UNDERSTANDING PHYSICIAN DECISION MAKING: THE CASE OF DEPRESSION

JANET M CURRIE* AND W. BENTLEY MACLEOD**

Abstract. Faulty physician decision making has been blamed for everything from medical errors to excessive procedure use and wasteful spending. Yet medical treatment is often complex, requiring a sequence of decisions that may involve trade offs between selecting the choice with the highest expected value or selecting a choice with higher possible payoffs. We show that the best choice depends on a physician's diagnostic skill so that the optimal treatment can vary even for identical patients. Bringing the model to patient claims data for depression, we show that doctors who experiment more with drug choice achieve better patient outcomes, except when physician decisions violate professional guidelines for drug choice.

1. Introduction

In 2000, the National Academy of Medicine published a report entitled "To Err is Human," highlighting the importance of medical errors which, according to the report, kill 98,000 U.S. patients annually. In addition, commentators such as Fuchs (2004) have blamed "the idiosyncratic beliefs of physicians [and] the parochial character of much clinical practice," that is, poor physician decision making, for much wasteful spending. For example, Finkelstein et al. (2015) report that health care expenditures on the average elderly person in Miami, FL were $14,423 in 2010, adjusted for age, sex, and race, compared to $7,819 for the average elderly person in Minneapolis, MN. Other researchers have found that these large differences in spending and utilization of care are largely unrelated to health outcomes (Fisher et al. (2003)).

This paper examines variations in physician decision making in the context of the prescribing of anti-depressant medications in the U.S. Anti-depressants are one of the largest and fastest growing classes of drugs. In the U.S., the fraction of the adult population that have taken an anti-depressant in the past 30 days has doubled from 6.8% to 13% between 1999/2000 and 2011/12 (Kantor et al. (2015)).[1] Depression has been blamed for rising suicide rates in the U.S., with suicide ranking as the 10th leading cause of death in 2016 (https://www.nimh.nih.gov/health/statistics/suicide.shtml). In

Date: August 14, 2018.
This paper was originally prepared for the Fisher-Schultz lecture of the Econometric Society delivered in Lisbon on August 23, 2017. We are grateful for the helpful comments received at that time, as well as from audiences at Yale, Stanford, University of Mannheim, University of Pittsburgh, Rochester University, Tel Aviv University, University College London, and the International Association of Applied Economists. The authors thank Allen Campbell of IQVIA and the members of the Blue Cross/Blue Shield Association for working with us to help us access their data. Xuan Li, Sara Shahanaghi, Haowen (Alice) Wu and Emily Cuddy provided outstanding research assistance. We thank the NIA for support under P30-AG024928-14. The authors are solely responsible for the content of the paper.
[1]Other countries have seen a similar increase in anti-depressant use over time. For example, the EU average utilization rate rose from just under 3% to almost 6% between 2000 and 2010 (http://www.oecd.org/els/health-systems/Item10_Trends-in-antidepressant-consumption_NIPH.pdf).

addition to its overall importance, several features of the anti-depressant market make it an especially interesting context for studying physician decision making as reflected in physician practice style. First, prices are likely to be relatively unimportant in decision making about anti-depressants today since most anti-depressants are now available as generics and many patients face the same cost (a small co-pay) for many drug choices. Hence, anti-depressants belong to the large and interesting category of markets that clear largely without the aid of price signals (Roth (2018)). Second, as Frank and McGuire (2000) observe, the assessment of patient condition is often much more difficult in the case of mental illness than for many physical illnesses, suggesting that physician diagnostic skill plays an important role.[2]

Third, there are many anti-depressants available (32 separate molecules over our study time period), and since every patient responds differently, and there is no one drug that dominates all others, finding an optimal treatment will necessarily involve experimentation with different drugs over time. This classic multi-armed bandit problem, introduced in economics by Rothschild (1974), implies a trade off between choosing the treatment with the highest expected value at a point in time, and experimenting to learn more about treatments that may be better for a particular individual. Since the physician cannot perfectly predict which drug is best for a patient, effectiveness cannot be known until a patient starts taking a drug.[3] The multi-armed bandit problem is about finding the best rule for choosing the sequence of drugs to prescribe as a function of the observed effects of past treatment. One important feature of the problem is that the effects of a drug may be correlated with other drugs. Thus we have what is known as a correlated armed problem for which there is no known closed form solution.[4]

The goal of this paper is to better understand the sources of variation in physician practice style in such a market, and how these variations contribute to the performance of the health care system. To do so, we develop a tractible Bayesian model using the upper confidence bound (UCB) algorithm from the machine learning literature to explore the relationship between short run exploitation of current knowledge (what Brezzi and Lai (2002) call Bayesian myopic behavior) and experimentation to learn more.[5] Reverdy et al. (2014) show that this algorithm does a good

[2]See for example the guidelines which are quite vague regarding best practices. The main tool used in a psychiatrist's office is the patient interview. In contrast, a person seeing a physician with a physical ailment often has a battery of diagnostic tools available, including comprehensive blood tests, x-rays, MRI scans, heart monitors, fetal monitors and so on. In future, genetic testing may be used to choose treatments for depression, but such screening is still in the experimental stage (See Hicks et al. (2015)).

[3]The multi-armed bandit can be viewed as choosing an arm to pull at a slot machine. One does not learn the consequence of a choice until the armed is pulled. Then one decides whether to continue pulling the arm or to switch to a different arm. In our context, the patient corresponds to a slot-machine and the arm is a drug choice.

[4]See Gittins et al. (2011) and Cappé et al. (2013) for a discussion of the limitations of the single index approach. There are some results for extreme cases, such as Klein and Rady (2011). See Bergemann and Valimaki (2006) for a review of the earlier literature and Kendrick et al. (2014) for a recent review of the use of dynamic programming in economics, and the so called curse of dimensionality.

[5]The machine learning literature has developed a class of easy to implement heuristic algorithms known as upper confidence limit (UCL) or upper confidence bound (UCB) algorithms. There is a growing literature using such quasi-Bayesian approaches to model observed human behavior for bandit type problems, including Acuna and Schrater (2010), Mathys et al. (2011), Lee et al. (2011), and Reverdy et al. (2014). The literature on the UCB algorithm begins with Lai and Robbins (1985) who provide bounds on the speed of convergence to the long run optimal solution. Recent extensions include Auer et al. (2002), Brezzi and Lai (2002), Abernethy et al. (2007) , Cappé et al. (2013) and Srinivas et al. (2012).

job of modeling observed human choice in laboratory settings. It is computationally much less complex than the Gittins index, yet there are a number of important results, beginning with Lai and Robbins (1985), which imply that the algorithm converges to the optimal solution at the rate $O\left(\frac{1}{log(T)}\right)$.[6] The algorithm also allows for correlations in treatment effects.

We are able to model the physician's taste for experimentation (or degree of satisficing, see Simon (1955)) using a single parameter τ, known in the UCB literature as the *tuning parameter*. When the tuning parameter is zero, ($\tau = 0$), the physician chooses the option with the highest expected value. When the tuning parameter is positive ($\tau > 0$), the payoff from each choice is given by the expected value of treatment plus τ times a term that increases with the variance of treatment. The decision maker may select options that have high variance rather than the highest expected mean payoff. After a decision is made, the physician observes the outcome, and updates her beliefs. When the tuning parameter is large, then the physician experiments with more drug possibilities. Hence, the model can integrate the different physician practice styles discussed in, for example, Frank and Zeckhauser (2007), into a single framework.

A novel feature of our model is that the trade-off between current and future returns to experimentation depends not only on physician tastes, beliefs, and the time horizon with the patient (or alternatively, a discount factor), but also on physician diagnostic skill. If a physician is not able to correctly assess a patient's condition then the patient may be better off without experimentation. However, if a physician is highly skilled in the sense of being able to correctly assess a patient's condition, i.e. has excellent diagnostic skill, then the gains from learning are larger. One insight of the model is that there will not necessarily be a single correct treatment even for identical patients since the best treatment will also depend on the doctor's characteristics. Other things being equal, a skilled diagnostician should be more experimental than a less skilled diagnostician.

The model is used to illustrate how patient health is likely to vary with physician diagnostic skill and taste for experimentation, and how the two features interact. These results are then used to motivate the use of an entropy score to access the importance of physician practice style for patient outcomes in our data. The model makes predictions regarding the relationship between a physician's practice style and observed variability in prescribing behavior: Physicians who are more experimental use a wider variety of drugs for their patients, and physicians with better diagnostic skills stand to learn more from experimentation and so will be more experimental. Experimentation may make it more likely that the physician finds the optimal treatment, but it may also make it more likely that the physician violates prescribing guidelines, with potentially negative consequences for patients.

Explorations of the relationship between physician practice style, guidelines, and patient outcomes are carried out using novel data formed by merging information on all anti-depressant prescriptions for each doctor to newly available national patient claims data on hundreds of thousands of patients treated with anti-depressants from Blue Cross/Blue Shield (BCBS). The prescription data is from IQVIA (formerly known as QuintilesIMS) and comes from their Xponent data base.

