Efficient Provision of Experience Goods: Evidence from Antidepressant Choice *

Michael J. Dickstein

September 30, 2021

Abstract

In the market for medical care, physicians often face uncertainty about how a newly diagnosed patient will respond to available treatments. I design a framework to analyze how price and promotion influence the learning process as the patient and physician jointly search for the most effective treatment. The dynamic model I employ accommodates large choice sets and permits learning to be correlated within clusters of choices. Applying this model to depression care, I ask how the design of a health insurance plan, including the required patient out-of-pocket costs by drug, might interact with the physician's learning process. In the data, patient costs largely correspond to the drug's wholesale cost. In contrast, I design a new drug pricing schedule that lowers the patient cost for those drugs that the model suggests are best to sample early in the search process. By using these price incentives to redirect the search process, I find physicians identify the optimal treatment faster, leading to lower overall costs, improved adherence, and ultimately better patient health.

Keywords: Learning, dynamic discrete choice, pharmaceutical demand, health insurance design, multi-armed bandits

*I owe special thanks to David Cutler, Guido Imbens, Ariel Pakes, and Dennis Yao for their guidance and suggestions. I thank Gary Chamberlain, Andrew Ching, Han Hong, Greg Lewis, Eduardo Morales, Julie Mortimer, Amanda Starc, and many seminar participants for helpful comments. I gratefully acknowledge support from the Harvard Business School and the Cowles Foundation at Yale University, where I conducted part of this research.
Dickstein: New York University and NBER, michael.dickstein@nyu.edu.

1 Introduction

In markets for experience goods, buyers learn about the quality of available products through experimentation. By testing out an option, they can reevaluate the alternatives to identify those that better match their preferences. In their next decision, buyers may choose one of the preferred products or may select a less well understood option to sharpen their opinions further–a trade-off between "exploitation" and exploration". Many markets feature this search behavior, including consumers' purchases of non-durable goods; manufacturers' choices of suppliers; and physicians' recommendations of prescription drugs for their patients.[1] The market for prescription drugs is a particularly important example because the pace and precision of the patient and physician's learning process has consequences for health. The health benefits come in two forms: patients may realize better outcomes while on more tailored treatment and may adhere to the proper drug regimen at higher rates.

I develop a dynamic model of prescription drug demand to capture the key features of the physician and patient's joint learning process. I estimate an empirical version of this learning model that allows both drug prices and promotion to influence the search process. Using the estimates, I consider alternative insurance plan designs that alter the relative cost-sharing levels between available drugs. I show that if the incentive design lowers out-of-pocket costs for treatments with better expected outcomes and tolerability, physicians and patients will search among those drugs most likely to relieve symptoms and least likely to cause side effects that discourage patients from continuing treatment. The identity of the drugs sampled changes and the time needed for exploration falls.

I focus specifically on the antidepressant market. The potential improvement in depression care is substantial, as few patients continue treatment to the six month threshold recommended by the American Psychiatric Association. In my empirical setting, half of the patients discontinue treatment by the first month and over 90% exit care before the recommended six months.[2] In

[1] For empirical models of such markets, see Erdem and Keane (1996) and Ackerberg (2003) on consumer products and Crawford and Shum (2005), Narayanan et al. (2005), and Chintagunta et al. (2009) on demand for prescription drugs.

[2] Berndt et al. (2002) find that such short treatment spells limit the probability of short-run recovery. Inadequate treatment duration also nearly doubles the rate of illness relapse (Melfi et al. (1998)). The resulting increase in

addition, physicians face uncertainty in this market about how a new patient will respond to the treatments available. Six classes of drugs compose the choice set, each affecting a distinct set of chemicals in the brain; physicians cannot predict ex ante which mechanism will work best for a particular patient.

To analyze this dynamic choice problem, I begin with a duration analysis. Using panel data on the treatment choices within a patient's illness spell, I identify the characteristics of drugs that more often cause patients to switch treatments. The results from this flexible hazard model suggest that a drug with greater side effects, more frequent doses per day, and higher out-of-pocket costs prompt patients to quit the treatment more rapidly. Patients and physicians prefer branded treatments, and, at the population level, have no clear favorite: the predicted probability that a drug will be effective for a new patient range from 40% to 70%, with seven treatments in the range of 65% to 70% predicted effectiveness. The piece-wise hazard estimates also suggest the likelihood of finding a drug ineffective varies by drug and changes in a nonlinear fashion over the course of an illness spell.

The uncertainty at the initial prescribing stage and the rich switching patterns call for a more detailed model of the patient and physician's dynamic choice problem. While this modeling requires assumptions on the agents' learning process, it will allow me to measure the effect of counterfactual pricing and promotion policies designed to improve adherence and patient outcomes.

The model involves three components. First, I specify the physician and patient's prior beliefs on the effectiveness of each option. In the model, sampling a drug treatment in a month is a Bernoulli trial, where a successful outcome means the drug provided "effective" relief. As typically defined in the medical literature, effectiveness encompasses more than simply whether a treatment improves the patient's symptoms. An effective treatment must also be convenient to administer, easy to acquire, and cause relatively tolerable side effects. I assume the effectiveness of a drug is a patient-specific endowment that may vary across patients depending on the severity of their mental illness but is fixed over time. The goal of estimation is to recover the probability that a drug is effective— that is, to identify the parameter of the Bernoulli distribution.

depressive symptoms inflicts serious pain: in surveys, patients equate 10 years living with depression to only 6 years living without it (Fryback et al. (1993)).

To characterize the joint distribution of these probabilities, I specify a prior distribution on each drug's probability of being effective that depends on a drug's price, branded status, side effect profile, and dosing requirements. I allow the influence of these drug characteristics to differ for patients with more severe illnesses and for patients that visit psychiatric specialists rather than general practitioners. Furthermore, because most of the antidepressants I study entered the US pharmaceutical market several years before my sample period, I assume the patient and physician have rational expectations. At the time they make a drug selection, their priors on the effectiveness of a treatment equal the parameters of the distribution that generates outcomes in the population. However, there is still substantial heterogeneity across patients in a drug's effectiveness; physicians and patients must search to identify the best match.

To complete the model, I must specify a learning process for the patient and physician as well as a decision rule for selecting a drug in each period. I choose a simple learning process: after observing a Bernoulli draw on a drug's effectiveness, the patient and physician update their priors in a Bayesian fashion. For the decision rule, two features of the market complicate the application of existing tools and call for methodological innovation. First, physicians select from among nineteen unique products in the antidepressant class. The agent—and the econometrician—must hold in memory the expected outcomes and the covariance in outcomes across the nineteen drugs when evaluating possible treatment regimens. This large number of required state variables reflects the "curse of dimensionality"; to handle a problem of this size, traditional methods require strong assumptions on the state variables to simplify computation.[3] Second, when searching for the drug best suited to a particular patient's illness, physicians may learn about the quality of each drug in a correlated fashion. After a poor outcome on drug A, for example, the physician and patient may avoid drugs B and C if they share drug A's mechanism of action in the brain. The optimal sequence of products to sample thus depends directly on the similarity among subsets of treatments within the choice set. If several products share a characteristic, physicians and patients might start with a drug in this set to learn rapidly about the patient's match to all of the drug's close substitutes.

I accommodate these features with a new estimation approach. To deal with computational

[3]See Aguirregabiria and Mira (2010) for a survey of dynamic discrete choice models and a discussion of the computational requirements.

roadblocks, I avoid direct dynamic programming solutions to the agent's sequential choice problem. These approaches find solutions via backward induction, which in my choice setting requires computing high-dimensional integrals for each patient-physician pair. I avoid computing these multidimensional integrals by appealing to an alternative solution to the physician's sequence problem developed in the statistics literature by Gittins and Jones (1979) and applied previously in economics contexts by Miller (1984) and Bergemann and Valimaki (1996), among others. In the drug choice setting, the actual effectiveness of a treatment is constant over time; the physician and patient can always return to an option they overlooked initially and it will yield the same outcome regardless of its place in the choice sequence.[4] The physician and patient can therefore apply forward induction to solve the sequence problem.

The classical forward induction solution, whose properties were first proven by Gittins and Jones (1979), provides a decision rule appropriate for a setting in which the choices are independent. As I show later in the data, treatments in my setting appear to be correlated in an observable pattern. When physicians and patients switch treatments, they do not shift according to a drug's share in the market. Rather, they more commonly switch in relation to the initial drug's class designation. Thus, in the antidepressant setting, the independence assumption likely fits the data structure poorly.

Rather than employ Gittins' index rule directly, I follow the suggestion of Pandey et al. (2007), who provide a two-level index rule that explicitly accounts for the dependence across choices. The approach loses the directly connection between Gittins' index rule and dynamic programming solutions, but nonetheless introduces a structure that better fits the depression context. Under the two-level rule, the decision maker first selects a drug class and then selects a drug within the class. After observing an outcome on a drug within the class, the patient and physician update their priors on the effectiveness of the drug sampled and on the class as a whole. In this way, a poor outcome on a drug within a class may diminish the perceived effectiveness of all drugs that are members of the class.

[4]This assumption would be incorrect in settings in which a treatment changes the physiology of the body such that it forecloses use of other treatments in the future. This type of sequential foreclosure is not a usual feature in the antidepressant choice setting.

The two-level index rule approach has three attractive features. First, the strategy vastly reduces computation for both the agent and the econometrician relative to a full dynamic programming solution that applies backward induction. Second, it captures optimizing behavior using a policy that appears more like a "rule of thumb" that physicians may follow in practice. Finally, the decision rule allows forward-looking behavior, as in Crawford and Shum (2005) and Ching (2010), but also permits spillovers in the learning process. A contribution of this paper is allow both correlated learning and experimentation in the discrete choice of prescription drug.

With the dynamic model, I can evaluate the likely effects of "value-based" care for depression. Specifically, I test the effect of insurance design and informational campaigns that aim to lower the insurer's long-run costs while maintaining patient health, translated to dollar terms. Here, the target of the counterfactual analysis includes patient prices and promotion to physicians, which enter the model through my specification of the patient and physician's prior probability that a product will be effective.

I find "value-based" insurance designs that set lower copayments for drugs with higher effectiveness can steer physicians and patients toward treatments with a higher probability of providing effective relief. With better matching, patients remain in treatment longer. I convert this increased rate of adherence to improved health using previous estimates from the medical literature. "Value-based" designs outperform commonly used tiered copayment policies both in terms of improving health and minimizing costs. For promotion, I test the effect of two informational campaigns. In the first, I allow policymakers to restrict manufacturers from labeling reformulations of branded drugs as distinct products, apart from differences in dosing. In this world, both adherence and costs fall slightly. Second, I endow general practitioners with the same preferences as psychiatric specialists. The result is a small improvement in adherence and greater use of generic treatments.

Finally, I use the predictions of the structural model on the effectiveness of each drug to form a list of recommended treatments. Observed switching patterns in the data identify seven treatments that have higher probability of a successful outcome. Existing protocols built from clinical trial outcomes recommend a larger list of products, some of which far underperform these seven treatments in terms of adherence. Thus, in the case of judging drug efficacy—and possibly for

policy evaluation more generally—my results suggest observed adherence rates provide valuable information to combine with the treatment effects found in randomized trials.

I proceed in the paper by first describing the market for depression care and transition patterns observed in the data in Section 2. In Section 3, I describe the components of the learning model built to explain the observed transitions. In Section 4, I describe the data in detail and specify the econometric model. I describe the results and fit of this model in Section 5. Finally, I conduct counterfactual simulations to test the effect of insurer policies on costs and patient health in Section 6. Section 7 concludes.

2 Background and Preliminary Analyses

2.1 Depression Care

Major depression affects 6.5% of adults in the United States each year. The illness causes patients to suffer significant impairment in their productivity, with surveys finding 60% have symptoms severe enough to keep them from performing daily tasks (Kessler et al. (2003)). The volume of diagnoses in the United States leads to a large market for drug treatment. In 2008, patients filled 164 million monthly prescriptions for antidepressants. As a class, only cholesterol regulators and pain medicines exceeded this volume of sales. In dollar terms, sales of antidepressants reached $9.6 billion in 2008, roughly 3.3% of the U.S. prescription drug market.[5]

The antidepressants I study fall into six distinct classes according to their effect on the concentration of the chemicals serotonin, norepinephrine, and dopamine in the brain. Patients react idiosyncratically to changes in the concentrations of these chemicals, creating uncertainty about the efficacy of any one biological mechanism for a patient (Murphy et al. (2009)). The current set of treatments entered the market in two waves. The first generation entered production in the late 1950s and 1960s, and includes a class known as tricyclic antidepressants (TCAs). They provide symptomatic relief, but at the cost of poor tolerability. In the late 1980s and 1990s, researchers developed treatments with "selective" effects on the brain chemistry, providing similar efficacy to the

[5] "Top Therapeutic Classes by US Sales", *IMS National Sales Perspectives*. IMS Health, 2008.

first generation of treatments but with lower risk of harm from overdose and fewer interactions with drugs for other conditions. These "second generation" treatments include the most popular drug classes today: selective serotonin reuptake inhibitors (SSRIs); serotonin-norepinephrine reuptake inhibitors (SNRIs); norepinephrine and dopamine reuptake inhibitors (NDRIs); noradrenergic and specific serotonergic antidepressants (NASSAs); and serotonin antagonist and reuptake inhibitors (SARIs) (Gartlehner et al. (2007)). In my sample period, the patient and physician can choose among two TCAs, eight SSRIs, three SNRIs, two NDRIs, one NaSSA, and two SARIs.

2.2 Preliminary analysis of switching behavior

Identification of the learning model parameters requires a sufficiently rich set of observed switches across products and classes with distinct characteristics. Although I defer a detailed description of the data until Section 4.1, here I summarize the switching patterns to motivate the structure imposed in the learning model. Importantly, because I focus on patients with more severe forms of depression, I interpret a switch as evidence that the drugs failed to provide effective relief at a reasonable cost. For the conditions I study, the American Psychiatric Association recommends at least six months of treatment; 90% of the observed switches or exits occur within this timeframe and so are unlikely to reflect successful treatment.[6]

2.2.1 Observed transitions

In Table 1, I illustrate the observed transitions in treatment from the initial drug choice to the choice picked in the second period of treatment. I report these switches by drug class. Decision points, t, roughly correspond to one month of treatment. In cases in which the physician writes a 90 day prescription, roughly 10% of all observed prescriptions, the decision point corresponds to three months. When a patient returns for follow-up before completing the initial prescription, the decision point corresponds to the observed number of days between office visits.

