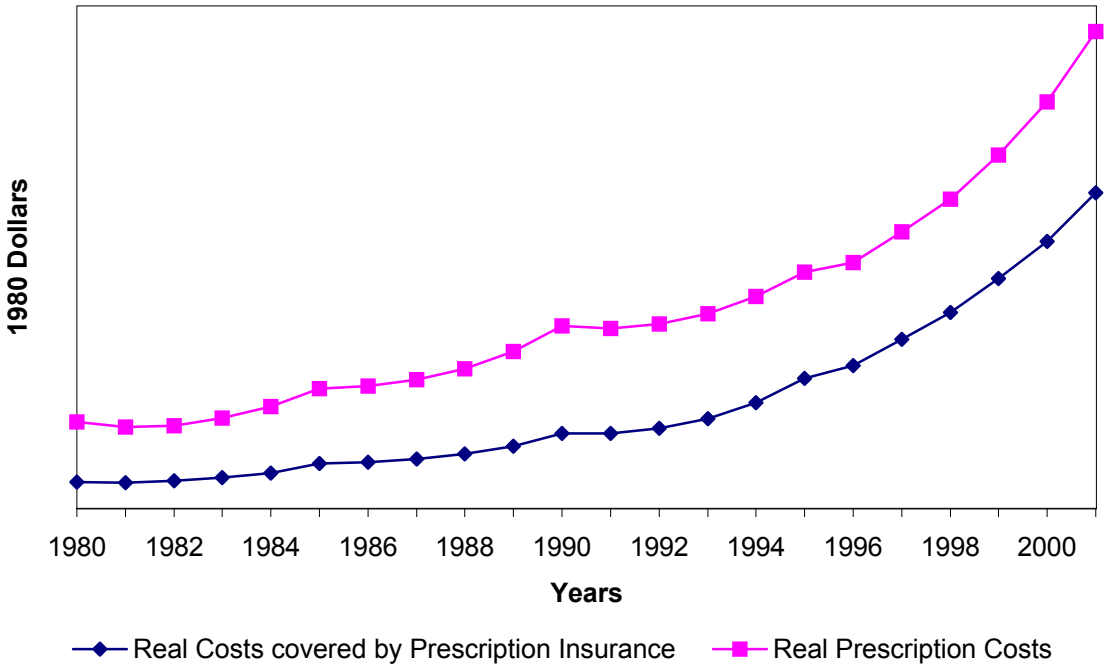


Figure 11a: Prescription Costs (All Patients)