[6] In particular this implies that the solution maximizes average payoff in the long run, $lim_{T\to\infty} U(T)/T$, where $U(T)$ is the total, undiscounted return until time T.

One advantage of these data is that we can examine emergency room (ER) visits and hospitalizations as outcomes, rather than relying on the suicides as the outcome measure. Suicides are thankfully comparatively rare, making it difficult to compare doctors in terms of this patient outcome.

We show that doctors differ greatly in their propensity to experiment with different anti-depressant drugs. We proxy skill using the specialty of the prescribing physician (since in the U.S. most anti-depressants are prescribed by specialties other than psychiatry) and find that as predicted, higher skilled providers are more experimental. We are able to follow a patient through a treatment episode, and find that seeing a more experimental provider improves a patient's outcomes measured using total costs, non-drug costs, emergency room visits and hospitalizations. However, our results also suggest limits on the value of experimentation in that patients whose doctors violate prescribing guidelines have worse outcomes. This implies that the optimal management of provider care should consist of a combination of constraints upon some choices, while allowing or even promoting discretion over a set of treatments whose potential benefits are patient specific. Moreover, choice constraints should be relaxed for specialists relative to general practitioners.

The rest of this paper is laid out as follows: Section 2 provides some necessary background information. Section 3 lays out our model and theoretical results. The data is introduced in Section 4, and the empirical results appear in Section 5. Section 6 concludes.

2. Background: Previous Work on Physician Decision Making and Practice Style

We focus on physician diagnostic skill, tastes, and beliefs and how they impact physician decision making. Most previous work on prescribing focuses on other aspects of the problem. Dickstein (2015) asks how drug choices are affected by differences in copayments across insurance plans. Over his time period (2003-2005), branded drugs were a larger share of the anti-depressant market and the insurer's price paid varied from \$8.00-\$110.00 per month. However, copays only varied from ~\$10.00 to ~\$20.00 per month, and drugs with a wide range of prices had similar copays, suggesting that price differences could only be part of the explanation for why one drug was selected over another. His data is not well suited to examine differences in physician practice style, since even if physicians were identified, there would only be a few patients observed per physician.

Differences in patient health or patient tastes are an obvious potential demand-side driver of differences in medical decision making. Crawford and Shum (2005) and Dickstein (2015) study prescribing decisions in models in which there are no differences in physician practice style. As Crawford and Shum (2005) (page 1147) state: "... all doctors in our model have the same probability of prescribing a given drug to a patient with a given diagnosis in a given time period" implying that all differences in treatment must be driven by patient needs or preferences.

However, there is increasing evidence that supply-side variation is important. In an innovative study using vignettes from patient and physician surveys, Cutler et al. (2013) assess the hypothesis that regional variations in procedure use are driven by differences in patient demand across areas. Like Fuchs (2004), they conclude that patient demand is a relatively unimportant determinant of regional variations and that the main driver is physician beliefs about appropriate treatment that

are often unsupported by clinical evidence. Similarly, previous studies have found little evidence that patient demand is driving the large differences in C-section rates across U.S. providers (McCourt et al. (2007)). Finkelstein et al. (2015) address the same question using longitudinal Medicare claims data that allow them to track patients as they move through different healthcare markets. They suggest that about half of the observed variation in procedure use among these elderly movers is due to supply-side factors, while half is due to patient-level, or demand-side factors.[7]

Fear of litigation is another frequently cited reason for variations in physician decision making. The idea that physician decisions are shaped primarily by fear of litigation is popular, but has been repeatedly de-bunked. If fear of litigation were a primary driver, then one might expect legal reforms that limit liability to have major effects on practice. However, Baicker and Chandra (2005) find no evidence that treatment responds to changes in such liability, except for some screening procedures.

Currie and MacLeod (2017) and Currie et al. (2016) develop a framework for studying individual doctor decision making in contexts with zero/one choices. Their framework allows for two dimensions of physician practice style: Relative to a reference physician, doctors can be more or less aggressive on average, and they can also be more or less responsive to a patient's medical needs. They show that there is a great deal of variation in both responsiveness and aggressiveness across doctors, and that these characteristics of doctors are fairly stable over time. The model developed below builds on these ideas in a context with multiple possible treatments, and highlights the taste for experimentation and diagnostic skill as key dimensions in which doctors differ.

Previous research on prescribing practices (e.g. Berndt et al. (2015); Frank and Zeckhauser (2007)) focuses on concentration in prescribing as a measure of practice style and suggests that most physicians have a favorite drug or a small number of favorites which they prescribe to most patients with a given condition. Differences in prescribing practices have led some observers to call for stronger practice guidelines, especially in psychiatry. Meehl (1954), Grove et al. (2000) and Kahneman and Klein (2009) argue that in general an algorithm could do as least as well as a psychiatrist in the treatment of mental illness. [8] On the other hand, Frank and Zeckhauser (2007) argue that "excess clustering of physician behavior may be reinforced by the proliferation of practice guidelines accompanied by increased public reporting on performance relative to recommended care patterns." That is, they express concern that guidelines could prevent doctors from providing care

[7]An alternative explanation for sub-optimal treatment decisions is spillovers between doctors. Chandra and Staiger (2007) study the choice of surgery versus medical management of cardiac patients. Physicians in areas that specialize in surgery are assumed to become better at surgery and worse at medical management and vice-versa. These spillover models have two important empirical implications which are that one should see more uniformity within areas than across areas, and that doctors should converge to a regional practice style over time. However, neither Epstein and Nicholson (2009) nor Dranove et al. (2011) find convergence in practice styles among physicians in the same hospitals over time. Similarly, Chan (2015) finds that the practice styles of attending physicians have little impact on those junior to them in the same hospital. On the other hand, Molitor (2016) studies cardiologists who move and finds rapid convergence to the practice style of the destination area within one year. It is conceivable that physicians choose to move to places with a more congenial practice style rather than the practice style in the destination place having an immediate effect on their own style. As Cuddy et al. (2018) show, we find little evidence of area-level spillovers in treatments for depression.

[8]There is ongoing research looking the question of the optimal sequence of treatments following an algorithm (see Adli et al. (2017)).

that is sufficiently individualized. More generally, management practices have been shown to be strongly related to outcomes in health care settings, and the question of how best to structure guidelines for physicians is an important one (see Bloom et al. (2015), Tsai et al. (2015) and McConnell et al. (2013)). We measure concentration using an entropy score, and show that it increases with the doctor's taste for experimentation, and falls when doctors follow guidelines. An important contribution is to relate concentration in prescribing and adherence to guidelines to patient outcomes in a setting that allows us to control for differences between patients.

3. Conceptual Framework for Understanding and Measuring Variation in Physician Practice Style

Much previous work, as discussed in Frank and Zeckhauser (2007), views physicians as having practice styles that range from norm-based to hyper-rational Bayesian. Our goal is to provide a single theoretical framework that can capture this variation, and link variation in physician behavior to patient outcomes. This section is divided into three parts. In the first, we discuss the experimentation-exploitation trade-off faced by physicians. After working through a simple two-drug, two-period problem that has a closed form solution and offers some useful insights, we turn to the multi-drug, multi-period problem and discuss the properties of the upper confidence bound algorithm. We show that the UCB algorithm allows one to parameterized the experimentation-exploitation trade-off with a single parameter and we provide conditions under which the optimal decision is achieved. Reverdy et al. (2014) use the UCB algorithm to model the behavior of human subjects playing a multi-armed bandit problem with correlated arms and find that a great deal of heterogeneity in individual behavior can be captured by variations in the tuning parameter.

The second section introduces the (normalized) entropy score. In each period we can measure the fraction of patients for whom a physician prescribes a particular molecule. This yields a probability vector in the 32 dimensional simplex, $\Delta^{32} = \left\{\mathbf{p} | \sum_{i=1}^{32} p_i = 1\right\}$. Following Theil (1980), we use the Shannon entropy score to measure the diversity in decision making.

Finally, in the third part we connect the learning model with the entropy score. We use data from medical trials to provide a rough parameterization of beliefs about the drug effects. We then illustrate how variation in the taste for experimentation and in physician diagnostic skill are reflected in physician practice style and how they affect patient payoffs over time.

3.1. **Learning.**

The Experimentation-Exploitation Trade-off. Physicians know that every patient is different, and that what works for one may not work for another. At the same time, physicians have data from clinical drug trials that can be used to determine which drug has the highest expected payoff for an average patient. Hence, physicians face a trade off between picking the drug that is best for the average patient (unless it has been proven not to work for that patient), and experimenting with different drugs in an attempt to find the one that works best for a particular patient.

This trade-off can be illustrated concretely with a simple two-drug, two-time period example. Suppose a physician is choosing between two drugs, A and B with independent effects over two periods, $t \in \{0, 1\}$.