Table 1 contains three matrices: (1) all transitions, (2) only transitions to a drug treatment, excluding the outside good of no drug treatment, and (3) only transitions to a drug treatment that

[6]The American Psychiatric Association recommends treatment for a 6-8 week acute phase followed by a 16-20 week continuation phase to prevent relapse (Karasu et al. (2000)).

Table 1: Observed transitions between treatments after one period of care

Class in period 1	Class in period 2						
	Exit	TCA	NDRI	SSRI	SNRI	NaSSA	SARI
Panel A: All Transitions							
TCA	30.43	42.39	2.17	21.74	3.26	0.00	0.00
NDRI	29.27	0.39	62.90	5.78	0.77	0.13	0.77
SSRI	24.92	0.62	1.49	71.33	0.80	0.18	0.67
SNRI	21.05	0.25	1.37	2.99	73.35	0.12	0.87
NaSSA	28.30	0.00	1.89	11.32	1.89	54.72	1.89
SARI	36.17	4.26	5.67	15.60	2.84	0.00	35.46
Panel B: Only transitions to drug treatments (excludes quitting)							
TCA		60.94	3.13	31.25	4.69	0.00	0.00
NDRI		0.54	88.93	8.17	1.09	0.18	1.09
SSRI		0.82	1.99	95.00	1.06	0.24	0.89
SNRI		0.32	1.74	3.79	92.90	0.16	1.10
NaSSA		0.00	2.63	15.79	2.63	76.32	2.63
SARI		6.67	8.89	24.44	4.44	0.00	55.56
Panel C: Only transitions to a different product (excludes quitting)							
TCA		0.00	9.38	78.13	9.38	0.00	3.13
NDRI		3.29	33.55	49.34	7.89	0.66	5.26
SSRI		7.04	19.68	51.62	12.64	1.26	7.76
SNRI		1.82	20.91	40.91	25.45	0.91	10.00
NaSSA		0.00	10.00	70.00	10.00	0.00	10.00
SARI		13.04	19.57	56.52	10.87	0.00	0.00

Notes: This table illustrates observed switches recorded for the 10,000 patient sample drawn from the MarketScan database for use in estimation. Roughly 35% of patients receive no drug care at their initial visit and so do not appear in the above switch statistics. Selected antidepressant classes shown. The number of individual drugs modeled in each subclass are: SSRI, 8; NDRI, 2; SNRI, 3; TCA, 2; NaSSA, 1; SARI, 2. Thus, no patients switch 'within class' when starting on a NaSSA.

is distinct from the treatment chosen in the initial period. Panel A illustrates that roughly 20-35% of patients quit treatment after the first month, though the rate depends importantly on the class of the drug chosen. For example, patients beginning treatment on a drug in the SNRI class have the lowest rate of exit, at 21%, while patients on drugs in the TCA or SARI class quit at rates over 30%. Panel B shows that, for those patients remaining in treatment, the majority remain on the same drug or the same drug class chosen initially, particularly if the initial treatment was a second generation SSRI or SNRI product. Panel C focuses only on those patients who switch to a new drug treatment.

The patterns of switching in Panel C motivate my decision to specify the expected probability of success under a drug as a function of its class and to permit correlation across products in the learning model. First, it is clear in the switching statistics that patients do not switch to new drugs based on the average market shares in the choice set. For example, of those patients who begin treatment on a drug in the SSRI class and then switch to another drug, 54% choose another SSRI; if the drugs were uncorrelated and patients chose the next treatment in proportion to market shares, we would expect around 44% of patients to choose an SSRI. A similar pattern exists for drugs in the SNRI and NDRI classes. These frequencies suggest a possible clustering of choices by class. It also suggests a hazard model may be insufficient to explain observed behavior. In particular, the pattern of switching likely depends on both the degree of correlation of product quality within a class and the differentiation in quality across classes for a particular patient. Even if a patient has a poor outcome on a drug in one class, he may switch to a second drug within the class if the best option in any alternative class still has worse expected quality. I develop a learning model later that features this correlated learning structure.

To illustrate the timing of switches, I include in matrix form in Table 2 the share of patients who remain on the initial choice at different decision points within an episode. Conditional on the level of adherence, about 10-15% of patients switch at the first decision point. Physicians and patients appear to learn quickly about a patient's match to a treatment. This variation will help identify the precision in the physician and patient's priors.

2.2.2 Hazard of switching

Finally, before specifying the learning model, I conduct a duration analysis to highlight the key individual and drug characteristics that drive the decision to quit or switch treatments. I use a flexible piecewise proportional hazard model to capture changes in the duration on treatment at different points in the patient's treatment spell. In this setting, I interpret the estimated probability of switching as the probability that the patient and physician together find the focal drug ineffective.[7]

[7]In Section A.2 of the technical appendix, I use the hazard estimates in a two-stage model of drug selection. I combine the hazard estimates on switch timing with a simple logit model of drug choice for the periods in which

Table 2: Share of patients who remain on the treatment chosen at their initial diagnosis (in %), subset by observed treatment length

Count in episode	Length of episode (# monthly prescriptions dispensed)							
	1	2	3	4	5	6	7	8
1	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2		85.1	86.8	88.9	89.4	89.6	90.4	91.4
3			84.4	84.2	84.7	85.8	86.4	88.2
4				82.3	82.0	83.0	83.9	86.4
5					80.2	81.0	82.2	84.4
6						79.6	80.1	82.6
7							79.3	80.7
8								79.9
Share of dataset	56.7%	12.9%	9.3%	6.4%	4.4%	3.1%	2.1%	1.7%

For each patient, I group the observed durations into month-long intervals. I use a piecewise constant proportional hazard, allowing for censoring at the end of treatment.[8] When patients leave treatment, I assume they do not return and thus I end the treatment spell in that month.

The estimates suggest that patient and physicians believe most treatments will be effective with probability ranging from 40-80%, with drugs in the SNRI and SSRI classes most likely to be effective. Increases in out-of-pocket costs, in the number of doses needed per day, and in the rate of side effects decrease the perceived probability of successful treatment on a drug. The hazard of finding a drug ineffective increases with the severity of the patient's diagnosis. These characteristics will motivate the specification of the priors in the later econometric model.

3 Learning Model

The preliminary analysis of the transition data in Section 2.2 suggests that patients quit treatment sooner when prices are high, when side effects are more severe, and when a drug simply provides little symptomatic relief for the patient. I now model this learning process explicitly. The cost is im-

the patient and physician switch treatment. I compare the predictions of this simple model to the predictions of the main learning model to highlight the gains from introducing Bayesian learning.

[8] I provide more detail on the specification of the hazard in Section A.2 of the technical appendix. I report the estimates from the model in Table A1. In Table A2, I report the predicted probability that a drug is effective over three periods of treatment, as implied by the hazard estimates.

posing several assumptions on the agent's decision rule. The benefit is that we can predict choices— and therefore patient health outcomes and medical costs— in counterfactual policy environments.

In specifying a learning model for patients and physicians, the goal is to capture the richness of the switching patterns observed in each patient's treatment episode while keeping the decision rule sufficiently parsimonious that it is computable both for decision-makers and for the econometrician. I describe the model in the following two sections. I start by describing the latent outcome variable and the patient and physician's updating process. I then describe the decision rule the patient and physician use to select a treatment.

3.1 Latent Outcomes and Updating

I describe three components to the updating process in the model: specification of the patient and physician's priors, the distribution of outcomes, and the updating process.

Before the patient begins treatment, the physician and patient together form priors on how effective each drug will be for the patient. These priors vary only according to the observable characteristics of the patient, and so patients with the same characteristics begin with the same priors. Setting the prior probabilities for the patient and physician in this setting is akin to the initial conditions problem discussed in detail in Heckman (1981) and summarized in Aguirregabiria and Mira (2010). In my baseline model, I simply impose that in the first period of the structural model, all individuals begin in the same unobserved state. Intuitively, I consider each patient at the time of his diagnosis as having no history: the physician and patient start the search for an effective treatment together in the initial period with no experience.[9]

To link these prior probabilities to observed outcomes, I assume rational expectations. The patient and physician's priors on the parameters of the distribution for the probability that a drug is effective, p_{ij}, equals the parameters in the population of patients with the same observable characteristics as patient i. What the patient and physician must learn about in this setting is precisely how the focal patient will respond to a treatment.

[9]Keane and Wolpin (1997) use a similar assumption in the labor economics setting, presuming that all individuals enter their dynamic model at age 16 with no work experience in previous years.

Formally, I assume there exists a discrete latent outcome, Y_{ijt}, for individual i, drug j, and time period t that is drawn from a Bernoulli distribution with the probability of a successful outcome equal to p_{ij}:

$$Y_{ijt} \backsim p_{ij}^{k}(1-p_{ij})^{1-k}, k \in \{0,1\} \tag{3.1}$$

where $k = 1$ if drug j proves effective in period t.[10] For the prior distribution of p_{ij}, I choose a beta distribution with parameters $(a_{ij,0}, b_{ij,0})$, and mean and variance equal to:

$$\mu_{ij,0} = \frac{a_{ij,0}}{a_{ji,0} + b_{ij,0}} \tag{3.2}$$

$$v_{ij,0} = \frac{a_{ij,0}b_{ij,0}}{(a_{ij,0} + b_{ij,0})^2(a_{ij,0} + b_{ij,0} + 1)} \tag{3.3}$$

where $a_{ij,0} > 0$ and $b_{ij,0} > 0$.

In modeling outcomes as discrete events, I depart from previous empirical learning models, including Erdem and Keane (1996), Crawford and Shum (2005), and Ching (2010). I do so primarily to simplify the estimation procedure. In particular, with discrete outcomes I can set up an exact likelihood without simulation, which is helpful in settings in which choice sets are large such that most choices have small market shares. In addition, updating in this setting with discrete outcomes is very simple. Given that the beta distribution is the conjugate prior for the Bernoulli likelihood, the posterior distribution of p_{ij} also follows a beta distribution. After t trials of treatment j, simply replace $(a_{ij,0}, b_{ij,0})$ in the above formulas with $(a_{ij,t}, b_{ij,t}) = (a_{ij,0} + s, b_{ij,0} + (t - s))$, where s is the number of successes in t trials on treatment j. That is, add to $a_{ij,0}$ the number of successes observed and add to $b_{ij,0}$ the number of failures observed in t trials on drug j. In the empirical application, the number of success and failures are not observed; I integrate over the discrete number of possible realizations when computing the likelihood.

[10] Time period t refers to the interval between observed prescriptions, which generally equals 30 days. In the empirical exercise, I adjust the definition of t depending upon whether I observe prescriptions of 90 day duration and whether patients switch treatment in less than 30 days.

3.2 Decision Rule

The physician and patient can follow one of two approaches in their treatment selection. First, they may learn from experience but fail to consider the future implications of their decision. I label this behavior Bayesian-myopic, following the nomenclature of Brezzi and Lai (2002).[11] Second, the physician and patient may be forward-looking, considering the future implications of today's choice in their decision rule. Under this assumption, the physician and patient may experiment with lesser-known options to identify those that produce a superior outcome. This dynamic experimentation process takes the form of the classical multi-armed bandit problem. Gittins and Jones (1979) and Gittins (1979) showed that the optimal solution to a sequential search problem of this form is an index rule.

In economics, early applications of multi-armed bandit rules include Rothschild (1974) on pricing under demand uncertainty, and Jovanovic (1979), Miller (1984), and Mortensen (1986) on decision-making in labor markets. I follow this stream of research and show how an index rule solves the physician and patient's treatment decision problem.

I describe first the Bayesian-myopic rule followed by two versions of Gittins' index rule. The first of the latter two dynamic rules follows the classical multi-armed bandit assumptions that require the choices to be independent. In the second version, I drop this assumption and instead group the choices into related clusters by an observed characteristic. This version permits spillovers in learning across related treatments. In my later empirical implementation, I apply the model testing approach of Vuong (1989) to look for support in the data for allowing correlation in the learning process in the dynamic model.

3.2.1 Bayesian-myopic rule

The Bayesian-myopic physician and patient maximize the expected probability of successful treatment under choice j, given past experience and idiosyncratic shocks to the value of treatment:

$$\max_{j\in 1,\ldots,J} E(p_{i,j,T+1}|a_0, b_0, \widehat{Y}_{ij}) + \varepsilon_{ijt} = \max_{j\in 1,\ldots,J} \mu_{j,T+1} + \varepsilon_{ijt} \tag{3.4}$$

[11] Several papers in the marketing and economics literatures use Bayesian-myopic models of learning in the context of pharmaceuticals, including Chintagunta et al. (2009), Narayanan et al. (2005), and Coscelli and Shum (2004).

Here, the decision-maker has a history of experience over T periods and must select a treatment to sample in the next period, labeled $T+1$. The patient and physician update their priors at $T+1$ using the vector of realizations, $\widehat{Y}_{ij}$.[12] As written, the expected probability of success under j depends only on realized outcomes on j and the agent's priors; there is no change in the expected utility from treatment j following experience on options $l \neq j$. However, it is possible to relax this assumption and allow the experience from other products to affect the agent's posterior probability on drug j. I describe how to allow correlation in learning in Section 3.2.3 in the context of the fully dynamic model.

The idiosyncratic shock, ε_{ijt}, captures features like advertising for j at t that change over time but affect the choice at t. I assume it follows a Type I extreme value distribution.[13] The probability of a choice will therefore take a logit form.

3.2.2 Dynamic rule with independent options

In the data, physicians and patients appear to experiment with treatments, switching to new medications after a trial period. To rationalize this observed behavior, I move beyond the fairly straightforward Bayesian-myopic learning model and solve the agent's fully dynamic problem. In this model, I allow physicians and patients to consider the future implications of their current decisions.

Given the set of choices, $j \in \{1, ..., J\}$, the physician and patient choose a sequence of drugs to maximize the expected discounted sum of outcomes, Y_t:[14]

$$\int ... \int E_{p_1,...,p_J} \left(\sum_{t=1}^{\infty} \delta^{t-1} Y_t \right) d\Pi^{(1)}(p_1) \cdots d\Pi^{(J)}(p_J) \tag{3.5}$$

[12] If the patient and physician have not yet tried drug j, the vector $\widehat{Y}_{ij}$ is empty and the agents' perceived probability that drug j will be effective equals their prior beliefs.