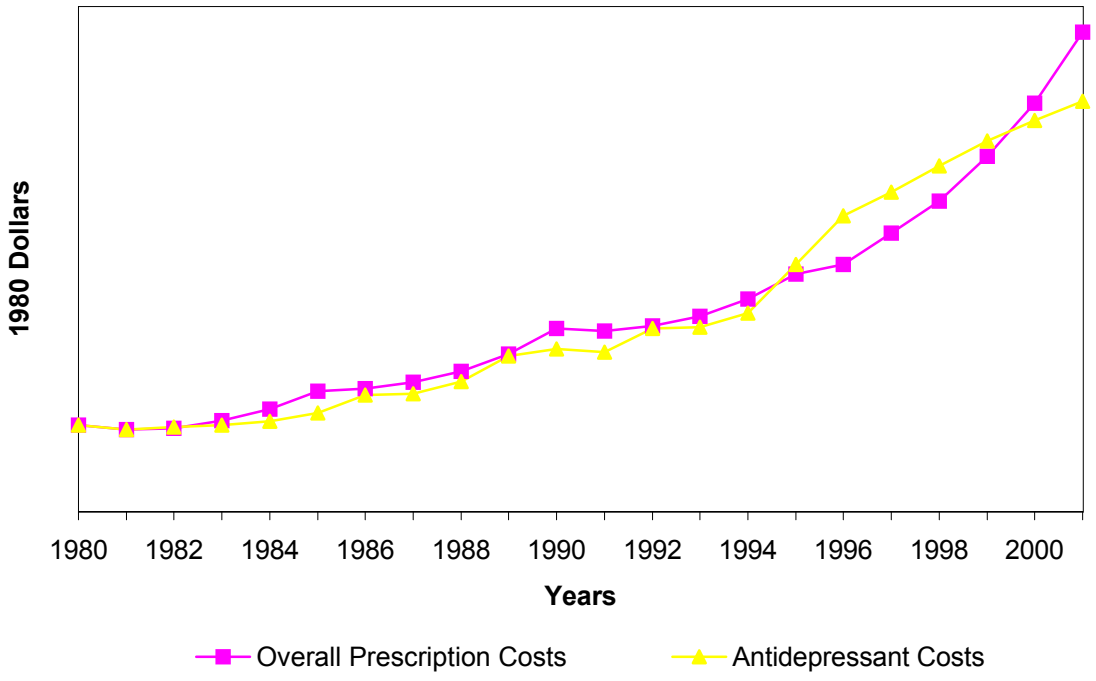


Figure 11b: Prescription Costs Per Patient

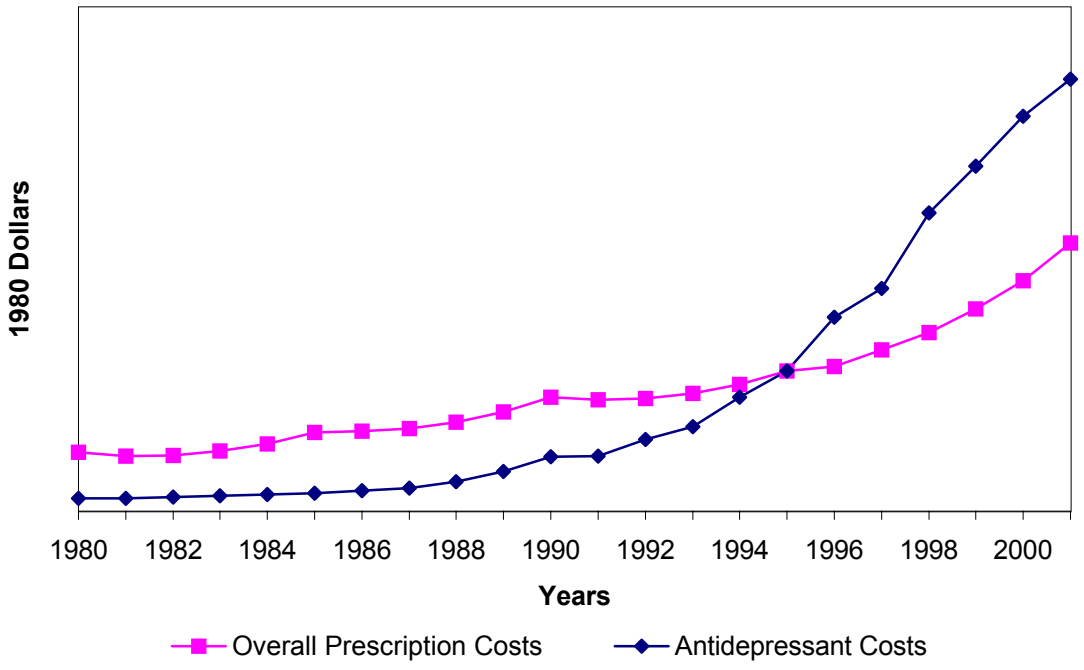


Table 1

ANTIDEPRESSANT AGENTS

Comparison of Usual Dosage, Mechanism of Action, and Side Effects of Antidepressants													
Drug	Initial Dose	Usual Dosage (mg/d)	Dosage Forms	Average Dosing Frequency	Half-Life	Side Effects							
						Fatality	Anti-Cholinergic	Drowsiness	Insomnia/ Agitation	Orthostatic Hypotension	Cardiac Arrhythmias	GI Distress	Weight Gain
				FREQ	HALF	FAT	AC	DR	IA	OH	CA	GID	WTG
MONOAMINE OXIDASE INHIBITORS (MAOI)													
Isocarboxazid (Marplan®)	20 mg	20	T	1	2	4	1	1	2	2	0	1	2
Phenelzine (Nardil®)	15 mg	15-90	T	1	2	4	2+	2+	2	2+	1+	1+	3+
Tranylcypromine (Parnate®)	10 mg	10-60	T	3	2	4	2+	1+	2	2+	1+	1+	2+
TRICYCLIC ANTIDEPRESSANTS & RELATED COMPOUNDS (TCA)													
Amitriptyline (Elavil®, Endep®)	25-75 mg	100-300	T, I	1	24	2	4+	4+	0	4+	3+	1	4+
Amoxapine (Asendin®)	50 mg	100-400	T	1	10	1	2+	2+	2	2+	2+	0	2+
Clomipramine (Anafranil®)	25-75 mg	100-250	C	1	24	2	4+	4+	1	2+	3+	1+	4+
Desipramine (Norpramin®, Pertofrane®)	25-75 mg	100-300	T	1	18	2	1+	2+	1	2+	2+	0	1+
Doxepin (Adapin®, Sinequan®)	25-75 mg	100-300	C, L	1	17	2	3+	4+	0	2+	2+	0	4+
Imipramine (Janimine®, Tofranil®)	25-75 mg	100-300	T, C, I	2	22	2	3+	3+	1	4+	3+	1+	4+
Imipramine Pamoate (Tofranil-PM®)	75 mg	300	T, C, I	1	22	2	3+	3+	1	4+	3+	1+	4+
Maprotiline (Ludiomil®)	25-75 mg	100-225	T	1	43	1	2+	3+	0	2+	2+	0	2+
Nortriptyline (Aventyl®, Pamelor®)	25-50 mg	50-150	C, L	1	26	2	2+	2+	0	1+	2+	0	1+
Protriptyline (Vivactil®)	15 mg	15-60	T	4	76	2	2+	1+	1	2+	3+	0	0
Trimipramine (Surmontil®)	25-50 mg	100-300	C	1	12	2	4+	4+	0	3+	3+	0	4+