In period 0 the physician observes a measure of the patient's condition, denoted by y_0, the log of a numerical measure.[9] This measure is normalized so that $y_0 < 0$ indicates that a person is depressed and would benefit from treatment. The doctor chooses a drug $d_0 \in \{A, B\}$, that has a true *unobserved* effect, $e_{d_0} \in \Re$.

Period 1 : The patient returns to the physician after taking the drug, and the physician observes the patient's condition:

$$y_1 = e_{d_0} + \epsilon_1,$$

where $\epsilon_1 \sim N\left(0, \sigma^2 = \frac{1}{\rho}\right)$ is a noise term that represents the ability of the physician to correctly diagnose the patient. In other words, the physician observes the true condition plus some noise. We can think of the precision ρ as a physician-specific variable that is larger for better physicians.

The physician has some prior beliefs about drug effectiveness. Suppose the physician believes that the effectiveness of each drug is normally distributed and uncorrelated with the effectiveness of the other drug:

$$e_d \sim N\left(\mu_d, \sigma_d^2 = \frac{1}{\rho_d}\right),$$

where ρ_d is the *precision* of the physician's beliefs about the effect of drug d on a person randomly draw from the potential patient pool. Hence, ρ_d reflects the physician's uncertainty about the likely effectiveness of the drug for a particular individual. The physician can use the signal y_1 to update her beliefs. The assumption that drug effects are believed to be uncorrelated implies that the physician can update her beliefs for drug d_0, but not for the other drug, $d \neq d_0$. Suppose that $d_0 = A$. Then, after observing y_1, Bayesian updating implies:

$$\mu_A(y_1) \equiv E\{e_A|y_1\} = \frac{\rho y_1 + \rho_A \mu_A}{\rho + \rho_A}, \tag{1}$$

$$\rho_A(y_1) = \frac{1}{var\{e_A|y_1\}} = \frac{1}{\rho + \rho_A}. \tag{2}$$

If drug B is prescribed in period 0, then we have the analogous expressions for the updated beliefs for drug B, $\mu_B(y_1)$ and $\rho_B(y_1)$.

Given the data $\{d_0, y_1\}$, in period 1, the physician makes a choice $d_1 \in \{A, B\}$, that in turn results in an outcome:

$$y_2 = e_{d_1} + \epsilon_2,$$

where it is again assumed that $\epsilon_2 \sim N\left(0, \sigma^2 = \frac{1}{\rho}\right)$. We can denote the physician's choice function in period 1 by $\delta_1(y_1, d_0) \in \{A, B\}$, and hence a physician's overall decision rule in period 0 is defined by:

$$\delta_0 = \{d_0, \delta_1(y_1, d_0)\}.$$

In period 0 a physician might place different weights on each period of treatment. Suppose that physician preferences are given by:

$$U(\delta, \zeta) = E\{(1-\zeta) y_1 + \zeta y_2|\delta\},$$

[9] In the simulations below we measure the patient's condition using h, the Hamilton 17 score, an outcome measure that is commonly used in clinical trials of anti-depressants. A score of $h > 7$ means that the person is depressed, so we define outcomes to be $y = log(7) - log(h)$ so that persons with $y > 0$ are not depressed. See Hieronymus et al. (2015) for a recent discussion of using clinical trials to measure the effectiveness of anti-depressant drugs.

where $\zeta \geq 0$ is the weight given to the future well being of the patient. This formulation of the utility function implies that the physician is altruistic and only cares about the well being of the patient. However, physicians can differ in terms of the weights they place on the present vs. the future outcomes of the patient. If $\zeta = 0$, then in period 0 the physician will chooses the drug that maximizes $E\{y_1|d_0\} = \mu_{d_0}$. Let us label the drugs so that $\mu_A > \mu_B$, and hence when $\zeta = 0$ the physician always chooses A.

When $\zeta = 1$, the choice in period 0 depends on the *value of information* from that choice. If drug A is chosen in period 0, then the physician updates her beliefs regarding the effectiveness of drug A. If $E\left\{e^A|y_1, d_0 = A\right\} < \mu_B$ then she will switch to drug B in period 1. However, if the variance of the effectiveness of drug B is higher than the variance in the effectiveness of drug A, then it is possible that choosing drug B in period 0 will have a larger impact on the expected effect of treatment in period 1.

The value of information can be computed as follows.[10] First, compute the value of choosing drug A in period 0 and then using this information to decide on treatment in period 1. The value of information is computed relative to no information. When the physician has no information (NI) from period 0, then her utility in period 1 is:

$$U_1^{NI} = max\left\{E\left\{y_2|d_1 = A\right\}, E\left\{y_2|d_1 = B\right\}\right\} = max\left\{\mu_A, \mu_B\right\} = \mu_A.$$

If she chooses to treat with drug A in period 0, she then observes the outcome $y_1 = e_A + \epsilon_1$. Then, using (1) the physician chooses:

$$\delta_1(y_1, A) = argmax_{A,B}\left\{\mu_A(y_1), \mu_B\right\}.$$

The computation can be completed by taking the expected value of the outcome in period 0. Notice that:

$$y_1 = \mu_A + \xi_{A1} + \epsilon_1, \tag{3}$$

where ξ_{A1} is a mean zero, precision ρ_A, normal random variable representing the physician's uncertainty about the effect of drug A in period 0. The payoff from choosing treatment $\delta_0^A = \{A, \delta_1(y_1, A)\}$ is:

$$U\left(\delta_0^A, \zeta = 1\right) = E\left\{max\left\{\mu_A(y_1), \mu_B\right\}\right\}.$$

The value of information from choosing A in period 0 is given by:

$$V_A = U\left(\delta_0^A, \zeta = 1\right) - \mu_A \geq 0.$$

Before observing y_1, the expected value of y_2 given $d_1 = A$ is μ_A, and hence the ability to observe y_1 after choosing $d_0 = A$ provides the option to choose $d_1 = B$ if one learns that it is likely to be better. Hence $V_A \geq \mu_A$. Similarly, the value of information from choosing strategy $\delta_0^B = \{B, \delta_1(y_1, A)\}$ is given by:

$$V_B = U\left(\delta_0^B, \zeta = 1\right) - \mu_A \geq 0.$$

[10] See Raiffa and Schlaifer (2000) for a discussion of the value of information, particularly section 5.10. They do not have the explicit formula provided here - the derivation is in the appendix.

When $d_0 = B$, the physician can always get μ_A from choosing A in period two, hence learning about B can only increase her future payoffs.

For this two-period case one can explicitly compute the value of information (proofs are in the appendix):

Proposition 1. *The value of information from choosing drug $d \in \{A, B\}$ is:*

$$V_d = \sigma_d L\left(\frac{\mu_A - \mu_B}{\sigma_d}\right),$$

where $\sigma_d^2 = \frac{\rho}{\rho_d(\rho_d+\rho)}$, and, for $x \geq 0$,$L(x)$ is the unit-normal linear loss function defined by:

$$L(x) = E\{max\{x, \gamma\}\} - x = (1 - F(x))(\phi(x) - x),$$

$\gamma \sim N(0,1)$, $F(x)$ is the cumulative distribution function for the Normal distribution, and $\phi(x) = E\{\gamma|\gamma \geq x\}$ is the expected value of a lower truncated Normal distribution. For $x < 0$ we have $L(x) = L(-x) = E\{max\{x, \gamma\}\}$.

This result provides an exact computation of the value of information in this two drug example. The function $L(x)$ has a maximum value at $L(0)$ which decreases to zero for large positive values of x. Thus we have the following corollary:

Corollary 1. *Experimenting with drug B has more value than giving drug A if and only if the precision of the prior for drug B, ρ_B, is lower than the precision associated with drug A, ρ_A. The value of experimentation also falls with the difference in the expected effects of the two drugs ($|\mu_A - \mu_B|$).*

The uncertainty term, $\sigma_d^2 = \frac{\rho}{\rho_d(\rho_d+\rho)}$, depends upon both the precision of the signal y_1 and the prior precision of the physician's beliefs, ρ_d. If $\rho_A > \rho_B$ then a physician who cares only about long term outcomes will choose drug B in period 0 because the gain in information is greatest in that case, regardless of the value of the mean effect. Let us now consider decisions when the value of $\zeta \in [0, 1]$. The payoffs to the physician from choosing drugs A or B in period 0 are then given by:

$$U(A, \zeta) = (1 - \zeta)\mu_A + \zeta(\mu_A + V_A), \tag{4}$$

$$U(B, \zeta) = (1 - \zeta)\mu_B + \zeta(\mu_A + V_B). \tag{5}$$

Taking the difference between these two expressions shows that the physician chooses treatment A over B if and only if:

$$(1 - \zeta)(\mu_A - \mu_B) + \zeta(V_A - V_B) \geq 0, \tag{6}$$

$$\zeta^*(\mu_A, \rho_A, \mu_B, \rho_B, \rho) \equiv \frac{1}{1 + \left(\frac{V_B - V_A}{\mu_A - \mu_B}\right)} \geq \zeta. \tag{7}$$

Thus, physicians who put a weight $\zeta > \zeta^*(\mu_A, \rho_A, \mu_B, \rho_B, \rho)$ will choose treatment B in period 1. The function ζ^* is decreasing in $\left(\frac{V_B - V_A}{\mu_A - \mu_B}\right)$. Hence taking the distribution of ζ as given, physicians are more likely to experiment with B when there is more uncertainty about the likely effect of B, or when the expected value of B, μ_B, is closer to the expected value of A.