[13] Including an idiosyncratic shock with a Type I extreme value distribution also serves as a computational simplification, as it generates smooth probabilities of each choice in the likelihood function. As noted by Crawford and Shum (2005), previous authors, including Rust (1987) and Hotz and Miller (1993) have introduced extreme value errors in part for computational convenience.

[14] I drop i subscripts here to simplify notation.

The discount rate, δ, is given and $p = (p_1, ..., p_J)$ is the vector of probabilities that a drug $j \in 1, ..., J$ is effective. Here, p is unknown; the agent forms independent priors, Π, on the elements of p. The state variables include the number of successes and failures under each choice.

There is a long literature in econometrics, industrial organization, and labor economics that proposes solutions to single-agent dynamic discrete choice problems of this form. Early empirical work directly applied dynamic programming to solve the sequence problem in (3.5), using the methodology Rust (1987), Hotz and Miller (1993) or Keane and Wolpin (1994) propose. Subsequent authors, including Arcidiacono and Miller (2011), extended this framework to allow persistent unobserved heterogeneity.

I exploit features of the drug choice setting that allow me to avoid the computational burden of backward induction. Instead, I apply a solution that uses forward induction. Under forward induction, I can transform the multidimensional integrals in the dynamic problem into a series of simpler one-dimensional integrals.

Specifically, a forward induction rule requires two maximizations. In an inner maximization, the agent considers each choice as if it were optimal and determines a stopping time at which she would prefer to retire rather than continue using the option. The agent solves a one-dimensional optimal stopping problem of this type for each choice. The sum of returns under each choice, were it taken for the optimal length of time, is the "index" of the choice. In an outer maximization step, the agent selects the option with the largest index. This procedure repeats in each period after the agent updates her priors using past outcomes. I provide a formal description of this procedure in Section A.3 of the technical appendix.

Gittins and Jones (1979) prove that the forward induction rule solves the sequence problem in (3.5) if the following conditions hold: (1) the decision-maker selects one option at t; (2) the options not selected do not contribute to the individual's outcome; (3) the options not selected will produce the same average outcome in later periods as they would if chosen in the initial period; and, (4) the options are independent.[15] Independence in this context implies that the probability of a successful outcome from treatment j is independent of the probability of success on treatment k.

[15]See Mahajan and Teneketzis (2007). Whittle (1981) proves that Gittins' index rule is also optimal in settings in which new products enter the choice set, provided the entry is independent of the decision maker's past choices.

The setting of antidepressant choice satisfies conditions (1) and (2): physicians recommend only one drug per period and patients do not suffer side effects or health benefits from drugs they do not take. Condition (3) also holds in this setting because the effectiveness of a drug is individual-specific, not time-specific. The true effectiveness of drug A for patient i remains stable over time. What changes over time is simply what the patient believes about the drug's quality based on past experience.[16] Condition (4) holds by assumption. I drop this assumption in the following subsection and propose an alternative decision rule when choices are correlated.

To apply a forward induction rule, I must compute the one-dimensional optimal stopping problem for each unique treatment choice and at each period in the patient's episode. Here, it is helpful to use an approximation. I employ a closed form approximation to the optimal stopping problem in the inner maximization step, as described in Brezzi and Lai (2002).[17] This numerical approximation is asymptotically optimal as the discount rate, δ, approaches 1. I provide more detail on this final computational step in Section A.4.1 of the Technical Appendix.

The index rule takes the following form, where $\Pi_t^{(j)}$ denotes the posterior distribution of the probability of success for individual i under choice j after t periods. Suppressing i subscripts:

$$G(\Pi_t^{(j)}) = \mu_{j,t} + \sqrt{v_{j,t}} * \left[\psi \left(\frac{v_{j,t}}{h(\delta) * \sigma^2(\mu_{j,t})} \right) \right] \tag{3.6}$$

As described in Section 3.1, $(\mu_{j,t}, v_{j,t})$, are the mean and variance of the posterior beta distribution for p_j, the probability that drug j is effective. Equations (3.2) and (3.3) contain the mean and variance for the prior distribution of p_j; the mean and variance of the posterior distribution

[16] In the antidepressant example, ignoring minor discontinuation periods, physicians can prescribe drugs in any order without changing their treatment effects.

[17] This approximation allows the researcher to quickly compute the index in the discounted infinite horizon setting. Previous work, including Gittins (1989) approximates the index rule in the infinite horizon setting by choosing a finite horizon, N, and then computing the finite horizon analog using backward induction. When the discount rate is close to 1, the researcher would need to choose a large N to find a good approximation for the index rule, which is computationally costly. Brezzi and Lai (2002) simplify this step in the computation by providing a closed form approximation to the optimal stopping problem. They do so by using a diffusion approximation, which involves computing the Gittins' index for a Wiener process (See Section 2.1 of Brezzi and Lai (2002)).

take a similar form:

$$\begin{aligned} \mu_{j,t} &= \frac{a_{j,t}}{a_{j,t}+b_{j,t}} \\ v_{j,t} &= \frac{a_{j,t}b_{j,t}}{(a_{j,t}+b_{j,t})^2(a_{j,t}+b_{j,t}+1)} \\ \sigma^2(\mu_{j,t}) &= \mu_{j,t} * (1-\mu_{j,t}) \end{aligned}$$

Here, $(a_{j,t}, b_{j,t}) = (a_{j,0} + s, b_{j,0} + (t-s))$, where s is the number of successes in t trials under treatment j and $(a_{j,0}, b_{j,0})$ are the parameters of the prior beta distribution for p_j. The function $\psi(.)$ represents the closed-form numerical approximation to the boundary of the one-dimensional optimal stopping problem for each drug; I describe its form in Section A.4.1 of the Technical Appendix. Again, δ is the discount rate, and $h(\delta) = -\ln(\delta)$. Finally, $\sigma^2(\mu_{j,t})$ is the variance of the distribution of Y_{ijt}, which equals $E[p_j(1-p_j)] = \mu_{jt} * (1-\mu_{jt})$, given that Y_{ijt} is a Bernoulli draw with parameter p_j.

The distinction between the Bayesian-myopic rule and this index rule comes solely from the second term in expression (3.6) above, which represents the incentive to experiment.[18] Intuitively, this incentive will be small in three cases. First, when physicians and patients discount the future more heavily, the experimentation term decreases in size. That is, there is less incentive to experiment if future outcomes matter less to the decision-maker. Second, when past experience diminishes $v_{j,t}$, the incentive shrinks. Finally, when $\sigma^2(\mu_{j,t})$ is large, the experimentation incentive will be small. $\sigma^2(\mu_{j,t})$ is maximal when $\mu_{j,t} = 0.5$. That is, when success or failure is is equally likely, there is little information to gain from experimentation.

As in the Bayesian-myopic decision rule, I allow an additive idiosyncratic shock, ε_{ijt}, to influence value of treatment under drug j. The complete decision rule is therefore:

$$\max_{j \in 1,\ldots,J} G(\Pi_t^{(j)}) + \varepsilon_{ijt} = \max_{j \in 1,\ldots,J} \mu_{j,t} + \sqrt{v_{j,t}} * [\psi(.)] + \varepsilon_{ijt} \tag{3.7}$$

In contrast to the Bayesian-myopic rule in (3.4), the index rule in (3.7) accounts for the variance

[18] The idea that agents face a trade-off between "exploration" and "exploitation" in their decision rule has a long history in the economics and statistics literature. Lai and Robbins (1985) and Lai (1987) provide a early formulation of the multi-armed bandit problem and the key properties of the optimal solution to such problems.

in outcomes under each choice. In short, the physician and patient balance the impulse to choose the drug with the highest expected outcome against the incentive to experiment with lesser-known but potentially superior treatments.

3.2.3 Dynamic rule with correlated options

The index results due to Gittins (1979) hold under the assumption that the treatment options are independent. In general, allowing correlation across choices is a more difficult problem, though work by Pandey et al. (2007) and Rusmevichientong and Tsitsiklis (2010) provide guidance in cases in which the set of choices can be grouped into clusters or recast as a linear parameterization of characteristics.

In the antidepressant setting, there is a natural clustering of the available treatments according to their mechanism of action in the brain. As detailed earlier, the probability of success on a particular treatment relates strongly to the drug class used, as each class affects the brain's chemistry differently. I therefore develop an estimation approach that employs the index rule that Pandey et al. (2007) define for the case in which dependent choices can be grouped into clusters.

I run a two-level algorithm. The patient and physician first select a drug class and then select one of the drug options within the class. Formally, the patient and physician first calculate an index by cluster, pooling together past successes and failures on all drugs within the class. Once they select the class with the largest class-level index, they must choose a drug from the class selected. To do so, the patient and physician again calculate indices, but now at the drug level; the expected success of a drug in the class depends only on the past successes or failures realized on that particular drug. The patient and physician choose the drug at this step with the largest index value.

Updating in this model allows for spillovers or correlation across products. If the patient experiences a negative outcome on a drug in class c, she will lower her expectation of successful treatment on all drugs within the class. In the likelihood, the probability of choosing a drug within that class will fall, and so the next choice is more likely to come from a different class. The degree to which the share declines depends on the similarity between clusters. Decreasing the index on

one drug within a cluster may not cause an immediate switch across clusters if the next best option still lies within the cluster. Pandey et al. (2007) provide a formal analysis of the optimality of index rules of this form in the discounted infinite horizon setting.[19] Similar to the nested logit model in the static discrete choice context, this two-level approach applies most readily to settings with a natural taxonomy to the choice set.

As in the independent Gittins' index case, however, the optimal index rule proves difficult to compute exactly in most settings; the size of the state space depends on the number of choices and the set of possible outcomes that can result. Pandey et al. (2007) instead provide an approximation to the optimal policy with an error that is bounded. The approximation, which I adapt below to fit the market for depression treatment, parallels the algorithm used in Section 3.2.2 for the independent choice case. In both cases, I choose the policy developed for settings in which there are only a limited number of observations in the time series dimension of the data and the discount rate lies above .95. For the case in which I allow correlation, Pandey et al. (2007) provide some analysis of how well the approximation performs as a function of the cluster definition. The authors show that the two-level policy performs best when the clusters have greater separation and are more cohesive. That is, if agents follow an optimal rule, my approximation that pre-specifies clusters will work best when the choices within a cluster are similar in quality and the clusters themselves are differentiated in quality for each individual.

The two-level model involves the following steps. First, collect the prior experiences of success or failure for all drugs within a class into $(a_{c,t}, b_{c,t})$, where $c = 1, ..., C$ represents the drug's class designation defined in the medical literature:

$$\begin{aligned} a_{c,t} &= \sum_j 1\{j \in c\} * a_{j,t} \\ b_{c,t} &= \sum_j 1\{j \in c\} * b_{j,t} \end{aligned}$$

Next, compute the dynamic allocation index by class. The index rule takes the same form as

[19] Pandey et al. (2007) show in Theorem 1 of their paper that one can solve the full dynamic programming problem set out in Equation (3.5) by examining each cluster in isolation, in a manner parallel to Gittins' approach in the independent bandit problem.

in the independent case in Equation (3.6), but $\Pi_t^{(c)}$ now denotes the posterior distribution of the probability of success under all drugs in class c after t periods:

$$G(\Pi_t^{(c)}) = \mu_{c,t} + \sqrt{v_{c,t}} * \left[\psi\left(\frac{v_{c,t}}{h(\delta) * \sigma^2(\mu_{c,t})}\right)\right]$$

Here, $(\mu_{c,t}, v_{c,t})$, are the mean and variance of the posterior beta distribution for p_c, the probability that class c is effective.

The probability of choosing class c in period t, allowing again for Type 1 extreme value shocks to the index under each cluster, takes the form:[20]

$$\text{Prob}_{c,t} = \frac{\exp(G(\Pi_t^{(c)}))}{1 + \sum_{s=1}^{C-1} \exp(G(\Pi_t^{(s)}))}$$

In the second level, we use the same index policy rule, but only for those treatments in the class chosen at the first level. Here, the index is at the drug level, as in (3.6). The probability of choosing $j \in c$ now equals the conditional probability of choosing j, conditional on first choosing class c, multiplied by the probability of choosing c:

$$\begin{aligned} \text{Prob}_{j \in c,t} &= \text{Prob}_{c,t}(\text{Prob}_{j,t}|1\{\text{c chosen}\}) \\ &= \text{Prob}_{c,t} * \frac{\exp(G(\Pi_t^{(j)}))}{\sum_{k \in c} \exp(G(\Pi_t^{(k)}))} \end{aligned}$$

4 Econometric Model

4.1 Data

I draw a sample from Thomson Reuter's *Commercial Claims and Encounters* MarketScan database, collecting its patient-level insurance claims data for outpatient visits and prescription drug services in the years 2003 through 2005. I link this data to patient demographics and plan design fea-

[20] In the choice of class, I set the index for the 'outside good' class equal to 0 as a level normalization.

tures from the *Benefit Plan Design* MarketScan database. The individuals recorded in the data include active employees working for a group of large, self-insured firms in the United States. The employees' dependents and some classes of retirees enter the database as well.

To form the sample, I identify patients in the outpatient data who receive a new diagnosis in one of five categories of depression.[21] A diagnosis is "new" if the patient has not received any treatment for depression in the six months before the beginning of the new illness episode.[22] Conditional on observing a patient with a depression diagnosis in an office visit, I collect panel data on the patients' prescription drug use over the remaining months of the sample. There are two benefits to conditioning on a diagnosis in creating the sample. First, I can restrict attention to more severe diagnoses, avoiding patients prescribed antidepressants for off-label uses or prescribed for conditions that may not strictly require drug treatment. Second, I can study the extensive margin, as my data will contain individuals diagnosed with major depression but who do not fill a prescription.

The initial filters lead to a dataset of 102,780 unique patient episodes of depression care. The individuals are members of 307 insurance plans, each with a distinct set of required drug copayments. I report summary statistics on the variables in the dataset in Table 3. In the sample collected from the MarketScan database, 27% of patients suffer from major depression, the most severe diagnosis examined. Of those patients diagnosed with depression, 25% visit a psychiatrist, 47% visit general practitioners and the remaining share visit other specialists, such as obstetricians.

[21] The depression diagnoses include major depression; dysthymia and depression with anxiety; prolonged depressive reaction; adjustment disorder with depressed mood; and, depression not otherwise specified. These match the International Classification of Diseases (ICD-9-CM) codes of: 296.2, 296.3, 300.4, 309.0, 309.1, and 311. Melfi et al. (1998), Pomerantz et al. (2004), and Akincigil et al. (2007) use similar diagnostic codes in selecting a sample of depression sufferers.