Table 1 (continued)
ANTIDEPRESSANT AGENTS

Comparison of Usual Dosage, Mechanism of Action, and Side Effects of Antidepressants													
Drug	Initial Dose	Usual Dosage (mg/d)	Dosage Forms	Average Dosing Frequency	Half-Life	Side Effects							
						Fatality	Anti-Cholinergic	Drowsiness	Insomnia/ Agitation	Orthostatic Hypotension	Cardiac Arrhythmias	GI Distress	Weight Gain
				FREQ	HALF	FAT	AC	DR	IA	OH	CA	GID	WTG
SELECTIVE SEROTONIN REUPTAKE INHIBITORS (SSRI)*													
Citalopram (Celexa™)	20 mg	20-60	T	1	37	1	0	0	2	0	0	3+**	1+
Fluoxetine (Prozac®, Sarafem™, Prozac Weekly)	10-20 mg	20-80	C, L, T	1	168	1	0	0	2	0	0	3+**	1+
Fluvoxamine (Luvox®)	50 mg	100-300	T	1	15	1	0	0	3	0	0	3+**	1+
Paroxetine (Paxil™)	10-20 mg	20-50	T, L	1	24	1	1+	1+	2	0	0	3+**	1+
Sertraline (Zoloft™)	25-50 mg	50-150	T	1	22	2	0	0	2	0	0	3+**	1+
NEW GENERATION ANTIDEPRESSANTS DOPAMINE-REUPTAKE BLOCKING COMPOUNDS													
Bupropion (Wellbutrin®, Wellbutrin SR®)	100-150 mg	300-450	T	3	14	1	0	0	2	0	1+	1+	0
SEROTONIN/NOREPINEPHRINE REUPTAKE INHIBITORS													
Venlafaxine (Effexor®, Effexor-XR®)	25 -37.5 mg	75-375	T	3	8	1	1+	1+	2	0	1+	3+	0
5HT2 RECEPTOR ANTAGONIST PROPERTIES													
Nefazodone (Serzone®)	100 mg	300-600	T	2	3	1	1+	1+	1	0	0	1+	0
Trazodone (Desyrel®)	50 mg	150-600	T	3	8	1	0	4+	0	3+	1+	1+	2+
NORADRENERGIC ANTAGONIST													
Mirtazapine (Remeron®, Remeron-Soltab®)	15 mg	15-45	T	1	32	1	1+	3+	2	0	0	0	3+

Sources: Depression Guideline Panel (1993), Physician's Desk Reference Generics (1996), Drug Information Handbook (2001).

* Flat dose response curve, headache, nausea, and sexual dysfunction are common side effects for SSRIs
** Nausea is usually mild and transient.

Key:
ACH anticholinergic effects (dry mouth, blurred vision, urinary retention, constipation)
GI gastrointestinal
WTG weight gain more than six kg.
0 - 4+ absent or rare - relatively common
T, L, I, C Tablet, Liquid, Injectable, Capsule

Table 2a
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	-	-	None
Amoxapine	Asendin®	1980	-	-	1989
Maprotiline	Ludiomil®	1981	-	-	1988
Clomipramine	Anafranil®	1990	-	-	1996
NEW GENERATION ANTIDEPRESSANTS					
Trazodone	Desyrel®	1982	-	-	1986
Bupropion	Wellbutrin®	1989	Wellbutrin SR®	1996	1999
Venlafaxine	Effexor®	1994	Effexor-XR®	1997	None
Nefazodone	Serzone®	1995	-	-	None
Mirtazapine	Remeron®	1996	Remeron Soltab®	2001	None
SELECTIVE SEROTONIN REUPTAKE INHIBITORS (SSRI)					
Fluoxetine	Prozac®	1988	Sarafem™, Prozac Weekly®	2000 2001	2002
Sertraline	Zoloft™	1992	-	-	None
Paroxetine	Paxil™	1993	-	-	None
Fluvoxamine	Luvox®	1994	-	-	2000
Citalopram	Celexa™	1998	-	-	None

Sources: IMS America, Inc. and Food & Drug Administration (Annual)
** Entry is taken to mean the year when drug sales first occurred and not the FDA approval date.
** Data upto September, 2002