As discussed above, physician diagnostic skill is an important physician attribute which is here parameterized by the precision of diagnosis, ρ. As the physician becomes less skilled ($\rho \to 0$), then the value of information approaches zero for both drugs, and $\zeta^* \to 1$. In other words, we should expect to see less experimentation by lower ability physicians.

Choosing among many alternatives. The purpose of this section is to extend the analysis to the multi-drug case. Suppose there are $j \in J$ physicians and each physician has patients I_{jt} in period t. Patient $i \in I_{jt}$ is described by a state $\mathbf{x_{it}} \in X$, but the physician can only observe $\mathbf{z}_{ijt} \in Z$. The physician's task, given condition $\mathbf{z}_{ijt}$, is to recommend a course of action. An action involves choosing a drug $d \in D = \{d^0, d^1, ..., d^m\}$, where m is the number of drugs, d^0 means no drug, and $d^1, ..., d^m$ are one of the available drugs. Physician behavior can be represented as a mapping from the observed patient condition to a drug treatment:

$$\delta_{jt} : Z \to D.$$

Since physicians may learn from experience, we allow their behavior to change over time, though we assume that behavior is fixed within a time period. Even when the number of choices is small, physician behavior is potentially complex.

For example, one patient attribute that is typically contained in $\mathbf{z}_{ijt}$ is the drug that the patient was prescribed in the previous period (which might be "no drug"). With 32 drugs, then for each drug history the physicians would have 32 possible choices, and hence the number of possible physician behaviors is at least $32^{32}!$[11] Our goal is to not only provide a tractable model of this choice, but also to provide a low dimensional characterization of these decisions.

At date t, physician j observes y_{ijt}, the condition of patient $i \in I_{jt}$:

$$y_{ijt} = e_{id_{t-1}} + \epsilon_{ijt}, \tag{8}$$

where $e_{id}, d \in D$ is the effect of any drug that the patient is currently taking, and the error term is distributed $\epsilon_{ijt} \sim N\left(0, \sigma_j^2 = \frac{1}{\rho_j}\right)$. Larger values of ρ_j correspond to physicians who can more accurately assess patient condition. Different drugs are expected to have different effects on each patient, and the task of the physician is to choose the drug that is most helpful for each individual. A physician's behavior will depend on their assessment of the patient's condition and their beliefs about the effectiveness of medication. The distribution of beliefs about drug effects at time t is assumed to be given by:

$$\mathbf{e}_{ijt} \sim N\left(\mu_{ijt}, \mathbf{\Sigma}_{ijt}\right),$$

where physician beliefs are compactly represented by $\mathbf{B}_{ijt} = (\mu_{ijt}, \mathbf{\Sigma}_{ijt})$.[12] Before the patient arrives, a physician's training and experience endows them with some initial beliefs denoted by

[11] This is just a bit less than the estimated number of atoms in the planet earth - around 10^{49}.

[12] In particular, $\mu_{ijt0} = 0$, the expected condition of a patient who has not been treated, $\mathbf{\Sigma_{ijt}}[0, d] = \mathbf{\Sigma_{ijt}}[d, 0] = \sigma^2, d = d^0, \ldots, d^m$ is the variance of the underlying condition of the patient. For a matrix $\mathbf{A}$ the expression $\mathbf{A}[k, l]$ means the $k'th$ row and the $l'th$ column.

$\mathbf{B_{j0}}$. Note that this formulation allows for correlations in the expected effects between different drugs.[13]

The initial beliefs, $\mathbf{B}_{j0}$ play a key role because they determine the first drug that a physician tries. In principle, these beliefs should be guided by the medical evidence on drug effectiveness. There are many randomized control trials of anti-depressants (some results are discussed briefly in the appendix). In what follows we will assume that all physicians have initial beliefs based on the distribution of treatment effects from these clinical trials.

Treatment proceeds as follows. A patient specific effect $\mathbf{e}_i$ is drawn before the first visit to a physician and is given by:

$$\mathbf{e}_i \sim N\left(\mu_0, \Sigma_0\right),$$

where $\mathbf{B}_0 = (\mu_0, \mathbf{\Sigma}_0)$ represents the true distribution of patient effects in the population of individuals who seek treatment with a physician. For the time being, we assume that the physician is the sole decision maker, which may be realistic given limited patient information about the many drugs available. In the empirical work, we will allow for fixed differences between patients, which could include differences in patient tastes for medications. The physician decision maker believes that the true distribution is $\mathbf{B_{j0}}$, which may or may not be the same. The physician is assumed to observe the drug taken in period $t-1$, denoted d_{it-1}. Patients can change physicians, so the current physician may not have prescribed the current drug.

The state of a physician's information is completely defined by:

$$\mathbf{z}_{ijt} = \left\{\mathbf{B}_{ijt-1}, y_{ijt}, d_{t-1}\right\},$$

namely the current beliefs regarding the effect of treatment for patient i, $\mathbf{B}_{ijt-1}$, the observed condition of the patient, y_{ijt}, and the current drug regime d_{it-1}. The physician uses the data y_{ijt} to update her beliefs using Bayes rule. The formula is in the appendix. Here we write:

$$\mathbf{B}_{ijt} = \Pi\left(\mathbf{z}_{ijt}\right).$$

Once beliefs have been updated, then by definition:

$$E\left\{y_{ijt+1}|\mathbf{z}_{ijt}\right\} = \mu_{ijt}\left[d_{ijt}\right],$$

where $\mu_{ijt}\left[d_{ijt}\right]$ is the expected value of drug d following the update.

[13]Dickstein (2014) explicitly allows for such correlation in his model by supposing that physicians follow a two step procedure when selecting a drug - they first choose the drug class, and then choose a drug within the class. Since we allow for general correlation between all drugs, his model is a special case.

We assume that the preferences of the physicians are given by the difference between the expected consequence of treatment, and the best choice:[14]

$$
\begin{aligned}
0 \geq U_{ij}(T) &= E\left\{\sum_{t=1}^{T} y_{ijt}\right\} - e_i^* T, \\
&= E\left\{\sum_{t=1}^{T} e_{id_{t-1}} - e_i^* + \epsilon_{ijt}\right\}, \\
&= -E\left\{\sum_{d\in D} n_{id}\left(e_i^* - e_{id}\right)\right\}, \\
&= -E\left\{\sum_{d\in D\setminus d^*} n_{id}\left(e_i^* - e_{id}\right)\right\},
\end{aligned} \tag{9}
$$

where T is the period of treatment, and n_{id} is the number of times the physician chooses strategy d. The final expression $E\left\{\sum_{d\in D\setminus d^*} n_{id}\left(e_i^* - e_{id}\right)\right\}$ can be thought of as the negative of a physician's *regret* because it measures the distance between the optimal and the actual choice. In this setup the best the physician can achieve is normalized to zero. The number of times the physician chooses the *wrong* treatment is given by:

$$
0 \leq N_{ij}(T) = E\left\{\sum_{d\in D\setminus d^*} n_{id}\right\} \leq T. \tag{10}
$$

This expression provides a formal link between preferences and the number of times the physician chooses a sub-optimal treatment.

If the physician expects to see the patient only once, then she sets $T = 1$, and chooses the treatment that has the highest expected effect. For example, ER doctors may be concerned about finding the best immediate treatment. Other physicians may be involved with long-term treatment of the patient and hence they may be more concerned with the average well being of the patient over time. In that case, the expected duration of the doctor-patient relationship may be long enough to justify experimentation with different treatments. We explore the impact of increasing the time horizon on physician behavior and performance in the next section.

Bounds on the Optimal Treatment. This section provides some predictions about the way observed behavior changes with characteristics of physicians. We first find an upper bound on patient utility, and then a lower bound. Then, we use these two results to show that the UCB algorithm will eventually achieve the optimum.

Our first result follows immediately from an important paper, Lai and Robbins (1985) who provide a lower bound on the number of times the physician chooses treatment d under any strategy

[14] See Lai and Robbins (1985) who introduce the notion of minimizing regret as an objective. In economics, it is more natural to think in terms of utility, and hence our analysis is in terms of maximizing the negative of regret, as given by $U_{ij}(T)$.

that maximizes $U_j(T)$ in the long run:

$$N_{ij}(T) \geq \left\{ \frac{2}{(e_i^* - e_{id})\rho_j} + O(1) \right\} log(T), \tag{11}$$

where $O(1) \to 0$ as $T \to \infty$.[15] This expression shows that when there is uncertainty regarding the effect of treatment, a physician pursuing the optimal treatment strategy necessarily makes errors. The lower bound on the number of errors is smaller for physicians who have higher diagnostic skill (ρ_j is larger), and when the gap between the optimal drug and other drugs is larger.