[22] I choose six months as my threshold for the pre-treatment period because the medical literature defines a "recurrence" as new symptoms that arise after a gap of at least six months (Melfi et al., 1998). I exclude left-censored observations for patients who recieve depression care within six months of entering the Marketscan dataset. Appendix A.1 details the steps used to form the analysis sample.

Table 3: Summary statistics on variables entering the empirical model

Panel 1: Drug Characteristics

Product Name	Class	Brand?	Reform?	Dosing	Insurer Cost ($/month)			Copay ($/month)			Mkt Share (%)		
					2003	2004	2005	2003	2004	2005	2003	2004	2005
None	None	-	-	-	0.00	0.00	0.00	0.00	0.00	0.00	35.19	39.27	34.21
Citalopram HBr	SSRI	No	No	1	0.00	36.02	19.88	0.00	7.16	8.68	0.00	0.21	4.24
Celexa	SSRI	Yes	No	1	72.95	73.97	75.32	22.56	31.66	35.36	4.14	2.66	0.12
Lexapro	SSRI	Yes	Yes	1	63.19	63.28	68.91	22.74	24.78	28.76	13.77	13.02	12.00
Fluoxetine HCL	SSRI	No	No	1-2	30.91	17.66	16.98	9.92	6.91	7.35	7.33	11.47	10.63
Prozac	SSRI	Yes	No	1-2	140.07	102.45	179.51	23.39	19.52	39.45	0.35	0.63	0.18
Paroxetine HCL	SSRI	No	No	1	68.68	49.30	36.99	11.04	8.34	8.83	1.68	4.27	4.44
Paxil CR	SSRI	Yes	Yes	1	81.45	83.90	87.50	21.82	20.97	25.05	6.40	3.65	1.60
Zoloft	SSRI	Yes	No	1	79.93	81.88	85.01	20.35	25.64	28.20	12.50	10.56	9.63
Cymbalta	SNRI	Yes	No	1-2	0.00	103.83	111.85	0.00	27.85	34.62	0.00	0.51	2.42
Effexor	SNRI	Yes	No	2-3	74.28	70.15	76.05	22.91	25.40	26.29	0.51	0.57	0.48
Effexor-XR	SNRI	Yes	Yes	1	109.53	116.02	118.71	21.56	22.45	26.09	8.56	7.52	6.99
Bupropion HCL	NDRI	No	No	3	38.58	49.85	50.67	8.45	9.02	9.78	0.26	3.42	4.45
Wellbutrin XL	NDRI	Yes	Yes	1	104.82	104.05	113.46	21.20	21.24	26.82	6.47	5.71	5.09
Amitriptyline HCL	TCA	No	No	1	5.87	4.24	4.02	6.93	5.30	5.26	0.72	0.91	0.72
Nortriptyline HCL	TCA	No	No	1	11.83	6.19	6.28	7.37	6.22	6.14	0.25	0.38	0.40
Mirtazapine	NaSSA	No	No	1	59.97	42.40	31.85	11.17	8.82	7.95	0.48	0.71	0.64
Nefazodone HCL	SM	No	No	2	46.65	38.74	32.91	7.87	8.41	8.20	0.10	0.28	0.15
Trazodone HCL	SARI	No	No	3	8.00	6.06	5.81	7.17	5.82	6.03	1.28	1.78	1.61
										Total patient episodes:	11,341	41,285	50,154

Panel 2: Individual covariates in product choice model

Variable	Share
Patient diagnosed with major depressive disorder	27.18
Patient visits a general practitioner (internal medicine/family medicine)	46.67
Patient visits non-psychiatric specialist (obstetrician, etc)	27.87
Patient visits psychiatrist	25.45
Total # unique patients:	102,780
Total # of observations:	267,390

Notes: Sample includes individuals newly diagnosed with depression in an outpatient office visit; the data include information on subsequent office visits and prescriptions dispensed over the patient's episode. I exclude pregnant women, individuals with psychiatric comorbidities, and patients whose episodes are right-censored in Marketscan data. Financial variables reflect the mean value across 307 unique plans in the data. Shares reflect only the initial prescription written for a patient. Dosing is reported in doses per day.

Table 4: Reported side effects by antidepressant class

	Drug class					
Side effect	TCA	NDRI	SSRI	SNRI	NSM	SM
All cardiac	X	X		X	X	X
Anticholinergic	X	X		X		
All neurological	X	X	X	X		
Sexual dysfunction	X		X	X		
Insomnia/Activation		X	X	X		
Nausea		X	X	X		
Sedation	X				X	X
Weight gain	X		X		X	

In addition to the MarketScan data, I collect aggregate information on a product's expected side effects and dosing requirements. Gartlehner et al. (2007) and Gelenberg et al. (2010) examine relevant citations from the medical literature, collecting information on treatment effects and side effects from each study examined. I report a summary of their findings in Table 4. I report the dosing requirements for each drug, as listed in Centers for Medicare and Medicaid Services, U.S. Department of Health and Human Services (October 2015), in Table 3.

In July 2003, 16 unique products competed in the market for antidepressants.[23] In November 2004, citalopram, the generic form of Celexa, entered the market. In August 2004, a new SNRI, Cymbalta (duloxetine) entered, providing 18 drug options and one non-drug outside option by the end of the sample period. I allow the patient to face the latest choice set available at the start of his treatment episode.

[23] In defining the choice set, I include only drugs maintaining at least a 0.3% market share. This cutoff rule excludes extremely rare, older generation treatments, such as MAOIs (monoamine oxidase inhibitors). Given their small shares, I cannot collect reliable price information across distinct plans.

4.2 Priors

The goal of estimation is to recover the patient and physician's prior beliefs on the probability that each drug in the choice set will be a successful match for a patient. Under rational expectations, these priors will also be the parameters of the distribution that generates the latent outcomes. Given the distributional assumptions chosen in the model, I design an estimation algorithm to recover the mean and variance of the beta prior distribution on the individual match probability, p_{ij}.

To better address the policy questions of interest, I parameterize the prior distribution of p_{ij} as a function of both choice characteristics and interactions of individual characteristics with these choice characteristics. The product characteristics include the patient's price, population-level reports of side effects on product j, an indicator for whether the drug requires multiple doses per day, class indicators, and indicators for whether a drug is branded or a branded reformulation.[24] The individual characteristics include an indicator for whether the patient has the most severe depression diagnosis and an indicator for whether the physician is a psychiatrist. With this parameterization, the probability p_{ij} reflects the quality of the patient's match in more dimensions than simply health. Key drivers of switching over time include the patient's recognition of the costs of the drug or its side effects.

Rather than directly parameterize $a_{i,0}$ and $b_{i,0}$, the two non-negative parameters of the beta distribution, I follow Ferrari and Cribari-Neto (2004) and choose a beta regression parameterization:[25]

$$p_{ij}|X_{ij} \backsim Beta(a_{i,0}, b_{i,0}) \tag{4.1}$$

$$\mu(X_{ij};\gamma_1) = \frac{a_{i,o}}{a_{i,0}+b_{i,0}} = \frac{\exp(X_{ij}\gamma_1)}{1+\exp(X_{ij}\gamma_1)} \tag{4.2}$$

$$\phi(\gamma_2) = a_{i,0}+b_{i,0} = \exp(\gamma_2) \tag{4.3}$$

[24] Reformulations typically come in the form of "controlled-release" or "extended release" versions of the original patented drug. If approved, the manufacturer of the reformulation receives three years of additional exclusivity to market the new product (Huskamp et al. (2008)).

[25] Under this parameterization, $a_{i,0} = \mu(X_{ij};\gamma_1)\phi(\gamma_2)$ and $b_{i,0} = (1-\mu(X_{ij};\gamma_1))\phi(\gamma_2)$.

The prior mean of p_{ij}, $\mu(X_{ij};\gamma_1)$, takes a logit form, as seen in (4.2). The prior variance of p_j takes a more complicated form under the beta regression parameterization:

$$V(p_{ji}|X_{ij}) = \frac{a_{i,0}b_{i,0}}{(a_{i,0}+b_{i,0})^2(a_{i,0}+b_{i,0}+1)} \tag{4.4}$$

$$= \frac{\mu(1-\mu)}{1+\phi} \tag{4.5}$$

$$= \frac{\exp(X_{ij}\gamma_1)}{(1+\exp(X_{ij}\gamma_1))^2(1+\exp(\gamma_2))} \tag{4.6}$$

This expression reveals the economic content of the second parameter, ϕ. Here, ϕ functions similarly to a precision parameter: for a fixed μ, the larger the value of ϕ, the smaller the variance in p_{ij}. The goal of estimation is to recover $\gamma = (\gamma_1, \gamma_2)$.

4.3 Likelihood

The formation of choice probabilities follows directly from the learning model described above. In the case in which we assume the products are independent, the likelihood for individual i in period t is:

$$\prod_{j=1}^{J} E_{\varepsilon_{i1t},...,\varepsilon_{iJt}}(1\{G_{ijt}(\Pi_t^{(j)}) + \varepsilon_{ijt} > G_{ikt}(\Pi_t^{(k)}) + \varepsilon_{ikt} \text{ for all } k \neq j\}^{d_{ijt}}) =$$

$$\prod_{j=1}^{J}\left(\frac{\exp(G_{ijt}(X_{ij},\widehat{Y}_{i,j,t-1};\gamma))}{1+\sum_k \exp(G_{ikt}(X_{ik},\widehat{Y}_{i,k,t-1};\gamma)}\right)^{d_{ijt}}$$

where G_{ijt} is the index rule, $d_{ijt} = 1$ if individual i chose drug j in period t, and $\widehat{Y}_{i,l,t-1}$ is a vector of realized outcomes under treatments $l = 1, ..., J$ during the previous $(t-1)$ periods of experience, if any.[26] The logit form follows from the assumption that the idiosyncratic errors for each (i, j, t), labeled ε_{ijt}, follow an extreme value distribution.

As the econometrician, I do not observe the vectors $(\widehat{Y}_{i,1,t-1}, ..., \widehat{Y}_{i,J,t-1})$. In the model, I assume each element of $\widehat{Y}_{i,j,t-1}$ is a Bernoulli trial, such that the distribution of the number of successes observed in $(t-1)$ trials is binomial. Therefore, to form the likelihood, I sum over the possible

[26] I normalize the index for the outside option to zero, such that $\exp(G_{i,Outside,t}(.)) = 1$. If the physician and patient have not yet sampled drug l, then $\widehat{Y}_{i,l,t-1}$ will be an empty vector.

sequences of observed outcomes, weighting by the probability of observing those sequences of choices up to period t, where $t \in \{1, ..., T_i\}$:

$$\sum_s \omega_{i,s} \prod_{j=1}^{J} \left(\frac{\exp(G_{ijt}(X_{ij}, \widehat{Y}^s_{i,j,t-1}; \gamma))}{1 + \sum_k \exp(G_{ikt}(X_{ik}, \widehat{Y}^s_{i,k,t-1}; \gamma)} \right)^{d_{ijt}}$$

Here, $\omega_{i,s}$ is the probability of observing one of $s \in S$ possible sequences of outcomes, $\widehat{Y}^s_{i,j,t-1}$, for $j = 1, ..., J$. $\widehat{Y}^s_{i,j,t-1}$ represents discrete counts of successes and failures realized over $(t-1)$ periods. The probability, $\omega_{i,s}$, follows a discrete binomial distribution. Under rational expectations, the parameters that underlie $\omega_{i,s}$ equal the parameters of the agents' priors. The number of points, S, represents all possible permutations of $\{0, 1\}$ for the number of periods up to t for which the patient tried a treatment.[27] Because I model the outcome variable as discrete, I set up an exact likelihood and avoid simulation error in my empirical implementation.

In the case with dependency across the drugs via clusters, I calculate the choice probabilities at two levels: the probability of a class being chosen and the conditional probability of a drug being chosen, conditional on the first-level class choice. The likelihood at period t in this case has the following form:

$$\sum_s \omega_{i,s} \prod_{c=1}^{C} \left[\left(\frac{\exp(G_{ict}(\Pi_t^{(c),s}))}{1 + \sum_m^{C-1} \exp(G_{imt}(\Pi_t^{(m),s})} \right)^{d_{ict}} \prod_{j \in c}^{J_c} \left(\frac{\exp(G_{ijt}(\Pi_t^{(j),s}))}{\sum_k^{J_c} \exp(G_{ikt}(\Pi_t^{(k),s})} \right)^{d_{ijt}} \right]$$

where drug j is a choice contained in class c.

4.4 Identification

Two sources of variation in the data separately identify the coefficients in the specification of the mean and precision of p_{ij}, the probability that drug j is effective for individual i. First, the data reveals the identity of the patient and physician's choices throughout the episode of treatment. This identifies the expected mean outcome under the available choices following standard arguments in

[27] In the implementation, I calculate the probability of all possible sequences of outcomes under each product for T_i periods, where T_i is the number of choices observed for individual i. I sum over the likelihood values for each of the possible sequences of latent outcomes, for the observed sequence of drugs individual i chooses.

the discrete choice literature. Second, I extract information from the timing of the observed switches to identify the precision of the agents' priors.

Specifically, I observe both the identity of the drug from the main dataset and I collect a vector of characteristics of each drug from the patient data and from external sources, such as the report from each drug's original clinical trial. Restricting to $T = 6$ periods in the panel data, there are 19^6 possible sequences of the 19 choices over six periods, sampled with replacement. With a large enough dataset, it is possible to estimate the probability that the decision vector equals each of these 19^6 vectors using their frequency in the data. In my empirical model, I estimate a much lower dimensional parameter space. Because the number of empirical probabilities is greater than the number of parameters in my specification, the approach meets a necessary condition for identification.

By choosing to parameterize the probability that a drug is effective as a function of X_{ij}, I gain identifying power when the covariates are sufficiently rich. Consider the case in which I observe only the choice in the patient's first period of treatment. A logit assumption on the unobservables will allow me to form choice probabilities and to identify J indices, one for each of the possible choices. However, conditioning on X_{ij} covariates, some of which are continuous, allows me to predict the choice probabilities over the J options at different values of X_{ij}. This requires variation in X_{ij} that in turn produces distinct drug shares. As I show in Table 3, observed copayments differ by drug, as do side effect profiles, class status, and other characteristics. The data show patients and physicians choose a distinct set of treatments depending on the patient's diagnosis, the physician's specialty, and on drug characteristics like price.