Table 2b
ENTRY IN THE ANTIDEPRESSANT MARKET

	Type of Entry			
Type	Molecule	Branded	Generic	All
MAOI	0	0	0	0
TCA	2	2	8	10
New Gen	5	8	2	10
SSRI	5	7	1	8
All Antidepressants	12	17	11	28

Table 3
WITHDRAWAL* FROM THE ANTIDEPRESSANT MARKET**

Generic Name	Name	Entry	Withdrawal Years
MONOAMINE OXIDASE INHIBITORS (MAOI)			
Isocarboxazid	Marplan®	1959	1995-1998
TRICYCLIC ANTIDEPRESSANTS & RELATED COMPOUNDS (TCA)			
Desipramine	Pertofrane®	1971	1988, 1990, 1992 - 2001
Doxepin	Adapin®	1973	2000
Imipramine	Janimine®	1975	1986 -1990
Amitriptyline	Endep®	1975	1989 -1990, 1999 - 2001
Trimipramine	Generic	1988	1992 - 1998, 2000 - 2001

Sources: IMS America, Inc. and Food & Drug Administration (Annual)

* Withdrawal is taken to mean that there were no reported sales of the drug in that year. This implies that quantities/revenues less than the IMS-required threshold may also be captured.

** Data upto August, 2001

Table 4

CHOICE SET IN THE ANTIDEPRESSANT MARKET**

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

* Types 1-4 stand for TCA, NewGen, SSRI and MAOI respectively.

** Union of all choice sets 1980-2001

Table 5a

Characteristics for Selected Antidepressants, 1980-2001

Drug	Type	Molecule	Branded	Generic Exists in Molecule (1980)	Generic Exists in Molecule (1990)	Generic Exists in Molecule (2000)	Average Dosing Frequency	Half-Life	Side Effects							
									Fatality	Anti-Cholinergic	Drowsiness	Insomnia/ Agitation	Orthostatic Hypotension	Cardiac Arrhythmias	GI Distress	Weight Gain
							FREQ	HALF	FAT	AC	DR	IA	OH	CA	GID	WTG
Elavil	1	1	1	1	1	1	1	24	2	4	4	0	4	3	1	4
Amitriptyline	1	1	0	1	1	1	1	24	2	4	4	0	4	3	1	4
Bupropion	2	13	0	0	0	0	1	32	1	1	3	2	0	0	0	3
Effexor	2	16	1	0	0	0	3	8	1	1	1	2	0	1	3	0
Fluoxetine	3	18	0	0	0	0	1	168	1	0	0	2	0	0	3	1
Imipramine	1	7	0	1	1	1	2	22	2	3	3	1	4	3	1	4
Nardil	4	23	1	0	0	0	1	2	4	2	1	2	2	1	1	3
Paxil	3	20	1	0	0	0	1	24	1	1	1	2	0	0	3	1
Prozac	3	18	1	0	0	0	1	168	1	0	0	2	0	0	3	1
Tofranil	1	7	1	1	1	1	2	22	2	3	3	1	4	3	1	4
Wellbutrin	2	12	1	0	0	1	3	14	1	0	0	2	0	1	1	0

Sources: IMS Health, Food & Drug Administration, *Depression Guideline Panel* (1993), *Physician's Desk Reference Generics* (1996), *Drug Information Handbook* (2001).

Table 5b

Average Characteristics by Type of Antidepressants, 1980-2001

Drug	Type	Number of Molecules	Number of Drugs	Number of Generic Drugs	Year of First Entry in the Market		Average Dosing Frequency	Half-Life	Side Effects							
									Fatality	Anti-Cholinergic	Drowsiness	Insomnia/ Agitation	Orthostatic Hypotension	Cardiac Arrhythmias	GI Distress	Weight Gain
							FREQ	HALF	FAT	AC	DR	IA	OH	CA	GID	WTG
TCA	1	11	26	10	1959	Mean	1.27	25.9	1.85	2.69	2.92	0.58	2.50	2.50	0.35	2.69
						StdDev	*0.81*	*269*	*0.13*	*0.98*	*0.99*	*0.40*	*1.02*	*0.25*	*0.23*	*2.21*
NewGen	2	5	10	2	1982	Mean	2.50	14.1	1.00	0.50	1.70	1.50	0.60	0.70	1.20	1.00
						StdDev	*0.65*	*91.3*	*0.00*	*0.25*	*2.41*	*0.65*	*1.44*	*0.21*	*0.96*	*1.60*
SSRI	3	5	8	1	1988	Mean	1.00	77.1	1.13	0.13	0.13	2.25	0.00	0.00	3.00	1.00
						StdDev	*0.00*	*4996*	*0.11*	*0.11*	*0.11*	*0.19*	*0.00*	*0.00*	*0.00*	*0.00*
MAOI	4	3	3	0	1959	Mean	1.67	2.00	4.00	1.67	1.00	2.00	2.00	0.67	1.00	2.33
						StdDev	*0.89*	*0.00*	*0.00*	*0.22*	*0.00*	*0.00*	*0.00*	*0.22*	*0.00*	*0.22*
Antidepressants (All Drugs)	4	22	47	13	1959	Mean	1.51	30.6	1.68	1.72	2.06	1.15	1.64	1.57	1.02	2.02
						StdDev	*0.93*	*34.2*	*0.98*	*1.35*	*1.48*	*0.91*	*1.41*	*1.21*	*1.17*	*1.46*