The lower bound on the average number of errors approaches zero as the number of time periods increases:

$$lim_{T\to\infty} N_{ij}(T)/T \geq lim_{T\to\infty} \left\{ \frac{2}{(e_i^* - e_{id})^2 \rho_j} + O(1) \right\} log(T)/T = 0.$$

For $T < \infty$ this expression implies that in a diverse population of patients where physicians cannot perfectly observed patient condition, there is *necessarily* variation in choice. This result implies that there is an upper bound on physician payoffs that is increasing with physician skill:

Proposition 2. *For any period T of treatment, the expected payoff of a physician following an optimal treatment strategy satisfies:*

$$U_{ij}(T) \leq -log(T) \left\{ 2(m+1) \frac{1}{(e_i^* - e_{id})\rho_j} + O(1) \right\},$$

where $m+1$ is the number of possible treatments (including prescribing no drug).

We know from expressions 4 and 5 that the payoff from the choice of a drug has two components - the expected gain from treatment and the value of information. From Corollary 1 we know that the value of information increases with the variance of the estimated gains from treatment. The UCB algorithm builds on these two ideas and also assumes that the decision maker updates beliefs using Bayes rule. The physician has beliefs regarding patient i in period t that are given by $\mathbf{B}_{ijt} = \{\mu_{ijt}, \Sigma_{ijt}\}$. Given a parameter $\tau_t \geq 0$ we construct a score for choice d:

$$Q(d, \mathbf{B}_{ijt}, \tau_t) = \mu_{ijt}[d] + \tau_t \Sigma_{ijt}[d,d]^{1/2}. \tag{12}$$

Since the effect of treatment, e_{ijtd} is normally distributed, we can relate Q to the probability of a beneficial effect from treatment with drug d. Given physician beliefs $\mathbf{B}$, the physician has the following assessment regarding the likely effectiveness of treatment d:

$$Pr[e_d \geq Q(d, \mathbf{B}, \tau)] = 1 - F(\tau), \tag{13}$$

where F is the normal cumulative distribution function.

Thus for each treatment, the score Q defines the point at which there is the same odds, $1 - F(\tau)$ of the effect of treatment exceeding Q. For example, when $\tau = 0$, then $Q(d, \mathbf{B}_{ijt}, 0) = \mu_{ijt}[d]$, and the odds of treatment d being better than $Q(d, \mathbf{B}_{ijt}, 0)$ is 50%.

[15] The general condition is $E\{n_{id}^T\} \geq \left\{ \frac{1}{D(p_{id}||p_{id^*})} + O(1) \right\} log(T)$, where $D(p_{id}||p_{id^*})$ is the Kullback-Leibler divergence between the payoffs under d and d^* for patient i. This expression has a nice closed form for normally distributed errors. See Reverdy et al. (2014)

The UCB Algorithm Applied to Drug Choice. The goal of the UCB algorithm is to find the best long term choice as quickly as possible. The algorithm can be directly applied to the problem of drug choice as follows. Let B_{ij0} be the initial beliefs of physician j treating patient i at time $t = 0$ and let additional data about the patient be z_{ij0}. Decision making at time t proceeds as follows:

(1) Update beliefs using Bayes rule: $\mathbf{B}_{ijt} = \Pi(\mathbf{z}_{ijt})$.

(2) Given the tuning parameter $\tau_t \geq 0$, for each drug $d \in D$ compute the upper confidence limit :

$$Q(d, \mathbf{z}_{ijt}) = \mu_{ijt}[d] + \tau_{ijt}\mathbf{\Sigma}_{ijt}[d, d].$$

(3) Set treatment:

$$\delta^{UCB}(\mathbf{z}_{ijt}) = argmax_{d \in D} Q(d, \mathbf{z}_{ijt}).$$

Next, set the tuning parameter to increase over time:

$$\tau_t = F^{-1}\left(1 - \frac{1}{Kt}\right),$$

which ensures that:

$$Pr[e_d \geq Q(d, \mathbf{B}, \tau)] = \frac{1}{Kt}.$$

In other words as t grows, the probability that the effect is greater than the upper credible bound goes to zero. The physician can be said to have an uninformative prior if the variance of prior beliefs is infinite. Let $U^{UCB}(T)$ denote the payoff when the physician follows the UCB algorithm and has initial beliefs with precision zero (found by computing behavior with positive precision and then taking the limit to zero). In this case we have the following result (the proof is in the appendix):

Proposition 3. *Suppose that the effects of the drugs are uncorrelated and $K \geq \sqrt{2\pi e}$, then with an uninformative prior:*

$$0 \geq U_{ij}^{UCB}(T) \geq -\sum_{d \in D} \frac{\gamma_1(K, T)}{(e_i^* - e_{id})\rho_j} + (e_i^* - e_{id})\gamma_2(K, T),$$

where $lim_{T\to\infty} \frac{\gamma_1(K,T)}{log(T)} = 2$ and $lim_{T\to\infty} \frac{\gamma_2(K,T)}{log(T)} = \frac{2}{K}$.

This proposition provides a lower bound on the UCB algorithm. Notice that an increase in physician skill (ρ_j is bigger) results in a sharper lower bound. In other words we can expect higher skilled physicians to learn more quickly. Combining the upper and lower bounds:

Proposition 4. *Suppose that a physician has uninformative priors, then in the limit performance under the UCB algorithm achieves the highest possible payoff:*

$$lim_{T\to\infty} \frac{1}{T} U_{ij}^{UCB}(T) = 0.$$

This result follows immediately from Propositions (2) and (3) and implies that physicians who follow the UCB algorithm eventually find the optimal treatment. But in order to find it they *must* experiment, as shown by proposition (2).

Behavioral Interpretation of the UCB Algorithm and Experimental Physicians. The properties of the UCB algorithm are typically determined with specific assumptions on how to increase the tuning parameter over time. This in turn results in organized search over the arms of the multi-armed bandit in a way that leads to the long run optimal choice. As Reverdy et al. (2014) show, the UCB algorithm can also be used as a model of human behavior where one can suppose that different individuals have *fixed* tuning parameters, that in turn have well defined behavioral interpretations.

For example, a *short-run* physician can modeled by setting $\tau = 0$. In that case, she chooses the option that is best for the patient in the current period, though she continues to update her beliefs as a function of the information she receives. This may be a reasonable way to characterize physicians, such as obstetricians, who are treating patients with post partum depression that is expected to be temporary.

Conversely, a chronically ill patients has a longer horizon, and hence may benefit from exploration to find the optimal treatment. However, from Proposition (4) it must be the case that the posterior precision for each choice is unbounded in the long run, which in turn implies that experimentation never stops. In practice, patients have finite lives, and hence at some point experimentation should stop. This can be modeled by setting the tuning parameter to a fixed $\tau > 0$ that corresponds to an *experimental* physician whose preferences are precisely defined as follows:

Proposition 5. *Suppose physician j treats patients according to the UCB algorithm with a fixed tuning parameter τ. Then as $T \to \infty$, experimenting stops and choice settles upon some drug d_{ij}^* such that the probability that any other treatment $d \neq d_{ij}^*$ is better than d_{ij}^* is less that $1 - F(\tau)$:*

$$Pr\left[e_d \geq e_{d_{ji}^*} | B_{ij}^{\infty}\right] \leq 1 - F(\tau), \tag{14}$$

where B_{ij}^{∞} is the limit as $T \to \infty$ of B_{ijt}.

Proof. Given that there are a finite number of choices, with probability one decision making settles upon a single choice d_{ji}^*, which in turn implies that physicians learn the true value of treatment $e_{d_{ij}^*}$ as $T \to \infty$. Under the UCB algorithm, it must be the case that for $d \neq d_{ij}^*$ $e_{d_{ij}^*} \geq Q\left(d, B_{ij}^{\infty}, \tau\right)$. From (13) we have:

$$Pr\left[e_d \geq e_{d_{ji}^*} | B_{ij}^{\infty}\right] \leq Pr\left[e_d \geq Q\left(d, B_{ij}^{\infty}, \tau\right) | B_{ij}^{\infty}\right] = \left(1 - F(\tau)\right).$$

□

For example, consider a short-run optimizer, that is a physician who chooses the treatment that maximizes the expected payoff each period. When experimentation ceases, the choice d_{ij}^* will have the feature that for any other choice, d, it must be the case that the expected return $\mu_{ij}[d] \leq e_{d_{ij}^*}$, and hence the probability that better treatment exists is less than 50%.