Finally, to identify the precision of the agents' priors, I exploit variation in the timing of observed switches, conditional on the expected outcome by treatment. Intuitively, the slower is the switching, conditional on the mean outcome, the higher is the uncertainty in the agent's prior probability. Slow switching thus implies the patient and physician have lower precision in their priors. I identify the common precision parameter under three assumptions from the model described above: there are no switching costs, any product characteristics not captured in the specification of the expected outcome are independently and identically distributed over time, and the unobservables have an

extreme value distribution, suitably normalized. The strength of these assumptions depends in part on the richness of the observable choice characteristics available to the researcher. With few observables, the unobservables in the model may not be idiosyncratic; one would then be unable to determine whether the lag before a switch were due to low precision or to a persistent unobservable.

Table 2 illustrates that the majority of the observed switches occur after the initial drug choice. Table 1 shows variation in the timing of the observed switches by drug class. This variation allows me to identify a common precision parameter.[28]

5 Results

5.1 Model estimates

In Table 5, I report the estimated parameters, $\gamma = (\gamma_1, \gamma_2)$ from three models: (1) the two-level dynamic model, with choices clustered by class; (2) the Bayesian-Myopic version of the two-level model, which lacks the experimentation incentive included in the fully dynamic index rule; and (3) the one-level dynamic model in which I assume the choices are independent. I fix the discount rate, δ, equal to .95 and calculate the standard errors of these maximum likelihood estimates using a bootstrap procedure.[29]

The sign, if not the level, of the parameters is informative. In model (1), the baseline model, patients and physicians begin with priors that suggest products with greater population rates of nausea side effects, greater dosing requirements, and higher patient costs are less likely to lead to successful outcomes. I use two indicator variables to capture demand for branded products: an indicator for 'branded' and an indicator for 'reformulation', which are new versions of branded products often requiring fewer doses per day. The coefficient on the reformulation indicator is large and positive, suggesting that within the set of branded products, patients and physicians believe reformulations are more likely to prompt a successful outcome.

[28] One could also specify a precision parameter that varies with patient or physician observables, provided there is variation in the timing of switches according to these characteristics.

[29] See Efron and Tibshirani (1993). I use 30 bootstrap samples of patient illness episodes selected with replacement in my bootstrap implementation.

Table 5: Model estimates

	(1)		(2)		(3)	
Covariates in prior mean	Est	S.E.	Est	S.E.	Est	S.E.
1{SSRI}	99.16	8.72	-0.84	6.75	181.64	35.66
1{SNRI}	-22.58	2.23	-69.23	10.74	-91.43	110.30
1{NDRI}	-43.09	5.77	-73.38	7.33	-171.71	157.10
1{NaSSA}	-35.36	8.47	-15.08	3.24	-57.68	15.09
1{SARI}	-40.03	2.04	-7.62	18.92	-56.81	14.79
1{TCA}	-72.44	4.36	-13.72	8.88	-115.53	29.32
1{SSRI}*1{Major Depression}	10.06	2.55	-0.48	1.62	9.11	52.90
1{SNRI}*1{Major Depression}	-0.67	1.89	40.79	6.08	-179.14	105.48
1{NDRI}*1{Major Depression}	-3.57	2.65	59.46	1.59	49.35	58.37
1{NaSSA}*1{Major Depression}	-4.01	2.11	-0.41	0.89	-5.69	1.65
1{SARI}*1{Major Depression}	-3.99	0.36	-0.12	4.40	-2.99	0.93
1{TCA}*1{Major Depression}	-9.44	0.69	-0.49	10.93	-13.49	3.69
1{SSRI}*1{psychiatrist}	16.85	4.98	-0.76	1.13	-66.48	28.81
1{SNRI}*1{psychiatrist}	-1.04	2.15	-20.76	13.81	59.57	57.31
1{NDRI}*1{psychiatrist}	-4.29	11.32	-57.70	1.74	-144.99	71.56
1{NaSSA}*1{psychiatrist}	-3.55	1.45	-0.42	0.45	-4.64	1.33
1{SARI}*1{psychiatrist}	-3.09	0.27	0.19	3.07	-2.18	0.63
1{TCA}*1{psychiatrist}	-8.59	0.28	-0.20	15.48	-11.72	3.25
1{>1 dose/day}	-76.77	3.66	-20.69	18.94	-74.60	43.14
% nausea in trial	-15.62	5.20	9.23	47.77	455.72	77.58
1{reformulation}	13.91	15.96	70.17	2.99	457.77	234.04
1{branded}	-9.80	4.93	3.23	1.54	-36.25	53.51
copayment, in $/day	-81.58	19.58	-3.42	2.26	-216.10	147.27
log(precision)	-3.28	2.61	-32.06	0.02	-24.39	30.81

Note: This table illustrates results from three model specifications: columns labeled (1) provide estimates of the two-level dynamic index rule, columns labeled (2) provide estimates from the two-level Bayesian myopic model, and columns labeled (3) provide estimates the one-level index rule or Gittins index model.

The parameter, $\phi = \exp(\gamma_2) = .037$, illustrates the prior uncertainty patients and physicians have at the time of the initial prescription. The variance of the priors equals:

$$v(p_j|X_{ij}) = \frac{\mu(1-\mu)}{1+\phi} = \frac{\exp(X_{ij}\gamma_1)}{(1+\exp(X_{ij}\gamma_1))^2(1+\exp(\gamma_2))}$$

where μ is the prior mean. The larger is ϕ, for a fixed μ, the smaller is the variance. Here, for the average patient pair in the data, the expected probability of successful treatment on the most common SSRI medications is 77%.

5.2 Model Fit

I assess model fit in two ways. First, I use the baseline model to predict the identity of the choice selected within each patient's illness spell and the timing of the patient's exit from care. I compare the observed choices against these predicted choice probabilities. In my setting, there are 19 treatment options. Therefore, rather than report whether the model's top prediction equals the observed choice, I allow the model some flexibility and instead report whether the model's top three or top five predictions match the observed choice in the data. Table 6 contains the predictions. I also include predictions from the two-level Bayesian-myopic model as a comparison.

The exit rates in Panel A for Table 6 show the relative quality of fit for the adherence rate. Relative to the raw data, the predicted initial quit rate of 32% under the dynamic model lies close to the level in the raw data of 36%. The Bayesian-myopic rule predicts a quit rate of 24%, far below the level in the raw data. The quality of this prediction is important for the counterfactual exercises, as greater adherence translates into better health outcomes. Comparing the fit of the model in predicting the choice of inside good, both the Bayesian-myopic and dynamic model perform similarly, matching the data at a rate of 45% to 75% when looking at the top three predictions and at a rate of 66% to 80% when looking at the top five predictions.

In a second measure of fit, I follow Vuong (1989) and carry out a classical approach to model selection to compare the one-level and two-level dynamic models. Using the Kullback-Leibler Information Criterion to measure the closeness of the model to the truth, I compute likelihood-ratio based statistics for testing the null hypothesis that the competing models are equally close to

Table 6: Model fit - Matching predicted product choices to observed selections by patient

Panel A: Share of patients who exit in the first 3 months			
Model	t=1	t=2	t=3
Raw data	36.2	50.8	59.6
Dynamic model, clustered by type	31.7	51.6	64.7
Bayesian myopic model, clustered by type	23.6	41.1	54.2
Panel B: Share of patients for whom the obs choice equals the model choice			
Examination	t=1	t=2	t=3
Compare to top 3 ranked by dynamic model	44.9	62.7	74.3
Compare to top 3 ranked by Bayesian-myopic model	48.5	65.7	76.7
Compare to top 5 ranked by dynamic model	70.8	72.8	78.9
Compare to top 5 ranked by Bayesian-myopic model	66.3	74.6	79.0

the true data generating process. I find a test statistic of 11.95; at the 95% critical value, the data favors using the two-level dynamic model over the one-level model.

6 Counterfactual Experiments

I return to the public policy question motivating the empirical work, employing the dynamic model to predict how copayment policies and informational campaigns affect the insurer's net drug costs and adherence.

I run two sets of counterfactuals. In the first, I consider two alternative copayment schemes: uniform pricing, under which all patient prices are set equal to $5, and "value-based" insurance design, as described by Chernew et al. (2007). I carry out these pricing changes under the assumption that the insurer acts as a central planner, changing the entire vector of prices for the choice set patients and physicians face. In the second set of counterfactuals, I simulate the effect of two informational campaigns. In the first, policymakers discourage use of reformulated drugs, which are variants of existed drugs for which the United States grants three additional years of patent protection. I simulate this policy by adjusting the reformulation indicator in the prior mean. In the second, I endow general practitioners with the same preferences as psychiatric specialists by adjusting the physician-specific elements in the prior mean.

To quantify changes in patient health under the alternative policies, I translate the rate of adherence to an expected number of weeks the patient suffers from depressive symptoms, assigning a dollar value to the relief from symptoms. I describe the calculation in Section A.5 of the Technical Appendix. In brief, I use results from a survey of psychiatric specialists that Berndt et al. (2002) conduct. The results include the panel's prediction of the median probability of longer run rates of full and partial recovery from depressive symptoms under broad treatment categories. They define categories by the length of time the patient adheres to the regimen and by whether the treatment is a first or second generation antidepressant. I use the dynamic model to predict the drug/adherence category for each patient in my sample within each counterfactual scenario; I assign patients the matching recovery probability from the survey. In this way, I use the observed heterogeneity in treatment choices and adherence for a patient of a particular diagnosis to predict longer run health outcomes. The number of patients in my sample that fall into each category will differ by counterfactual policy.[30]

In the baseline case, the distribution of treatments prescribed leads to an average reduction in depressive symptoms after three months worth $930, versus a drug cost of roughly $67. I compare the dollar value of symptomatic relief under the counterfactual policies against the cost of drug care in Table 7.

6.1 Insurance Design

I measure the effect of two alternative copayment policies on insurer costs and patient health. For the baseline case, I use the tiered pricing policies defined by the 307 benefit plans in the database. In the first counterfactual, I set a uniform copayment of $5 for all treatments. In the second, I set a "value-based" policy, as Chernew et al. (2007) suggest. I set the copayments for generic treatments to $5 for a 30 day supply. For the specific drug treatments that lead to the fewest switches and highest adherence in the raw data, I set the copayment of the generic versions to $0 and the branded versions of these high adherence drugs to $5. For all other branded treatments, including those

[30] In this calculation, I miss part of the heterogeneity in the heath effect that comes from being matched to the correct drug within a treatment generation, conditional on adherence. One could use the estimates from the structural model to determine the short run utility gain from more tailored treatment.

with both moderate and poor predicted adherence, I set the copayment equal to the insurer's cost for the drugs. In effect, the patient pays the entire cost for these non-preferred medications.

For both policies, I simulate rates of adherence over three decision points and calculate net costs to the insurer, which equal the cost charged by the manufacturer less the copayments paid by patients.[31] I translate adherence to a reduction in depressive symptoms, denominated in dollar terms, and judge the performance of each policy by its effect on the trade-off between insurer cost and patient health.

Looking first at the $5 uniform copayment policy, the main effect is to increase the use of branded drugs relative to their generic counterparts. Physicians in the data express strong preferences for particular brand names and so prescribe them readily when price effects disappear. The policy leads to a 1% increase in adherence by month three, stemming largely from the increased use of effective SSRIs and SNRIs. However, the improvement comes at an added cost equal to 45% of the baseline level. By lowering all copayments to $5, the insurer loses revenue from these patient contributions and fails to encourage the use of cost-effective treatments. The policy leads to a $10 increase in the value of health above the baseline average of $930 over three months of care, but at an added cost of $30.

The "value-based" design offers a three month payoff of $193. Patients shift to the most effective SSRIs and SNRIs and to generics more generally, incentivized by copayments of $0 or $5 per month. Consumers of branded drugs with moderate or poor expected adherence face the full insurer cost of the drug, typically $60-$80. As a result of this shift, adherence increases by nearly 20% into the third month of care. The change in the distribution of treatments prompts a $3 increase in costs, largely due to the increase in the initiation of care.

6.2 Informational campaigns

I use the model to test the effect of two informational campaigns. The goal of these campaigns, much like the marketing efforts of pharmaceutical firms, is to change the patient and physician's

[31] Copayments typically return to the insurer indirectly via a bargaining process with the drug manufacturer (See Levy (1999)).

Table 7: Patient health vs. drug costs, under counterfactual scenarios.

	Three month drug costs per patient (in $)	Value of utility gain from symptomatic recovery (in $)
Baseline	67.25	929.90
Effects of copayment policies		
	Change in three month drug costs per patient (in $)	Change in symptomatic recovery (in $)
All copayments set to $5	30.28	10.42
Value-based design	3.13	193.35
Effects of informational campaigns		
	Change in three month drug costs per patient (in $)	Change in symptomatic recovery (in $)
Discourage use of reformulation products	(0.04)	(1.85)
Endow all physicians with psychiatrists' priors	0.65	2.18

priors. If the campaigns lead to greater use of the most effective drugs, patients may adhere at higher rates and have better health outcomes.

In the first counterfactual campaign, policymakers discourage physicians and patients from using reformulations by prohibiting them from being marketed as distinct medications from their original branded formulations, apart from the change in dosing. To carry out this counterfactual scenario, I turn off the reformulation indicator for all drugs. I leave the branded indicator and the dosing indicator at their original values in the data. In reaction, the use of drug care falls about 1%, either because physicians write fewer prescriptions or patients decide not to fill a prescription in hand. The main effect, however, is to shift the distribution of prescriptions away from reformulations. There is a small cost savings of $.04, but health falls $1.85 due to the drop in adherence.

In the second counterfactual campaign, I endow general practitioners with the same priors as psychiatric specialists, say through new guidelines or through an electronic medical record system that sends alerts to the general practitioner about best practices. This campaign leads to a slight

improvement in adherence and health: costs rise $.65, but health rises $2.18 over 3 months.[32]

The simulations highlight a key feature of decisions in this market. After accounting for learning, the largest gains in patient outcomes occur when policies promote treatments with above-average efficacy and tolerability. Costs improve when the policies use either informational campaigns or lower prices to encourage the cheapest products within the set of effective options. This is precisely the approach of value-based insurance design, on the price dimension, and "academic detailing", as described by Soumerai and Avorn (1990), on the informational dimension. However, the policies differ in costs. Given information technology, the creation of multiple tiers for pricing involves little additional expense. Campaigns, however, are costly. A small-scale academic detailing program in Pennsylvania in 2006 required $1 million a year in funding.[33]

6.3 New Protocol

Finally, using the estimated probability of effectiveness by drug treatment from the baseline model, I build a new treatment protocol. Policymakers might employ a protocol of this type in informational campaigns directed to physicians to encourage them to use the most effective and least costly treatments available.