Sources: IMS Health, Food & Drug Administration, *Depression Guideline Panel* (1993), *Physician's Desk Reference Generics* (1996), *Drug Information Handbook* (2001).

Table 6a
Annual Revenue Shares of Types (%)

	TCA	MAOI	SSRI	New Gen
1980	98.46	1.51		
1981	98.52	1.48		
1982	92.67	1.50		5.83
1983	87.42	1.66		10.92
1984	84.02	1.90		14.08
1985	80.05	2.12		17.83
1986	77.15	2.26		20.59
1987	76.53	2.41		21.07
1988	61.56	1.97	20.83	15.65
1989	46.06	1.38	41.24	11.33
1990	36.02	0.95	53.30	9.74
1991	37.70	1.01	50.57	10.72
1992	31.87	0.90	58.10	9.13
1993	21.83	0.72	70.11	7.34
1994	12.35	0.45	78.63	8.56
1995	8.07	0.34	81.09	10.50
1996	5.60	0.27	82.54	11.58
1997	3.82	0.22	81.20	14.76
1998	2.62	0.18	77.56	19.64
1999	2.01	0.14	75.30	22.54
2000	1.44	0.12	73.82	24.62
2001	1.19	0.10	72.04	26.67

Table 6b
Annual Unit Shares of Types (%)

	TCA	MAOI	SSRI	New Gen
1980	98.50	1.50		
1981	98.56	1.44		
1982	96.39	1.48		2.14
1983	94.41	1.63		3.96
1984	93.59	1.83		4.58
1985	92.39	2.03		5.59
1986	90.78	2.35		6.87
1987	90.13	2.35		7.51
1988	79.70	2.01	10.87	7.43
1989	66.16	1.58	24.21	8.05
1990	57.10	1.14	34.76	7.00
1991	63.12	0.97	28.61	7.30
1992	54.67	0.88	37.30	7.15
1993	45.99	0.69	46.83	6.50
1994	35.74	0.45	55.87	7.94
1995	29.49	0.34	60.88	9.29
1996	28.75	0.34	59.07	11.84
1997	25.22	0.27	60.95	13.56
1998	22.07	0.24	62.25	15.44
1999	18.48	0.20	64.66	16.66
2000	15.66	0.18	66.31	17.85
2001	14.51	0.15	66.19	19.15