An experimental physician is one who continues to experiment until the probability of success is less than $1 - F(\tau)$. This behavior is related to Simon (1955)'s notion of satisficing. Simon

observes that search costs typically lead decision makers to continue to search until they reach some aspiration level.[16]

The insight of the literature on the UCB algorithm is that systematic exploration aimed at discovering the options that have the highest probability of success results in behavior that is close to optimal in the long run. Proposition (3) shows that if τ increases at an appropriate rate over time, then eventually the optimal choice is made. Proposition (5) shows that when the tuning parameter is fixed at τ, the physician continues experimentation until the probability of success fall below $(1 - F(\tau))$. The UCB algorithm captures the idea that a physician would experiment with a new drug if and only if it has a chance of success that is at least as great as $1 - F(\tau)$.

This approach fits in well with the existing literature on physician behavior. For example, Frank and Zeckhauser (2007), discuss three physician types:

1. Physicians who follow therapeutic norms. These doctors select treatments based upon a category rather than customizing treatment for each individual. The broader the category, the more patient-specific information is being disregarded in the treatment choice, so that one can think of this process of categorization as reflecting physician diagnostic skill, captured by ρ_j in our model. The main empirical implication is that physicians who use "crude" or less informative categories learn at a slower rate, as shown in Proposition (3). Currie and MacLeod (2017) provide direct evidence that physicians vary in the extent to which they respond to patient observables, consistent with this hypothesis.

2. Physicians who follow "sensible-use" norms. Frank and Zeckhauser (2007) use the example of chronic versus acute conditions suggesting that physicians find it sensible to use different norms in these two cases. In our model, this distinction arises naturally because the time horizon affects the return to experimentation. With an acute condition, the physician chooses the drug with the highest expected value given current information, and hence is characterized by $\tau = 0$. With a chronic condition, there can be a benefit from experimenting to find the best treatment, and hence have a $\tau > 0$.

3. Physicians who "do it my way." Some physicians regularly prescribe therapy that is quite different from the choices made by other physicians. This observation is an immediately implication of Bayesian learning in a large population of physicians. Given that outcomes are stochastic, in a large population we can expect one or two physicians to have a number of positive experiences with any particular treatment. Bayesian updating then leads the physician to have strong beliefs about the efficacy of the treatment which can be slow to change. Alternatively, if beliefs are very strong (the variance of the estimated effect of the drug is believed to be very precise) then rational decision making can still lead to poor decisions for extended periods.[17]

3.2. **Characterizing physician practice style.** The model developed above emphasizes that physicians can vary in a number of ways that will affect their decision making including their

[16] As Simon (1955) shows, the aspiration level corresponds to the reservation wage in a labor market search. The point is that even when we cannot directly observe a person's future labor market prospects, we can characterize their behavior in terms of a reservation wage that determines the point at which further search stops.

[17] As Diaconis and Freedman (1986) show under appropriate conditions a Bayesian decision maker may hold incorrect beliefs even in the long run with unlimited amounts of information.

ability to diagnose the patient, differences in beliefs about the effectiveness of different drugs, the doctor's time horizon with the patient, and their preference to relieve suffering in the short term vs. experimenting to find an optimum treatment. We cannot directly observe these physician characteristics in observational data, so we seek to develop proxies for them.

As we will see below, most psychotropic drugs in the U.S. are prescribed by physicians other than psychiatrists. General practioners are the most common prescribers, but specialities including obstetrician-gynocologists, cardiologists, and rheumatologists also frequently prescribe antidepressants. Doctors who are not psychiatrists will usually have little training in the use of antidepressant medications, and see fewer depressed patients per capita. Thus, it seems reasonable to assume that they will be less skilled diagnosticians on average, and therefore we use specialty as a proxy for ρ_j.

As our model makes clear, the return to experimentation will depend on the value of the information to be gained. In a world where there was one drug that was clearly superior for all patients, there would be little need to experiment, and prescribing would be expected to be highly concentrated. In a world where each of the 32 drugs was best for some patient, there could be a high return to gathering more information through experimentation.

The most popular measure of information is the Shannon entropy score, a measure Theil (1980) advocated for in economics and which is also a natural measure of concentration in this context.[18] We begin with a general definition of the information content in a practice style, and then discuss two applications. Let $\mathbf{p} \in \Delta^n = \{\mathbf{p} \in [0,1]^n \,|\, \sum_i p_i = 1\}$ be a probability vector. Let $n(\mathbf{p})$ denote the number of entries in the vector. We define the scaled entropy score:

Definition 1. Given a vector $\mathbf{p} \in \Delta^n$, then the *scaled entropy score* is given by:

$$\begin{aligned}\Phi(\mathbf{p}) &= -\sum_{k=1}^{n} p_k log(p_k)/log(n) \\ &= \sum_{k \in D} p_k log(1/p_k)/log(n), \\ &= \phi(\mathbf{p})/log(n),\end{aligned}$$

where

$$\phi(\mathbf{p}) = \sum_{k=1}^{n} p_k log(1/p_k)$$

is the Shannon entropy index.

In our empirical work the number of entries in the vector is held fixed for each set of regressions, and hence we normalize by the number of outcomes in $\mathbf{p}$. This ensures that our score is always in $[0,1]$, with 0 corresponding to a single choice, while $\Phi(\mathbf{p}) = 1$ implies that $p_k = 1/n$, where k

[18]Concentration of prescribing is a frequently examined aspect of practice style. Stern and Trajtenberg (1998) look at antidepressants and calculate Herfindahl indices of the concentration of prescribing behavior. Frank and Zeckhauser (2007) report that of 1,372 primary care physicians surveyed in 2004, the most prescribed medication for each of 9 different conditions was responsible for about 60% of a physicians' prescriptions for that condition. In contrast, patient demographics had little explanatory power. Berndt et al. (2015) use data on prescriptions of anti-psychotics from IQVIA (formerly IMSQuintiles). They show that most physicians have a favorite drug and that on average 66% of their prescriptions are for this drug.

indexes drugs. Here k is taken to be the number of drugs that are ever available over the sample period, and p_k is the share of patients who are taking drug k at time t. The scaled measure has values between zero and 1, and a unique maximum with $p_k = 1/k$. There are k minima, each corresponding to $p_k = 1$. We divide the entropy score by $log(n(\mathbf{p}))$ so that the score always lies between 0 and 1. In every application the dimension, $n(\mathbf{p})$, is held fixed, and hence our score is always in terms of the fraction of the maximum value possible - this allows easier comparison across cases.

In order to use the entropy score to describe practice style we begin by dividing time into periods. In period t, let n_{jt} be the number of patients that a physician treats, and let n_{djt} be the number of prescriptions of drug $d \in D = \{d^1, \ldots, d^m\}$, where m is the number of drugs, in period t. Let

$$p_{djt} = \frac{n_{djt}}{n_{jt}},$$

be the fraction of patients of physician j who take drug d in period t. This m-dimensional vector $\mathbf{p}_{jt} \in [0,1]^m$ is a measure of the physician's *static practice style* at time t.[19] It can be summarized using the normalized entropy score:

$$\Phi_{jt} = \Phi(\mathbf{p}_{jt}).$$

In the case where there is a unique optimal drug, entropy should be low and decreasing over time as physicians learn about the best drug and gravitate to it. Conversely, if matching patients to drugs is important, then higher entropy should be associated with better outcomes. These relationships will be illustrated below with a simulation using current evidence from randomized control trials as the starting point for initial beliefs.

3.3. **Effect of Learning on Practice Style.** Although the BCBS data is very rich, as described further below, it will not allow us to fully explore the implications of our model for patient welfare. In particular, the observed treatment window is quite short and so we cannot take account of what may be lengthy patient histories with depression. Hence, we conduct a simulation exercise to illustrate the effects of physician skill, physician tastes for experimentation, and time horizons with a patient on physician entropy scores and patient utility. Another advantage of simulations is that we can fix the distribution of patients to be exactly the same for each physician and thus abstract for now from the issue of heterogeneity between patients. Even though patients are *ex ante* identical, the simulation results illustrate that whether a patient wants the physician to experiment will depend on the diagnostic skill of the physician, so there there cannot be a unique, optimal treatment protocol.

Our simulations consider a 2x3 experiment with 6 physician types. The physician is either a short-run optimizer or an experimental type with a fixed tuning parameter, τ (or success probability cutoff $1 = F(\tau)$). In addition, we consider 3 levels of diagnostic skill, ρ_j. The short run optimizers are assumed to always give the treatment with the highest expected value in the current period. These physicians still update and learn from experience, but see no value in experimentation, so

[19]Here we have left off the "no treatment" choice since in the prescriptions data we do not see appointments that do not involve a prescription.

that $\tau = 0$. The experimental types do value experimentation and for our simulations we set $\tau_{Exp} = 1(1 - (1/T_{max}K))$, where $K = \sqrt{2\pi e}$, the constant used in Proposition (3), and T_{max} is the maximum number of periods used in the simulation. In addition, we allow physicians to vary in terms of diagnostic skill. Let $\rho_j \in \{10.0, 1.0, 0.1\}$, measure the accuracy with which the doctor assesses the patient's condition. Here $\rho_j = 10$ denotes high skill (H), and $\rho_j = 0.1$ denotes low skill (L).