Past theoretical and empirical work supports the notion that adherence or attrition rates can provide valuable insight on effectiveness. Philipson and DeSimone (1997), for example, argue that randomization and blinding in clinical trials may not actually produce unbiased treatment effects when trial subjects are Bayesian and actively learn about the effectiveness of the experimental drug. Attrition rates, in contrast, summarize the effect of all unobserved aspects of a treatment that make it undesirable. Chan and Hamilton (2006) judge the value of attrition rates empirically using data from a trial of HIV treatments. I extend this notion to adherence rates in insurance claims data. Since physicians in practice expend considerably less effort to ensure patient adherence relative to

[32] A limitation of this analysis is that I do not observe drug promotion. It is possible that psychiatrists receive more detailing than general practitioners, changing their priors. Ching and Ishihara (2012) and Leffler (1981) find evidence that detailing can be persuasive.

[33] Scott Hensley, "Negative Advertising: As Drug Bill Soars, Some Doctors Get An 'Unsales' Pitch," *The Wall Street Journal*, March 13, 2006, sec. A, p.1.

the managers of clinical trials, the observed rates of quitting treatment provide an externally valid measure of a drug's effectiveness.

In Table 8, I report the recommendations for antidepressant treatment from three published protocols. In the final columns, I include protocols built from the hazard model and the dynamic model's predictions. I use an expected probability of effectiveness of 55% at the initial decision point as a lower threshold for recommending a product.

There are two main distinctions between the model-based protocol and existing recommendations. First, the new protocol de-emphasizes drugs in the NDRI and TCA classes. Although clinical trials generally show little difference in efficacy between classes, these two classes prompt high rates of switching in the panel data. Second, while no existing protocol differentiates among available SSRIs, the model suggests using off-patent compound fluoxetine (Prozac) and paroxetine (Paxil) over alternatives. This preference arises both from the lower cost of these drugs and from greater tolerability.

The results suggest that analyses that use attrition rates can enhance existing protocols. Crismon et al. (1999), in their algorithm design project, argued for using such analyses, but found few credible studies to incorporate. The methodology developed here can help fill this evidence gap cheaply.

7 Conclusion

I specify a model of learning that permits patients and physicians to search within a large set of potentially correlated experience goods. To estimate the parameters of this learning process, I depart from standard dynamic programming solutions due to computational limitations. Employing a forward-induction rule, I design an estimation framework that accommodates learning over potentially correlated options.

With this framework, I measure the importance of a key trade-off in incentive design: policies which promote cheaper but less effective options may lead to decreased adherence. When patients quit treatment early, they suffer from depressive symptoms for longer periods. Patient health declines and the insurer's long-run costs may increase. Managing this trade-off requires "value-

Table 8: Treatment recommendations from published protocols vs. the dynamic model's recommendations. X's within a row indicate the protocol recommends the drug compound.

		(1)	(2)	(3)	(4)	(5)
Product	Class	Texas Medication Algorithm Project, Report for Major Depression (1998)	APA Guidelines, Second Edition (2000,2005)	Comparative effectiveness review; AHRQ (2007)	Hazard Model	Dynamic Model
Amitriptyline	TCA		X			
Bupropion	NDRI	X	X	X	X	
Wellbutrin XL	NDRI	X	X	X	X	
Citalopram	SSRI	X	X	X	X	X
Celexa	SSRI	X	X	X	X	X
Cymbalta	SNRI		X	X	X	
Lexapro	SSRI	X	X	X	X	X
Fluoxetine	SSRI	X	X	X	X	X
Prozac	SSRI	X	X	X	X	
Mirtazapine	NaSSA			X	X	
Nefazodone	SARI	X				
Nortriptyline	TCA		X			
Paroxetine	SSRI	X	X	X	X	X
Paxil CR	SSRI	X	X	X	X	X
Zoloft	SSRI	X	X	X	X	X
Trazodone	SARI					
Effexor	SNRI	X	X	X	X	
Effexor-XR	SNRI	X	X	X	X	

Notes:
1. I report the prior probability that a drug will be effective from: (1) the baseline dynamic model and (2) the proportional hazard model. I use the characteristics of 10,000 sample patients, applied against the model estimates, to predict the prior probabilities. Sources:
(1) Crismon, M. L., et al. (1999). The Texas medication algorithm project: Report of the Texas consensus conference panel on medication treatment of major depressive disorder. J Clin Psychiatry, 60, 142-156.
(2) Karasu, T. B., et al. (2000). Practice guideline for the treatment of patients with major depressive disorder: Second edition. Arlington, VA: American Psychiatric Association. Also, Fochtmann, L. J., & Gelenberg, A. J. (2005). Guideline watch: Practice guideline for the treatment of patients with major depressive disorder, 2nd edition. Arlington, VA: American Psychiatric Association.
(3) Gartlehner, G., et al. (2007). Comparative effectiveness of second-generation antidepressants in the pharmacologic treatment of adult depression. Agency for Healthcare Research and Quality, US Dept of Health and Human Services.

based" copayment schemes or informational campaigns that emphasize treatment regimens that balance high efficacy with tolerability.

References

Ackerberg, Daniel A., "Advertising, learning, and consumer choice in experience good markets: an empirical examination," *International Economic Review*, 2003, *44* (3), 1007–1040.

Aguirregabiria, Victor and Pedro Mira, "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 2010, *156* (1), 38–67.

Akincigil, Ayse, John R. Bowblis, Carrie Levin, James T. Walkup, Saira Jan, and Stephen Crystal, "Adherence to antidepressant treatment among privately insured patients diagnosed with depression," *Medical Care*, 2007, *45* (4), 363–369.

Arcidiacono, Peter and Robert A. Miller, "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity," *Econometrica*, 2011, *79* (6), 1823–1867.

Bergemann, Dirk and Juuso Valimaki, "Learning and strategic pricing," *Econometrica*, 1996, *64* (5), 1125–1149.

Berndt, Ernst R., Anupa Bir, Susan H. Busch, Richard G. Frank, and Sharon-Lise T. Normand, "The medical treatment of depression, 1991-1996: productive inefficiency, expected outcome variations, and price indexes," *Journal of Health Economics*, 2002, *21* (3), 373–396.

Brezzi, Monica and Tze Leung Lai, "Optimal learning and experimentation in bandit problems," *Journal of Economic Dynamics and Control*, 2002, *27* (1), 87–108.

Centers for Medicare and Medicaid Services, U.S. Department of Health and Human Services, "Antidepressant Medications: U.S. Food and Drug Administration-Approved Indications and Dosages for Use in Adults," October 2015.

Chan, Tat Y. and Barton H. Hamilton, "Learning, private information, and the economic evaluation of randomized experiments," *Journal of Political Economy*, 2006, *114* (6), 997–1040.

Chang, Fu and Tze Leung Lai, "Optimal stopping and dynamic allocation," *Advances in Applied Probability*, 1987, *19* (4), 829–853.

Chernew, Michael E., Allison B. Rosen, and A. Mark Fendrick, "Value-based insurance design," *Health Affairs*, 2007, *26* (2), 195–203.

Ching, Andrew T., "Consumer learning and heterogeneity: Dynamics of demand for prescription

drugs after patent expiration," *International Journal of Industrial Organization*, 2010, *28* (6), 619–638.

_ **and Masakazu Ishihara**, "Measuring the informative and persuasive roles of detailing on prescribing decisions," *Management Science*, 2012, *58* (7), 1374–1387.

Chintagunta, Pradeep K., Renna Jiang, and Ginger Z. Jin, "Information, learning, and drug diffusion: the case of Cox-2 inhibitors," *Quantitative Marketing and Economics*, 2009, *7* (4), 399–443.

Coscelli, Andrea and Matthew Shum, "An empirical model of learning and patient spillovers in new drug entry," *Journal of Econometrics*, 2004, *122* (2), 213–246.

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

Crismon, M. Lynn, Madhukar Trivedi, Teresa A. Pigott, A. John Rush, Robert M. A. Hirschfeld, David A. Kahn, Charles DeBattista, J. Craig Nelson, Andrew A. Nierenberg, Harold A. Sackeim, and Michael E. Thase, "The Texas Medication Algorithm Project: report of the Texas consensus conference panel on medication treatment of major depressive disorder," *Journal of Clinical Psychiatry*, 1999, *60*, 142–156.

Cutler, David M., *Your money or your life*, Oxford University Press, 2004.

Efron, Bradley and Robert J. Tibshirani, *An introduction to the bootstrap*, Chapman and Hall, 1993.

Erdem, Tulin and Michael P. Keane, "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets," *Marketing Science*, 1996, *15* (1), 1–20.

Ferrari, Silvia and Francisco Cribari-Neto, "Beta regression for modelling rates and proportions," *Journal of Applied Statistics*, 2004, *31* (7), 799–815.

Fryback, Dennis G., Eric J. Dasbach, Ronald Klein, Barbara E. K. Klein, Norma Dorn, Kathy Peterson, and Patrica A. Martin, "The Beaver Dam Health Outcomes Study: initial catalog of health-state quality factors," *Medical Decision Making*, 1993, *13* (2), 89–102.

Gartlehner, Gerald, Richard A. Hansen, Patricia Thieda, Angela M. DeVeaugh-Geiss,

Bradley N. Gaynes, Erin E. Krebs, Linda J. Lux, Laura C. Morgan, Janelle A. Shumate, and Loraine G. Monroe, *Comparative effectiveness of second-generation antidepressants in the pharmacologic treatment of adult depression*, Agency for Healthcare Research and Quality (US), 2007.

Gelenberg, AJ, MP Freeman, JC Markowitz, JF Rosenbaum, ME Thase, MH Trivedi, and RS Van Rhoads, "American Psychiatric Association practice guidelines for the treatment of patients with major depressive disorder," *Am J Psychiatry*, 2010, *167* (Suppl 10), 9–118.

Gittins, John C., "Bandit processes and dynamic allocation indices," *Journal of the Royal Statistical Society.Series B (Methodological)*, 1979, pp. 148–177.

_ , *Multi-armed Bandit Allocation Indices*, New York: Wiley, 1989.

_ **and David M. Jones**, "A dynamic allocation index for the discounted multiarmed bandit problem," *Biometrika*, 1979, *66* (3), 561–565.

Heckman, James, "The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process," in Charles F. Manski and Daniel L. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, MA, 1981, pp. 179–195.

Hotz, V. Joseph and Robert A. Miller, "Conditional choice probabilities and the estimation of dynamic models," *The Review of Economic Studies*, 1993, *60* (3), 497–529.

Huskamp, Haiden A., Julie M. Donohue, Catherine Koss, Ernst R. Berndt, and Richard G. Frank, "Generic entry, reformulations and promotion of SSRIs in the US," *PharmacoEconomics*, 2008, *26* (7), 603–616.

Jovanovic, Boyan, "Job matching and the theory of turnover," *The Journal of Political Economy*, 1979, pp. 972–990.

Karasu, T. Byram, Alan Gelenberg, Arnold Merriam, and Philip Wang, *Practice Guideline for the Treatment of Patients With Major Depressive Disorder: Second Edition*, Arlington, VA: American Psychiatric Association, 2000.

Keane, Michael P. and Kenneth I. Wolpin, "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The*

Review of Economics and Statistics, Nov. 1994, *76* (4), pp. 648–672.

_ **and** _ , "The career decisions of young men," *Journal of Political Economy*, 1997, *105* (3), 473–522.

Kessler, Ronald C., Patricia Berglund, Olga Demler, Robert Jin, Doreen Koretz, Kathleen R. Merikangas, A. John Rush, Ellen E. Walters, and Phillip S. Wang, "The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R)," *JAMA*, 2003, *289* (23), 3095–3105.

Lai, Tze Leung, "Adaptive treatment allocation and the multi-armed bandit problem," *The Annals of Statistics*, 1987, *15* (3), 1091–1114.

_ **and H. Robbins**, "Asymptotically efficient adaptive allocation rules," *Advances in Applied Mathematics*, 1985, *6* (1), 4–22.

Lave, Judith R., Richard G. Frank, Herbert C. Schulberg, and Mark S. Kamlet, "Cost-effectiveness of treatments for major depression in primary care practice," *Archives of General Psychiatry*, 1998, *55* (7), 645–651.

Leffler, Keith B., "Persuasion or information-the economics of prescription drug advertising," *The Journal of Law and Economics*, 1981, *24*, 45–74.

Levy, Roy, *The pharmaceutical industry: a discussion of competitive and antitrust issues in an environment of change*, Washington, DC: Bureau of Economics/Federal Trade Commission staff report, 1999.

Mahajan, Aditya and Demosthenis Teneketzis, *Multi-armed bandit problems* Foundations and Applications of Sensor Management, Springer, 2007.

Melfi, Catherine A., Anita J. Chawla, Thomas W. Croghan, Mark P. Hanna, Sean Kennedy, and Kate Sredl, "The effects of adherence to antidepressant treatment guidelines on relapse and recurrence of depression," *Archives of General Psychiatry*, 1998, *55* (12), 1128–1132.

Miller, Robert A., "Job matching and occupational choice," *The Journal of Political Economy*, 1984, *92* (6), 1086–1120.

Mortensen, Dale T., "Job search and labor market analysis," *Handbook of Labor Economics*,

1986, *2*, 849–919.

Murphy, Michael J., Ronald L. Cowan, and Lloyd I. Sederer, *Blueprints: Psychiatry*, 5 ed., Philadelphia, PA: Lippincott Williams Wilkins, 2009.

Narayanan, Sridhar, Puneet Manchanda, and Pradeep K. Chintagunta, "Temporal differences in the role of marketing communication in new product categories," *Journal of Marketing Research*, 2005, *42* (3), 278–290.

Pandey, Sandeep, Deepayan Chakrabarti, and Deepak Agarwal, "Multi-armed bandit problems with dependent arms," in "Proceedings of the 24th international conference on Machine learning" ACM 2007, pp. 721–728.

Philipson, Tomas and Jeffrey DeSimone, "Experiments and subject sampling," *Biometrika*, 1997, *84* (3), 619–630.

Pomerantz, Jay M., Stan N. Finkelstein, Ernst R. Berndt, Amy W. Poret, Leon E. Walker, Robert C. Alber, Vidya Kadiyam, Mitali Das, David T. Boss, and Thomas H. Ebert, "Prescriber intent, off-label usage, and early discontinuation of antidepressants: a retrospective physician survey and data analysis," *Journal of Clinical Psychiatry*, 2004, *65* (3), 395–404.