Table 7
Parameters from Demand Estimation Across Logit and Nested Logit Models

	OLS Logit	IV Logit	OLS NML1	IV NML1	OLS NML2	IV NML2	OLS NML3	IV NML3
Variable				Price Coefficients (α's)				
Price	-0.591**	-1.563**	-0.667**	-1.930**	-0.370**	-2.643**	-0.247**	-2.142**
	(0.113)	(0.286)	(0.112)	(0.307)	(0.109)	(0.352)	(0.102)	(0.475)
				Group Correlation Coefficients (ρ's)				
Type (ρ_t)			0.088	0.076	0.183*	0.229*	0.085**	0.132**
			(0.184)	(0.138)	(0.094)	(0.150)	(0.015)	(0.047)
Molecule (ρ_m)					0.259**	0.128**	0.586**	0.478**
					(0.093)	(0.061)	(0.065)	(0.095)
Brand (ρ_b)							0.858**	0.779**
							(0.150)	(0.126)
				Coefficients of Characteristics (β's)				
Half-Life	0.002	0.038*	-0.021	0.016	0.043*	0.078*	0.054	0.044
	(0.030)	(0.034)	(0.030)	(0.034)	(0.029)	(0.041)	(0.026)	(0.398)
Fatality	-1.672**	-2.244**	-1.538**	-2.221**	-1.032**	-2.614**	-0.953**	-2.196*
	(0.189)	(0.252)	(0.188)	(0.257)	(0.185)	(0.329)	(0.166)	(0.364)
Anti-Cholinergic	0.578*	0.449**	0.310*	0.055	0.650**	0.103	0.789**	0.260*
	(0.462)	(0.175)	(0.169)	(0.194)	(0.155)	(0.216)	(0.140)	(0.216)
Drowsiness	-1.017**	-0.899**	-0.835**	-0.623**	-0.933**	-0.610**	-1.038**	-0.744**
	(0.159)	(0.171)	(0.161)	(0.183)	(0.147)	(0.196)	(0.132)	(0.179)
Insomnia/ Agitation	-0.819**	-0.647**	-0.788**	-0.559**	-0.864**	-0.528**	-0.928**	-0.698**
	(0.143)	(0.158)	(0.141)	(0.163)	(0.131)	(0.177)	(0.118)	(0.157)
Orthostatic Hypotension	0.490	0.590**	0.483**	0.608**	0.158*	0.563**	0.163*	0.454**
	(1.119)	(0.129)	(0.117)	(0.132)	(0.114)	(0.159)	(0.103)	(0.146)
Cardiac Arrhythmias	-1.385**	-1.726**	-1.216**	-1.595**	-0.561**	-1.576**	-0.645**	-1.320**
	(0.162)	(0.194)	(0.163)	(0.197)	(0.165)	(0.258)	(0.146)	(0.244)
GI Distress	0.181*	-0.077	0.147*	-0.192*	0.152*	-0.393**	0.232**	-0.223*
	(0.098)	(0.125)	(0.097)	(0.130)	(0.090)	(0.140)	(0.081)	(0.149)
Weight Gain	-0.427**	-0.312**	-0.522**	-0.408**	-0.119	-0.110	-0.152*	-0.168*
	(0.126)	(0.137)	(0.126)	(0.140)	(0.121)	(0.158)	(0.109)	(0.136)
Average Dosing Frequency	-0.439**	-0.557**	-0.138	-0.187*	-0.506**	-0.637**	-0.566**	-0.650**
	(0.129)	(0.140)	(0.141)	(0.155)	(0.119)	(0.156)	(0.107)	(0.135)
Brand Dummy	1.360**	0.614**	1.320**	0.352*	0.931**	0.940**	0.146	1.174**
	(0.203)	(0.294)	(0.200)	(0.307)	(0.205)	(0.377)	(0.196)	(0.401)
Constant	1.589**	3.382**	-0.022	1.723*	-1.588	2.781**	-5.678**	-1.515*
	(0.684)	(0.868)	(0.752)	(0.911)	(0.691)	(1.093)	(0.710)	(1.338)
R^2 =	0.560	0.316	0.506	0.264	0.670	0.452	0.664	0.408

Notes:
Regressions included type, molecule and time dummies
No. of Observations = 644
Standard errors are in parentheses
* indicates t-statistic > 1
** indicates t-statistic > 2

Table 8
Parameters from Demand IV Estimation Across Models

	IV Logit	IV NML2	Random Coefficients (no demographics)	Random Coefficients (demographics)
Variable	Price Coefficients (α's)			
Price	-1.563**	-2.643**	-1.095*	-1.108**
	(0.286)	(0.352)	(1.045)	(0.253)
	Coefficients of Characteristics (β's)			
Half-Life	0.038*	0.078*	-0.673*	-0.343*
	(0.034)	(0.041)	(0.505)	(0.289)
Fatality	-2.244**	-2.614**	-5.078**	-5.159**
	(0.252)	(0.329)	(1.719)	(1.693)
Anti-Cholinergic	0.449**	0.103	0.021*	0.144**
	(0.175)	(0.216)	(0.019)	(0.013)
Drowsiness	-0.899**	-0.610**	-3.402**	-2.769**
	(0.171)	(0.196)	(0.360)	(1.021)
Insomnia/ Agitation	-0.647**	-0.528**	-7.701*	-7.768**
	(0.158)	(0.177)	(4.118)	(1.374)
Orthostatic Hypotension	0.590**	0.563**	-2.421*	-1.266**
	(0.129)	(0.159)	(2.322)	(0.502)
Cardiac Arrhythmias	-1.726**	-1.576**	-3.607**	-1.942**
	(0.194)	(0.258)	(1.094)	(0.184)
GI Distress	-0.077	-0.393**	1.796	0.131
	(0.125)	(0.140)	(2.785)	(1.004)
Weight Gain	-0.312**	-0.110	-2.925	-4.528**
	(0.137)	(0.158)	(3.844)	(1.241)
Average Dosing Frequency	-0.557**	-0.637**	-5.436*	-3.183**
	(0.140)	(0.156)	(3.221)	(1.287)
Brand Dummy	0.614**	0.940**	1.529	3.038**
	(0.294)	(0.377)	(2.747)	(1.540)
Constant	3.382**	2.781**	10.332**	60.790**
	(0.868)	(1.093)	(1.505)	(3.893)

Notes:
Regressions included type, molecule and time dummies
No. of Observations = 644
Standard errors are in parentheses
* indicates t-statistic > 1
** indicates t-statistic > 2