In our simulations we assume that physician priors are given by data about the mean and variance of the effects of drugs from clinical trials.[20] These data are described further in the Appendix and the assumptions about drug efficacy that we use to model physician beliefs are briefly summarized in Table 1. The table details the efficacy of each of the top 11 anti-depressant drugs (ranked in terms of market share in 2014) in terms of the effect on the improvement in the Hamilton17 (HAM-D) score.[21] Clinical trials for anti-depressants typically select a population of depressed patients and randomly assign them to treatment and control. The control group gets a placebo drug while the treatment group get the drug under investigation. The level of depression before and after the experiment is measured using the HAM-D score. It is worth highlighting that the placebo effect is quite large. In the case of Sertraline (generic for Zoloft), the most popular drug, the placebo effect is 80% of the total effect of treatment. The fact that the placebo effect is on average responsible for more than 50% or a drug's effect can help explain why it is hard to find the most effective treatment.

The simulation considers a physician who has a constant load of 300 patients, all drawn from the same k-dimensional normal distribution of "true" drug effects from the clinical trials. That is, each patient will have a different optimal drug, but at first all patients will appear identical to the physician. We simulate a doctor's entropy score over time as well as her utility, which depends on the number of deviations from optimal treatment. To summarize, beliefs are fixed by the clinical data, so that the simulations explore the effects of doctor diagnostic skill, the doctor's taste for experimentation, and the treatment time horizon (T).

Figure (1) illustrates the evolution of mean practice style over time as a function of physician characteristics. Consider first the high skilled physician. Notice that entropy for the short run decision maker is initially higher than for the experimental decision maker, but the experimental type settles on a higher entropy score. Initially, the experimental type will prescribe high variance drugs, and she will not switch until the physician has enough information to make the alternatives more attractive. In the long run the experimental type tries out more drugs, which implies higher entropy in observed practice style.

The other cases mirror this result. The difference is that physicians with lower skill take more time to learn. Thus, over the 3-year period in the simulation, the entropy score from the medium-skilled physician is always lower for the experimental type than for the short-run type. Eventually, the entropy of the experimental type will over take that of the short run physician. In the case

[20] Drugs prescribed in the US are typically evaluated using a randomized control trial. Meta-analysis of these studies appear every few years in the medical literature so that professionals have access to the current results. See for example Linde et al. (2015) and Cipriani et al. (2016).

[21] See Hamilton (1960).

of the low skilled physician, learning is even slower and entropy is virtually identical for both the short-run and experimental types over three years. What we are observing here is that the medium and low skilled experimental physicians stick with the high variance drugs until they are sure they are worse than the other alternatives. This example illustrates why some decision makers can be observed prescribing "non-traditional" treatments.

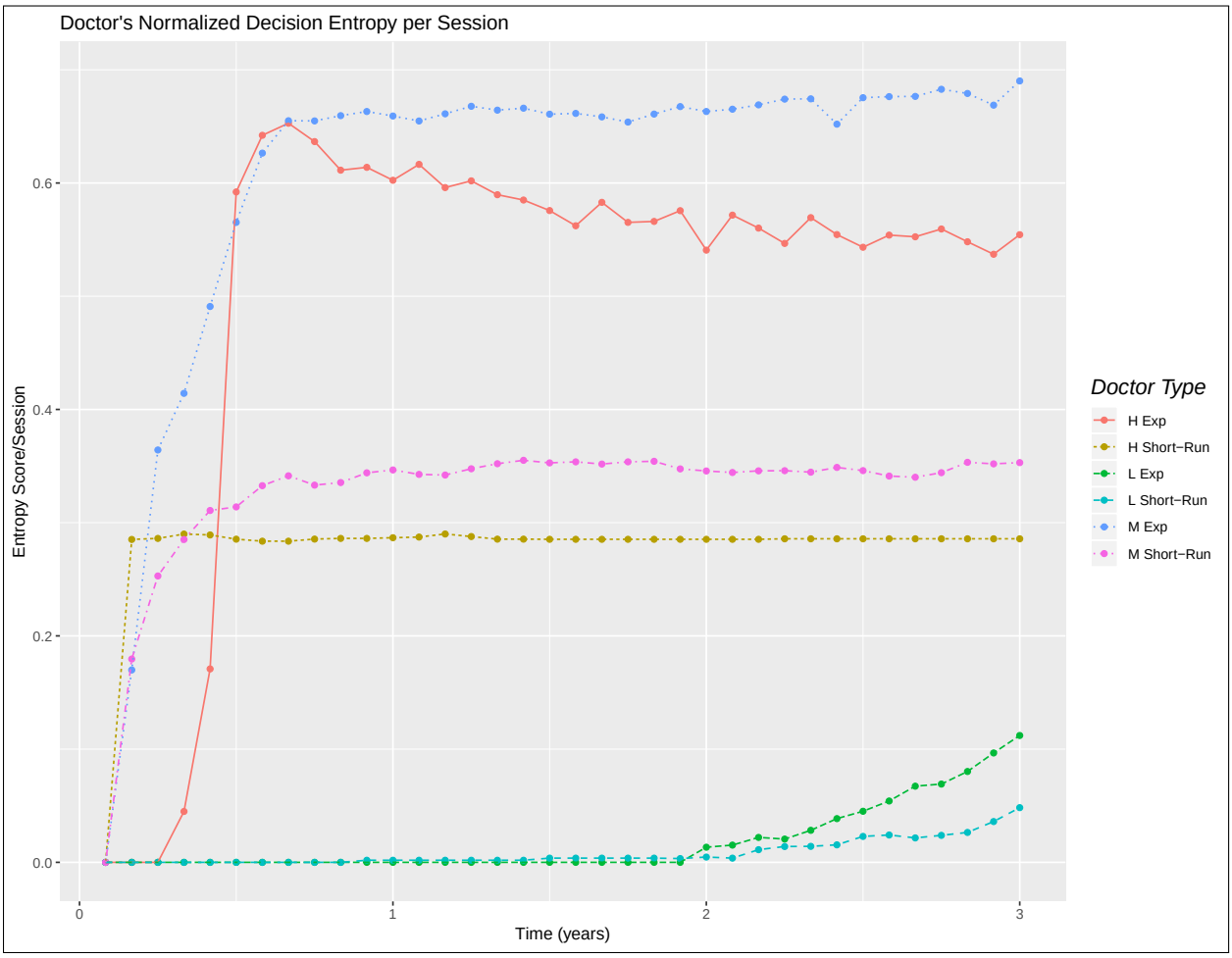

FIGURE 1. Effect of Physician Characteristics upon Entropy

In Figure (2) we plot the payoffs from treatment as a fraction of the maximum possible payoff. Again, starting with the high skilled physician, notice that patient well-being is initially lower than with the short-run physician, but eventually is very close to the maximum possible (1 in this diagram). Thus, the initial experimentation can hurt a patient, though in the long run the patient's outcome should be better. The case of the medium skill physician reflects what we observed with the entropy score - learning is slower, and in this simulation the outcome with experimentation is never better than with a short-run decision maker. Finally, with the low skilled physicians, learning is so slow that there is little difference in performance between short-run and experimental learning styles.