Rothschild, Michael, "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, 1974, *9* (2), 185–202.

Rusmevichientong, Paat and John N. Tsitsiklis, "Linearly parameterized bandits," *Mathematics of Operations Research*, 2010, *35* (2), 395–411.

Rust, John, "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, 1987, pp. 999–1033.

Soumerai, Stephen B. and Jerry Avorn, "Principles of educational outreach ('academic detailing') to improve clinical decision making," *JAMA*, 1990, *263* (4), 549–556.

Vuong, Quang H., "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica*, 1989, pp. 307–333.

Whittle, P., "Multi-armed bandits and the Gittins index," *Journal of the Royal Statistical Society.Series B (Methodological)*, 1980, *42* (2), 143–149.

_ , "Arm-acquiring bandits," *The Annals of Probability*, 1981, pp. 284–292.

A Appendix

A.1 Formation of the Sample

To form a sample, I identify patients in the outpatient data who receive a new diagnosis in one of the following depression categories, listed along with their International Classification of Diseases (ICD-9-CM) codes: major depression (296.2, 296.3); dysthymia and depression with anxiety (300.4); prolonged depressive reaction (309.0); adjustment disorder with depressed mood (309.1); and, depression not otherwise specified (311). Starting from this set of depression patients, I collect panel data on the patients' prescription drug use over all months of the sample. I impose the following additional screens in creating the analysis dataset: patients cannot have a concurrent diagnosis of bipolar disorder or schizophrenia or receive drugs that signal these conditions, as these more complicated illnesses require distinct prescribing behavior;[34] the patients' age must be between 18 and 64, the range for which the data are complete; patients must visit a health professional with the ability to prescribe all possible treatments in the choice set; and, patients must not be pregnant, a condition that involves serious drug contraindications.

Within the prescription history for each patient, I define episodes of care. Starting with the set of patients who are diagnosed with depression, I identify the episode start and end date using the following steps.

1. I flag and exclude patients with left-censored episodes. To do so, I first identify a subscriber's entry date in the Marketscan database using the patient's enrollment date in insurance; in this way, I do not require the patient to have an observed claim to identify her appearance in Marketscan. Using that entry date, I flag censored observations as those patients who have a diagnosis of depression or a prescribed antidepressant in the first six months upon entry in the dataset. I choose six months as my threshold based on the definition of a "recurrence" in the medical literature, as in Melfi et al. (1998).

[34] Patients excluded due to comorbidities have a diagnosis with one of the following ICD-9-CM codes: bipolar and manic disorders (296.0, 296.1, 296.4-.8) and schizophrenic disorders (295.0-295.9).

2. I define an episode start as either the date of the first office visit with a depression diagnosis and no drugs filled or the first observed antidepressant prescription filled. In this way, if a patient has a depression diagnosis but no prescription filled, she would appear in an episode with the choice of 'outside option.'

3. I define an episode end date as the last observed prescription–new or refill–where there is a gap of 90 days or more following the date of the last fill + the days supply of medication at that date. For example, if a patient last filled a 28 day prescription on July 1, followed by a gap in depression care of 90 days, the "last date" will be July 29th. If a patient episode begins with depression office visit followed by no follow-up care, the end date is 30 days after the office visit.

4. To define the number of periods within an episode, I count the number of 30 days intervals that appear between the start and end date of the episode. If the end date occurs at a date that is not a multiple of 30 days from the start date, we set the counter to add a period if the last segment is greater than or equal to 15 days and do not add an additional segment if the last counter is less than or equal to 14 days.

5. If after the 90 day gap, the patient appears with another depression office visit or another depression drug, we define this sequence as another episode and set its start and end date as above.[35]

Using this definition of an episode, 35% of patients do not fill a prescription in their episode; 22% of patients fill one; 13% fill two; 9% fill three; and 6% fill four. 15% have episodes involving between 5 and 30 filled prescriptions.

We define switches within an episode when a patient's prescription switches from one drug in the choice set to another. To count the number of periods on each drug, we collect the start date as the first time a drug is observed being filled in a patient's sequence and the end date as the last date a drug is filled plus its days supply. We divide the number of days in this interval by 30, and count as an additional period any partial set of days of length greater than or equal to 15.

[35]Roughly 8% of the unique patient identifiers repeat in a new episode over the course of my sample period. In a sensitivity, I remove these observations and find similar results.

The initial filters lead to a dataset of 102,780 unique patient episodes of depression care, comprised of 267,390 observations on antidepressant prescriptions. The individuals are members of 307 insurance plans. To estimate the econometric model, I use a smaller sample of 10,000 patient episodes drawn randomly from the original dataset using a multinomial model. I set the drug characteristics, including price, equal to the plan-specific and year-specific drug characteristics that correspond to the patient's plan at the time of his initial diagnosis. The rich variation in drug copayments appears in Figure A1.

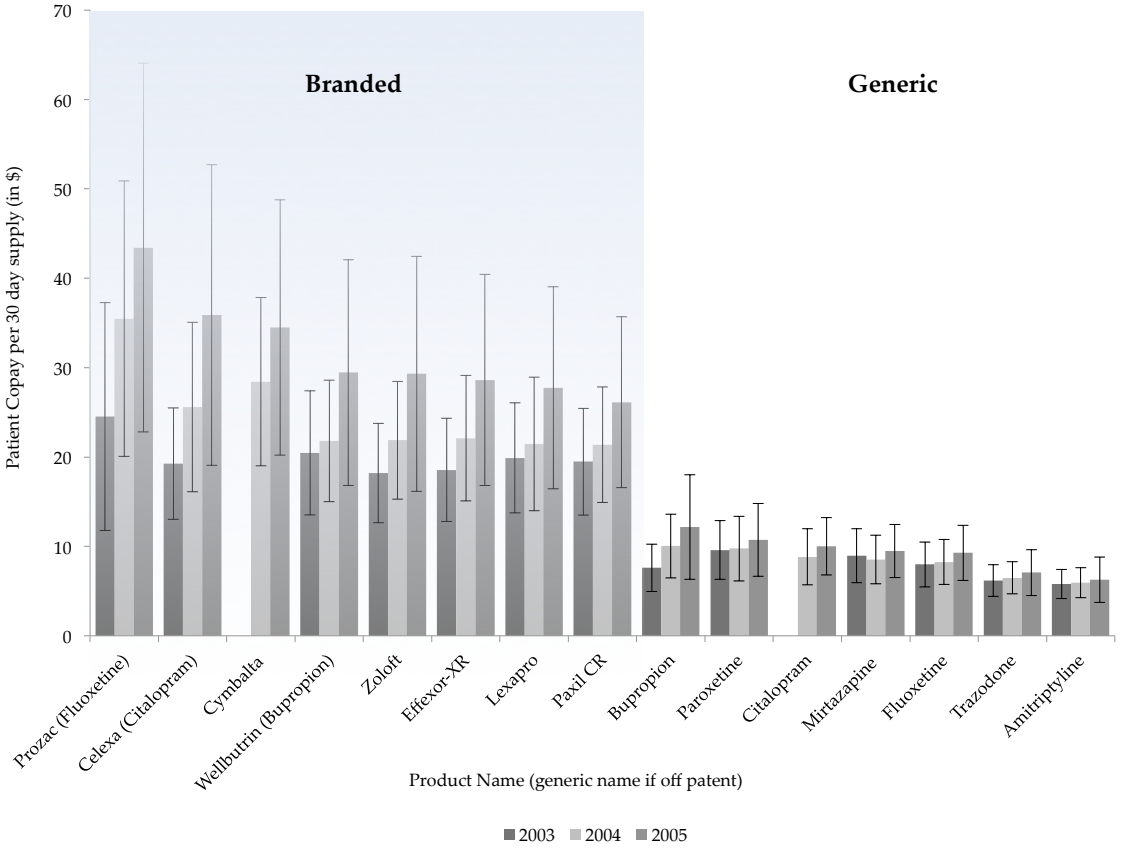

Figure A1: Patient copayments by product and year. Error bars show the standard deviation in copayments across insurance plans in the sample.

A.2 Two-stage panel data model without learning

The two-stage model uses the timing and identity of the observed transitions to estimate a hazard model of switching in piecewise fashion over each period of treatment. In the first stage, I estimate

a piecewise proportional hazard model. The probability that choice j is found ineffective in period 1 is:

$$\begin{aligned} P(0 &\leq t < 1 | t \geq 0, X) = 1 - \alpha_1(X, \theta) \\ \text{where } \alpha_m(X, \theta) &= \exp\left(- \int_{a_{m-1}}^{a_m} \exp(X\beta)\lambda_m ds \right) = \exp\left(- \exp(X\beta)\lambda_m (a_m - a_{m-1}) \right) \end{aligned}$$

In the case of $t = 1$:

$$\begin{aligned} P(0 &\leq t < 1 | t \geq 0, X) = 1 - \alpha_1(X, \theta) \\ &= 1 - \exp\left(- \exp(X\beta)\lambda_1 \right) \end{aligned}$$

The utility function for the agent's discrete choice incorporates this probability of finding a drug ineffective by weighting the deterministic component of utility by this probability:

$$\begin{aligned} U_{ijt} &= (1 - \Pr(j \text{ found ineffective at } t-1)) * Z_{ij}\theta + \varepsilon_{ijt} \\ &= \exp(-\exp(X\beta)\lambda_{t-1}) * Z_{ij}\theta + \varepsilon_{ijt} \end{aligned}$$

For periods in which the patient and physician choose a new drug, k, that they have not tried previously, the probability term, $\Pr(j \text{ found ineffective at } t-1)$ equals 0 and the expression for utility reduces to:

$$U_{ikt} = Z_{ik}\theta + \varepsilon_{ikt}$$

The choice probabilities for this two-stage model take a logit form. I illustrate the form of the choice probabilities for an example in which the patient and physician sampled drug j for m periods but have not yet sampled other drugs $k \neq j$:

$$\begin{aligned} P_{i,j,m+1} &= \frac{\exp\left(\exp(-\exp(X\beta)\lambda_m) Z_{ij}\theta\right)}{\exp\left(\exp(-\exp(X\beta)\lambda_m) Z_{ij}\theta\right) + \sum_{k \neq j} \exp(Z_{ik}\theta)} \\ P_{i,k,m+1} &= \frac{\exp(Z_{ik}\theta)}{\exp\left(\exp(-\exp(X\beta)\lambda_m) Z_{ij}\theta\right) + \sum_{k \neq j} \exp(Z_{ik}\theta)} \end{aligned}$$

I estimate the second-stage discrete choice model using maximum likelihood, conditioning on the estimates of (β, λ) from the first-stage piecewise proportional hazard model estimation described in Section 2.2.2. In Table A1 and Table A2, I report the estimates and predicted probabilities from the first stage hazard model; in Table A3 I report the estimates from the second stage of the model. I convert these estimates to predicted choices and report these in Table A4 for three periods of treatment. Comparing the predicted probabilities from this two-stage model to the predicted probabilities from the learning model in the main text, the two-stage model does, broadly, capture the choice behavior in the first period. It misses on the extensive margin, however, predicting 48% of patients will quit at the first month relative to 36% in the raw data. The transitions over time also miss the richer switching patterns seen in the data, instead predicting a slow decline in each drug's share as patients move toward the outside good of no drug treatment.

Table A1: Estimates, piece-wise proportional hazard model (in %)

Piecewise proportional hazard model estimates			
Covariate	Est	S.E.	T-stat
Constant	-1.63	0.25	6.63
1Diagnosis of major depressive disorder	-0.11	0.04	2.66
1SNRI*1Diagnosis of MDD	-0.04	0.09	0.42
1NDRI*1Diagnosis of MDD	0.06	0.11	0.54
1NaSSA*1Diagnosis of MDD	-0.26	0.34	0.77
1SARI*1Diagnosis of MDD	0.12	0.14	0.84
1TCA*1Diagnosis of MDD	-0.17	0.30	0.57
1SNRI	-0.58	0.11	5.25
1NDRI	0.00	0.05	0.07
1NaSSA	0.59	0.21	2.85
1SARI	0.43	0.17	2.58
1TCA	0.75	0.13	5.95
% of nausea reports in clinical trials	1.59	0.51	3.12
1frequency of dosing >1x/day	0.44	0.09	4.69
1reformulation	0.02	0.04	0.58
1branded	0.40	0.05	8.66
copayment, in $/day	-0.29	0.04	7.19
Time period 1	1.43	0.30	4.80
Time period 2	1.03	0.21	5.03
Time period 3	0.83	0.17	5.02
Time period 4	0.85	0.18	4.60
Time period 5	0.75	0.17	4.54
Time period 6	0.71	0.15	4.80
Time period 7	0.60	0.13	4.67
Time period 8	0.65	0.15	4.28
Time period 9	0.52	0.12	4.35
Time period 10	0.79	0.21	3.73
Time period 11	0.89	0.32	2.79

Table A2: Predicted probability that a product is effective, using estimates from the piece-wise proportional hazard model (in %)

				Periods since the patient's initial diagnosis		
Ingredient	Product	Subclass	Brand?	Period 1	Period 2	Period 3
Amitriptyline	Amitriptyline	TCA	N	51.7	62.2	68.2
Nortriptyline	Nortriptyline	TCA	N	53.0	63.3	69.2
TCA average				52.4	62.7	68.7
Citalopram	Citalopram	SSRI	N	72.0	79.0	82.7
Citalopram	Celexa	SSRI	Y	66.1	74.2	78.6
Escitalopram	Lexapro	SSRI	Y	64.6	73.0	77.6
Fluoxetine	Fluoxetine	SSRI	N	70.6	77.8	81.7
Fluoxetine	Prozac	SSRI	Y	65.5	73.8	78.2
Paroxetine	Paroxetine	SSRI	N	69.3	76.8	80.8
Paroxetine	Paxil CR	SSRI	Y	58.3	67.8	73.1
Sertraline	Zoloft	SSRI	Y	62.4	71.2	76.1
SSRI average				66.1	74.2	78.6
Duloxetine	Cymbalta	SNRI	Y	67.4	75.2	79.5
Venlafaxine	Effexor	SNRI	Y	60.2	69.4	74.5
Venlafaxine	Effexor-XR	SNRI	Y	71.1	78.2	82.0
SNRI average				66.2	74.3	78.7
Bupropion	Bupropion	NDRI	N	60.5	69.6	74.7
Bupropion	Wellbutrin XL	NDRI	Y	61.8	70.7	75.6
NDRI average				61.1	70.2	75.1
Mirtazapine	Mirtazapine	NaSSA	N	59.2	68.5	73.7
NaSSA average				59.2	68.5	73.7
Nefazodone	Nefazodone	SARI	N	41.7	53.3	60.2
Trazodone	Trazodone	SARI	N	45.7	56.9	63.5
SARI average				43.7	55.1	61.8