Table 9
Parameters from Random Coefficients 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:
Regressions included type, molecule and time dummies
No. of Observations = 644
Standard errors are in parentheses
* indicates t-statistic > 1
** indicates t-statistic > 2

Table 10
Random Coefficient Logit Elasticities

Table 10a

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

Table 10b

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

Table 11

Own- and Cross-Price Elasticities of Demand for Selected Antidepressants (2000-2001)*

Type***	Molecule	Drug	Drug Name	Cross-Price Elasticities					
				OPE	TCA	NewGen	SSRI	MAOI	ALL
					26**	10	8	3	47
1	1	1	Elavil	-0.209	0.355	0.437	0.302	0.402	0.366
1	1	3	Amitriptyline [G]	-0.456	0.381	0.179	0.116	0.102	0.275
1	2	4	Asendin	-0.089	0.707	0.543	0.393	0.496	0.605
1	2	5	Amoxapine [G]	-0.220	0.275	0.222	0.166	0.137	0.236
2	12	27	Bupropion [G]	-0.192	0.049	0.087	0.131	0.039	0.070
2	12	28	Wellbutrin SR	-0.157	0.254	0.405	0.308	0.181	0.291
2	12	29	Wellbutrin	-0.352	0.223	0.405	0.329	0.165	0.276
3	17	37	Celexa	-0.116	0.133	0.329	0.515	0.137	0.240
3	18	38	Prozac	-0.188	0.151	0.301	0.550	0.375	0.265
3	18	39	Prozac Weekly	-0.024	0.149	0.308	0.514	0.310	0.255
3	18	40	Sarafem	-0.158	0.156	0.298	0.573	0.384	0.272
3	19	41	Fluvoxamine [G]	-0.536	0.028	0.113	0.200	0.025	0.075
3	19	42	Luvox	-0.094	0.276	0.450	0.496	0.188	0.345
3	20	43	Paxil	-0.045	0.309	0.456	0.330	0.200	0.337
3	21	44	Zoloft	-0.065	0.407	0.415	0.312	0.218	0.380
4	23	46	Nardil	-0.313	0.552	0.447	0.308	0.752	0.501
4	24	47	Parnate	-0.436	0.616	0.494	0.343	0.780	0.554

* Average of elasticities for the last two years of the dataset: 2000 and 2001

** Number of drugs in each heading

*** Types 1-4 stand for TCA, NewGen, SSRI and MAOI respectively.

Table 12a

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 ($)	Per Unit ($)	From Annual Prescription ($)
1	*8*	*19*	*Ludiomil*	*Maprotiline*	*1981*	*943*	*49*	*77,200*	*0.00*	*1.46*
2	*15*	*33*	*Desyrel*	*Trazodone*	*1982*	7,391	661	352,453	0.03	11.51
1	5	12	Generic	Doxepin	1986	902	72	172,191	0.01	5.01
2	15	34	Generic	Trazodone	1986	23	30	5,433	0.01	2.54
1	4	8	Generic	Desipramine	1987	42	22	29,673	0.02	5.81
1	8	18	Generic	Maprotiline	1988	1,360	1,569	348,305	0.40	146.68
1	11	25	Generic	Trimipramine	1988	820	2,284	166,294	0.46	169.07
3	*18*	*38*	*Prozac*	*Fluoxetine*	*1988*	4,478 *	54,894	1,995,377 **	24.46	8928.90
1	2	5	Generic	Amoxapine	1989	6,396	27,930	1,353,365	5.91	2157.11
2	*12*	*29*	*Wellbutrin*	*Bupropion*	*1989*	*19,073*	*4,288*	*1,430,076*	*0.32*	*117.35*
1	*3*	*6*	*Anafranil*	*Clomipramine*	*1990*	*5,127*	*394*	*702,446*	*0.05*	*19.71*
1	9	21	Generic	Nortriptyline	1992	8,271	424	4,768,498	0.24	89.13
3	*21*	*44*	*Zoloft*	*Sertraline*	*1992*	*773*	*7*	*36,405*	*0.00*	*0.12*
3	*20*	*43*	*Paxil*	*Paroxetine*	*1993*	*7,952*	*151*	*734,202*	*0.01*	*5.07*
2	*16*	*35*	*Effexor*	*Venlafaxine*	*1994*	*5,579*	*139*	*296,643*	*0.01*	*2.69*
3	*19*	*42*	*Luvox*	*Fluvoxamine*	*1994*	*4,100*	*16,120*	*460,738*	*1.81*	*661.13*
2	*14*	*32*	*Serzone*	*Nefazodone*	*1995*	*1,068*	*37*	*212,130*	*0.01*	*2.70*
1	3	7	Generic	Clomipramine	1996	7,480	16,082	2,085,727	4.48	1636.67
1	10	23	Generic	Protriptyline	1996	5,452	22,016	1,082,250	4.37	1595.23
2	12	28	Wellbutrin SR	Bupropion	1996	11,757	12,881	1,362,284	1.49	544.77
2	*13*	*30*	*Remeron*	*Mirtazapine*	*1996*	*58,491*	*47,182*	*7,264,283*	*5.86*	*2138.78*
2	16	36	Effexor-XR	Venlafaxine	1997	20,861	4,990	1,435,885	0.34	125.36
3	*17*	*37*	*Celexa*	*Citalopram*	*1998*	*9,982*	*774*	*1,356,149*	*0.11*	*38.36*
2	12	27	Generic	Bupropion	1999	2,123	2,282	459,116	0.49	180.06
3	18	40	Sarafem	Fluoxetine	2000	37,998	9,481	3,352,705	0.84	305.35
3	19	41	Generic	Fluvoxamine	2000	45,325	40,902	8,669,484	7.82	2855.56
2	13	31	Remeron Soltab	Mirtazapine	2001	26,072	11,115	3,002,421	1.28	467.19
3	18	39	Prozac Weekly	Fluoxetine	2001	52,657	32,936	6,297,320	3.94	1437.69

* Billion dollars
** Thousand dollars
*** Types 1-4 stand for TCA, NewGen, SSRI and MAOI respectively.
**** 'Type' innovation is shaded
***** 'Molecule' innovation is in italics

Table 12b

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	
3	*18*	*38*	*Prozac*	*Fluoxetine*	*1988*	*1*	*1*	*1*	*1*	*30.68*
2	*13*	*30*	*Remeron*	*Mirtazapine*	*1996*	*2*	*2*	*3*	*4*	*1.49*
3	18	39	Prozac Weekly	Fluoxetine	2001	3	4	4	7	0.40
3	19	41	Generic	Fluvoxamine	2000	4	3	2	2	3.89
3	18	40	Sarafem	Fluoxetine	2000	5	11	6	11	0.61
2	13	31	Remeron Soltab	Mirtazapine	2001	6	10	7	10	0.33
2	16	36	Effexor-XR	Venlafaxine	1997	7	12	9	15	*0.15*
2	*12*	*29*	*Wellbutrin*	*Bupropion*	*1989*	*8*	*13*	*10*	*16*	*0.52*
2	12	28	Wellbutrin SR	Bupropion	1996	9	9	11	9	*0.74*
3	*17*	*37*	*Celexa*	*Citalopram*	*1998*	*10*	*17*	*12*	*18*	*0.06*
1	9	21	Generic	Nortriptyline	1992	11	19	5	17	*0.22*
3	*20*	*43*	*Paxil*	*Paroxetine*	*1993*	*12*	*21*	*15*	*22*	*0.01*
1	3	7	Generic	Clomipramine	1996	13	8	8	5	4.02
2	*15*	*33*	*Desyrel*	*Trazodone*	*1982*	*14*	*18*	*19*	*20*	*0.04*
1	2	5	Generic	Amoxapine	1989	15	5	13	3	*4.37*
2	*16*	*35*	*Effexor*	*Venlafaxine*	*1994*	*16*	*22*	*21*	*25*	*0.01*
1	10	23	Generic	Protriptyline	1996	17	6	14	6	5.05
1	*3*	*6*	*Anafranil*	*Clomipramine*	*1990*	*18*	*20*	*16*	*19*	*0.06*
3	*19*	*42*	*Luvox*	*Fluvoxamine*	*1994*	*19*	*7*	*17*	*8*	*1.59*
2	12	27	Generic	Bupropion	1999	20	15	18	12	0.49
1	8	18	Generic	Maprotiline	1988	21	16	20	14	1.06
2	*14*	*32*	*Serzone*	*Nefazodone*	*1995*	*22*	*25*	*22*	*24*	*0.01*
1	*8*	*19*	*Ludiomil*	*Maprotiline*	*1981*	*23*	*24*	*25*	*27*	*0.01*
1	5	12	Generic	Doxepin	1986	24	23	23	23	0.11
1	11	25	Generic	Trimipramine	1988	25	14	24	13	*1.20*
3	*21*	*44*	*Zoloft*	*Sertraline*	*1992*	*26*	*28*	*26*	*28*	*0.00*
1	4	8	Generic	Desipramine	1987	27	27	27	21	0.02
2	15	34	Generic	Trazodone	1986	28	26	28	26	0.01

* This is the ratio of the private marginal willingness-to-pay over the price of the drug (Table 12a)
** 'Type' innovation is shaded
*** 'Molecule' innovation is in italics
**** Types 1-4 stand for TCA, NewGen, SSRI and MAOI respectively.