Table A3: Estimates, Two-stage Discrete Choice Model with Hazard Weighting

Covariate	Est
1{Amitriptyline}	-4.80
1{Bupropion}	-2.41
1{Citalopram}	-2.72
1{Duloxetine}	-2.28
1{Escitalopram}	-0.80
1{Fluoxetine}	-2.21
1{Mirtazapine}	-4.78
1{Trazodone}	-4.84
1{Nortriptyline}	-5.09
1{Paroxetine}	-2.63
1{Sertraline}	-1.06
1{Venlafaxine}	-1.74
1{branded}	-0.80
copayment	-0.07

Table A4: Matching predicted shares under two-stage model to observed shares

			Raw Data			Two-stage		
Product Name	Subclass	Brand?	t=1	t=2	t=3	t=1	t=2	t=3
None	-	-	36.20	50.80	59.57	47.55	68.43	78.72
Amitriptyline	TCA	N	0.69	0.49	0.28	0.42	0.25	0.17
Bupropion	NDRI	N	3.50	2.15	1.52	4.48	2.62	1.74
Wellbutrin XL	NDRI	Y	5.22	3.55	2.52	2.18	1.44	1.01
Citalopram	SSRI	N	2.33	1.54	1.26	1.88	1.02	0.65
Celexa	SSRI	Y	1.53	1.00	0.78	1.42	0.84	0.60
Cymbalta	SNRI	Y	1.24	0.92	0.68	1.42	0.78	0.51
Lexapro	SSRI	Y	11.95	8.01	5.43	10.18	6.16	4.12
Fluoxetine	SSRI	N	9.87	6.69	4.39	5.91	3.72	2.50
Prozac	SSRI	Y	0.35	0.27	0.24	2.21	1.23	0.83
Mirtazapine	NaSSA	N	0.59	0.38	0.24	0.42	0.25	0.17
Nefazodone	SARI	N	0.11	0.10	0.08	0.36	0.20	0.13
Nortriptyline	TCA	N	0.44	0.25	0.19	0.30	0.17	0.11
Paroxetine	SSRI	N	3.90	2.63	1.86	3.54	2.13	1.44
Paxil CR	SSRI	Y	2.88	1.85	1.27	1.66	1.04	0.72
Zoloft	SSRI	Y	10.18	6.92	4.43	7.93	4.88	3.27
Trazodone	SARI	N	1.49	0.80	0.43	0.43	0.25	0.17
Effexor	SNRI	Y	0.44	0.28	0.17	3.62	2.01	1.34
Effexor-XR	SNRI	Y	7.09	5.14	4.00	4.09	2.58	1.79

A.3 Gittins' Index rule

I consider a market with J choices available. The outcome under j, Y_j, is a draw from a univariate distribution, $f(y;\theta_j)$, where θ_j is an unknown parameter vector. Gittins and Jones (1979) and the broader literature on "multi-armed bandits" sought to solve the following basic but non-trivial question: how should a decision-maker select from among the j choices in each period t to maximize the expected discounted sum of his payoffs:

$$E_{\theta_1,...,\theta_J} \sum_{t=1}^{\infty} \delta^{t-1} Y_t(\theta)$$

for a given discount rate, δ. Here, $\theta = (\theta_1, ..., \theta_J)$ are unknown. Agents form independent priors, $\Pi^{(j)}$, on θ_j for $j = 1, ..., J$. The optimal allocation rule maximizes:

$$\int ... \int E_{\theta_1,...,\theta_J} \left(\sum_{t=1}^{\infty} \delta^{t-1} Y_t(\theta) \right) d\Pi^{(1)}(\theta_1) \cdots d\Pi^{(J)}(\theta_J)$$

One can solve this sequence problem using a dynamic programming approach. Gittins and Jones (1979) and Whittle (1980) show that in the case in which agents form independent priors on $(\theta_1, ..., \theta_J)$ and several other conditions hold, there is a simpler solution. The agent can apply a forward induction rule or "index" rule. He computes J one-dimensional optimal stopping problems to determine the value of J indices in each period t. The agent then chooses the option with the largest index. I describe briefly the form of this index rule; the original contributions cited above provide detailed proofs.

In every period, the agent selects one of J choices and realizes an outcome, Y_{jt}. Let $\psi(t)$ represent the action the agent takes at t.[36] I denote the number of times option j has been selected up to and including period t as:

$$n_j(t) = \sum_{s=1}^{t} 1\{\psi(s) = j\}$$

I previously defined $\Pi^{(j)}$ as the agent's prior beliefs on the distribution of θ_j. The posterior

[36] From Chang and Lai (1987), the allocation rule, $\psi(t)$ can be specified by a sequence of random variables, $\psi(1), \psi(2), ...$, for periods $t = 1, 2, ...$ such that the event $\psi(t+1) = j$ for $j = 1, ..., J$ belongs to the σ-field generated by the past observed sequence $\psi(1), Y_1, ..., \psi(t), Y_t$.

distribution at time t for j after $n_j(t)$ outcome realizations under j is $\Pi^{(j)}_{n_j(t)}$. That is, the agent observes $(Y_{j,1}, ..., Y_{j,n_j(t)})$ for the $n_j(t)$ periods when he selects j and updates his prior beliefs on the distribution of θ_j from $\Pi^{(j)}$ to $\Pi^{(j)}_{n_j(t)}$.

Gittins' Index is a function of the prior distribution of each independent option, j:

$$G(\Pi^{(j)}) = \sup_{\tau} \left\{ \frac{\int E_{\theta_j} \left(\sum_{s=0}^{\tau-1} \delta^s Y_{j,s+1} \right) d\Pi^{(j)}(\theta_j)}{\int E_{\theta_j} \left(\sum_{s=0}^{\tau-1} \delta^s \right) d\Pi^{(j)}(\theta_j)} \right\}$$

where the supremum is over all stopping times $\tau \geq 1$ defined on $\{Y_{j,1}, Y_{j,2}, ...\}$. After $n_j(t)$ realizations under choice j, the Gittins' Index for j becomes $G(\Pi^{(j)}_{n_j(t)})$. The agent need only calculate $\{G(\Pi^{(1)}_{n_1(t)}), ..., G(\Pi^{(J)}_{n_J(t)})\}$ at t. He then chooses the option with the maximal $G(.)$.

An alternative way to express Gittins' Index given $\Pi^{(j)}$=$\Pi^{(j)}_0$ is as the infimum of the set of solutions M of the equation:

$$\sup_{\tau} \int E_{\theta_j} \left\{ \sum_{s=0}^{\tau-1} \delta^s \int \mu(\theta_j) d\Pi^{(j)}_s(\theta_j) + M \sum_{s=\tau}^{\infty} \delta^s \right\} d\Pi^{(j)}(\theta_j) = M \sum_{s=0}^{\infty} \delta^s$$

where $\mu(\theta_j) = E_{\theta_j}(Y_j)$, the mean outcome under j conditional on θ.

Intuitively, the agent solves an optimal stopping problem in which he decides between the choice j and a standard project that yields a constant reward, M.[37] This 'retirement' value is specific to each option. The agent maximizes the expected discounted value of playing choice j for the optimal number of periods τ given the posterior distribution $\Pi^{(j)}_{n_j(t)}$. In the remaining time from τ forward, he receives M. The infimum of the set of solutions for M is equivalent to the Gittins' Index, $G(\Pi^{(j)}_{n_j(t)})$ given a particular posterior distribution.

In the main analysis, outcomes take the following form:

$$Y_{ijt} \backsim p_{ij}(\theta)^k (1 - p_{ij}(\theta))^{1-k}, k \in \{0, 1\}$$

Again, agents do not know $p_{ij}(\theta)$ for $j = 1, ..., J$ at the time of the decision. They form a prior distribution over $p_{ij}(\theta)$, where $\theta = (\mu_{ij}, \phi)$. In the main text, μ_{ij} denotes the expected probability

[37] See Whittle (1980) and Brezzi and Lai (2002).

of success on drug j and ϕ denotes a measure of the precision of the agents priors, common across choices. I reduce the dimensionality of $\theta = (\mu_{i1}, ..., \mu_{iJ}, \phi)$ by letting $\mu_{ij} = \exp(X'_{ij}\gamma_1)$ and $\phi = \exp(\gamma_2)$. Then, θ becomes $\gamma = (\gamma_1, \gamma_2)$. Agents form priors $\Pi^{(r)}$ for the $r = 1, ..., R$ elements in γ.

A.4 Computational Details

A.4.1 Diffusion Approximation to the Index Rule

To compute the experimentation incentive term in the index rule, I rely on a numerical solution provided by Chang and Lai (1987). The authors numerically solve the optimal stopping problem underlying the index rule by using a change of variables to reframe the problem as a Brownian motion. Chang and Lai (1987) carry out Monte Carlo simulations to demonstrate the robustness of this numerical solution.

The closed form approximation to the optimal stopping boundary in the Brownian motion problem takes the following form in the finite horizon case:

$$h(s) = \left\{ \begin{array}{c} \{2\log(s^{-1}) - \log(\log(s^{-1})) - 1og(16\pi) + ... \\ .99\exp(-.038s^{-1/2})\}^{1/2} \text{ if } 0 < s \leq .01 \\ -1.58\sqrt{s} + 1.53 + 0.07s^{-1/2} \text{ if } .01 < s < .28 \\ -.576s^{3/2} + .299s^{1/2} + .403s^{-1/2} \text{ if } .28 < s \leq .86 \\ s^{-1}(1-s)^{1/2}\{.639 - .403(t^{-1} - 1)\} \text{ if } .86 < s \leq 1 \end{array} \right\}$$

where $s = t/T$, with T the finite horizon and t the point in the episode along the sequence to T.

In the discounted infinite horizon case, Chang and Lai (1987) replace the closed-form $h(s)$ with $\psi(s)$:

$$\psi(s) = \left\{ \begin{array}{cc} \sqrt{s/2} & \text{if } s \leq 0.2 \\ .49 - .11s^{-1/2} & \text{if } 0.2 < s \leq 1 \\ .63 - .26s^{-1/2} & \text{if } 1 < s \leq 5 \\ .77 - .58s^{-1/2} & \text{if } 5 < s \leq 15 \\ \{2\log(s) - \log(\log(s)) - \log(16\pi)\}^{1/2} & \text{if } s > 15 \end{array} \right\}$$

The discussion of s in the index rule in Section 3.2.2 follows from Brezzi and Lai (2002), who define the experimentation incentive as a function of $\psi(.)$.

A.4.2 Computational time

I include below statistics on the computational time necessary to find the maximum likelihood estimates for the three main specifications in the model. These computation times come from using Knitro optimization software run through Matlab, on a desktop computer with two 2.4 GHz Quad-Core Intel Xeon processors, with 16 GB 1066 MHz memory. Using an uninformed starting value, the time to compute the coefficients from maximum likelihood estimation of the two-level dynamic model is 61.91 hours with 10,000 patient episodes. With a "warm start" to the optimization, the time to convergence is considerably smaller, at 9.4 hours.

A.5 Determining the Dollar Value of Symptomatic Relief

To assign a dollar value to the relief of symptoms that prescription drug treatments provide, I carry out the following procedure.

First, I save the predicted sequence of treatments for the 10,000 patients from the simulation of the baseline policy and the counterfactual policies. I divide the sample of patients in each counterfactual into 5 categories based on the treatments used within their episode: (1) no drug care; (2) use of a TCA for one month or less, followed by exit from drug care; (3) use of a TCA for greater than one month; (4) use of any second generation (non-TCA) drug class for one month or less; and, (5) use of any second generation drug class for greater than 1 month. I choose these divisions to match a subset of the treatment categories studied by Berndt et al. (2002). They employ a panel of medical experts to rate the likelihood of both full and partial recovery given a patient's background and a treatment regime and duration. For example, in Table 1 of Berndt et al. (2002), treatment under an SSRI with 1-3 office visits produces a probability of full remission of .20, and a probability of partial remission of .45 over 16 weeks. If the SSRI treatment extends for greater than 30 days, those probabilities increase to .28 and .60, respectively.

Given the probabilities of full and partial remission over 16 weeks for each of the 5 categories above, I calculate the expected number of weeks out of 16 that an individual suffers either the full symptoms of depression or partial symptoms. I follow a similar procedure to that of Cutler (2004). Following the medical literature, I set the recovery rate at 0% in the first two weeks of treatment, before an antidepressant takes full effect. In every week for the next 16 weeks, I set a fixed rate of full and partial recovery for those who have yet to recover such that the share of full and partial recovery patients at the end of the 16 week period equals the median rate predicted by the expert panel in Berndt et al. (2002) for the category of treatment. I then use these per-week probabilities to calculate the expected number of weeks without full or partial recovery. In the case without drug care, the patient suffers from depression for 11.9 out of the 16 weeks, with partial recovery in an additional 2.8 weeks. The expected weeks of depression and of partial depression for drug treatment differ a bit by the type and duration care: with less than one month of TCAs or SSRIs, the expected number of weeks equals approximately 10.3 and 3.9; for longer duration on SSRIs, the expected number of weeks of full and partial depression change to 7.4 and 5.9, respectively.

Given the expected duration of complete or partial depressive symptoms, I multiply the savings in "depression weeks" against the quality disutility from depression, estimated in Lave et al. (1998) to equal -.41 for full depression. That is, patients equate 10 years living with depression to roughly 6 years living without depression. As in Cutler (2004), I set the disutility from partial depression at half the full depression rate. Multiplying this savings in utility against $100,000 as the value of a year of life, I find the per patient dollar value of treatment in each of the four treatment categories, normalizing by the dollar value of no treatment. I multiply this by the share of patients in each category in a particular counterfactual, summing across categories to get a total dollar savings.

In Table 7, I report these the dollar savings for three months of treatment under the counterfactual policies. The alternative policies change the distribution of treatment types, driving the observed differences in the dollar value gained. When a policy increases adherence and promotes use of SSRIs, for example, patients shift toward treatment regimens with higher recovery probabilities; the effect is to increase the expected number of weeks with relief of symptoms in the counterfactual relative to the baseline.