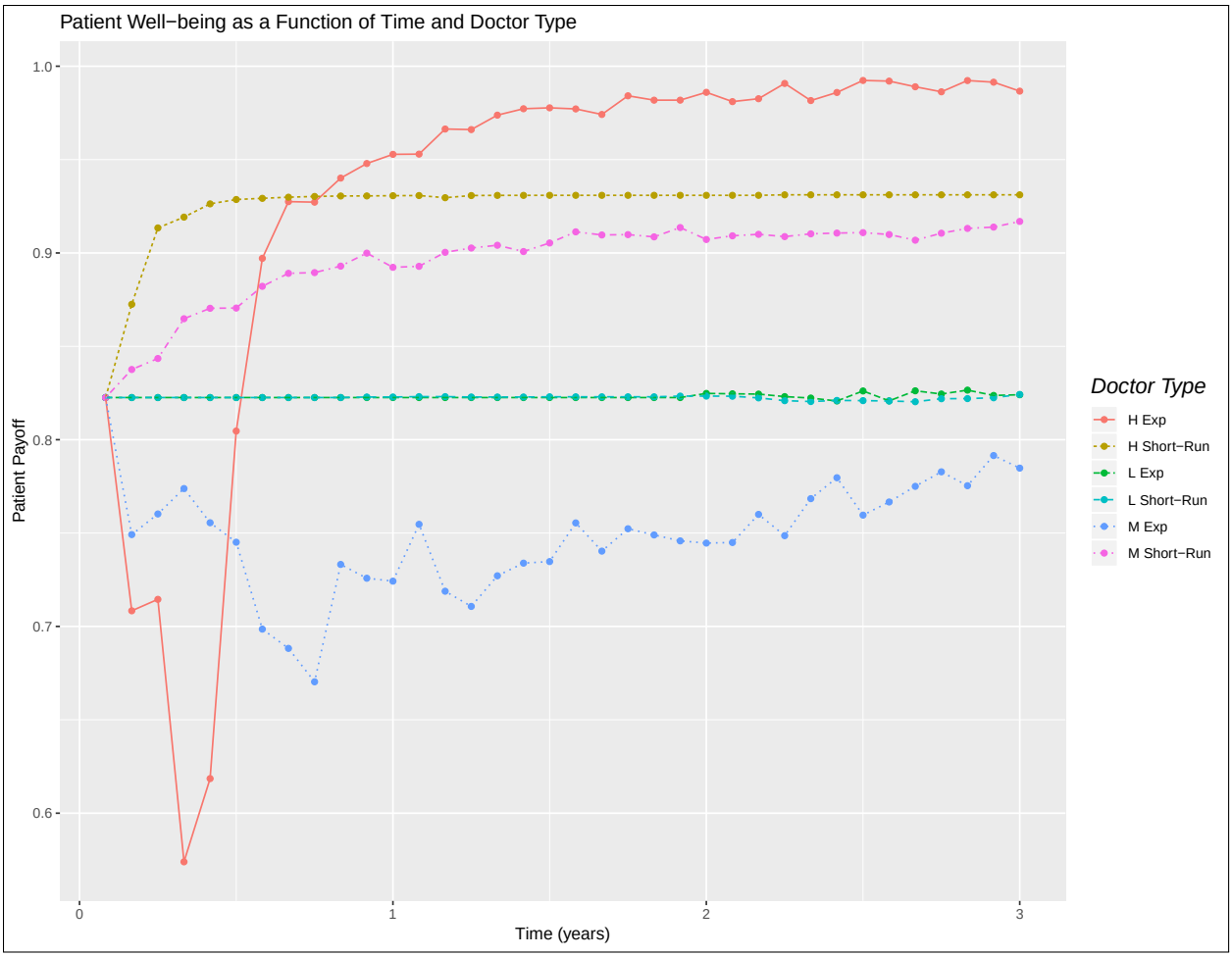

FIGURE 2. Effect of Physician Characteristics upon Patient Well Being

The point of these results is to illustrate that in the context of learning with noisy observations one cannot make simple predictions - there is a complex interaction between physician skill and preferences for learning through experimentation.

3.4. **Dynamic Practice Style.** Our framework can be naturally extended to deal with dynamic measures of practice style. One of the patient characteristics that physicians observe is the drug that patients are currently taking. For example, suppose that a physician begins with Sertraline and the patient has an adverse reaction. A natural question is what is the best next choice? For example, with antibiotics one typically starts with a common drug, and only progresses to more powerful drugs if there is failure to cure. This sequence of choices can be viewed as a Markov process - given the current drug, what is the probability of a particular drug being chosen subsequently? In our BCBS claims data we know whether a patient has been prescribed any anti-depressant in each month. Let $m = n+1$ be the number of drugs, where $d = 0$ corresponds to no drug. *Dynamic practice style* at date t is the vector $\mathbf{q}_{tj} \in \Delta^{m^2}$:

$$q_{tj}\left[k+m\times l\right]=fraction\ d^{k}\rightarrow d^{l},$$

the fraction of patients currently receiving drug k who are then prescribed drug l.[22] If $l=k$ then a decision has been made to continue with the same drug.[23]

Dynamic practice style is related to static practice style at date t as follows: Given $\mathbf{q}_{tj}$ the fraction of patients who are prescribed drug d by physician j in period t is given by:

$$p_{djt}=\sum_{k=0}^{m}q_{jt}\left[k+m\times d\right].$$

Define the vectors for each drug d:

$$\mathbf{m}_{djt}\left[i\right]=q_{jt}\left[i+m\times d\right]/p_{djt},$$

in which case we can write:

$$\mathbf{q}_{jt}=\left[p_{0jt}\mathbf{m}_{0jt},...,p_{njt}\mathbf{m}_{njt}\right].$$

The properties of the entropy function imply that:

$$\phi\left(\mathbf{q}_{jt}\right)=\phi\left(\mathbf{p}_{jt}\right)+\sum_{d=0}^{n}p_{djt}\phi\left(\mathbf{m}_{djt}\right).$$

In terms of normalized entropy we have:

$$\Phi\left(\mathbf{q}_{jt}\right)=\frac{1}{2}\left\{\Phi\left(\mathbf{p}_{jt}\right)+\sum_{d=0}^{n}p_{djt}\Phi\left(\mathbf{m}_{djt}\right)\right\},$$

and hence the normalized entropy of dynamic practice style is the average of the normalized entropy for static practice style and the weighted average of the normalized entropy of transitions to each drug observed.

From the concavity of the entropy function we have:

$$\Phi\left(\mathbf{p}_{j(t-1)}\right)=\Phi\left(\sum_{d=0}^{n}p_{djt}\mathbf{m}_{djt}\right)\geq\sum_{d=0}^{n}p_{djt}\Phi\left(\mathbf{m}_{djt}\right),$$

which implies:

$$\Phi\left(\mathbf{p}_{jt}\right)+\Phi\left(\mathbf{p}_{j(t-1)}\right)\geq2\Phi\left(\mathbf{q}_{jt}\right)\geq\Phi\left(\mathbf{p}_{jt}\right).$$

The vector $\mathbf{m}_{tjd}$ represents the pathways to drug d. For example, suppose that for each patient, the physician has a personal belief about the best drug, and never changes their choice, then $\Phi\left(\mathbf{m}_{djt}\right)=0$ and thus $2\Phi\left(\mathbf{q}_{jt}\right)=\Phi\left(\mathbf{p}_{jt}\right)$, the lower bound. On the other hand, suppose that the physician randomizes over her choices using the distribution $\mathbf{p}_{jt}$, then $\mathbf{m}_{djt}=\mathbf{p}_{j(t-1)}$, and we have $2\Phi\left(\mathbf{q}_{jt}\right)=\Phi\left(\mathbf{p}_{jt}\right)+\Phi\left(\mathbf{p}_{j(t-1)}\right)$. Thus, depending on the correlations between choices, the normalized entropy score for dynamic practice style can equal the upper or the lower bound.

[22] Alternatively, we could think in terms of an m*m transition matrix between drugs at *t-1* and drugs at *t*. Stacking the columns of the transition matrix would yield this vector.

[23] It is important to allow patients to remain with a previous medication, given that switching costs may be non-trivial. Previous work by Coscelli (2000) looking at ulcer medications shows that patients are even reluctant to change from branded drugs to equivalent generics.

We use dynamic practice style is to think about the extent to which physicians follow accepted practice as recommended by a national standards body (we use standards developed for the US, UK, and Canada). Standards are seldom specific enough to tightly specify exactly what a physician should do. However, many make suggestions about drug transitions, that is what to prescribe if the current drug does not seem to be working. We can map transitions that violate these guidelines into a vector:

$$\mathbf{P}_s \in \{0,1\}^{m^2},$$

were $\mathbf{P}_s[k + m \times l] = 1$ implies that under standards $s \in S = \{Canada, UK, US\}$, if a patient is taking drug d_k then they should not next be prescribed drug d_l. For each physician we can then examine the effect of transitions that violate standards s. One nice feature of our analysis is that we can compare the relationships between violations of different standards and outcomes using the same sample of doctors. Ultimately, examining the extent to which violations of guidelines affect patient outcomes will shed some light on utility of the guidelines themselves.

3.5. **Summary.** We have outlined a framework in which physicians must trade-off a higher probability of current treatment success against long-term success that can only be achieved via experimentation. We show that even if physician preferences are fixed, the benefits of experimentation vary with physician quality as measured by diagnostic skill, as well as with physician beliefs and with treatment windows. We have also suggested entropy as a proxy for a physician's propensity to experiment. In our data we can make use of the division between specialist and non-specialist prescribers as a proxy for physician diagnostic skill. We have also discussed the use of prescribing guidelines. In what follows, we examine the relationship between entropy, following guidelines, and patient outcomes.

4. Data

In order to examine the relationship between doctor practice style and patient outcomes, we access a new national sample of claims data from Blue Cross Blue Shield Alliance for Health Research (BCBS), a collaborative effort involving most of the regional BCBS plans. Specifically, we first selected a 10% sample of all of the member numbers (for members aged between 18 and 64 as of January 2013) in the system between January 2013 to September 2016. BCBS had about 99 million members aged between 18 and 64 who had any claims over our sample period. Of the 9.9 million members we selected, about 4.5 million show up in the pharmacy claims, and of these, 723,818 members were ever prescribed antidepressants over the sample period. These members constitute our core BCBS sample. For each of these members, we generate a panel of data with a record for each month and year that they appeared. In each time period we know whether they are taking any anti-depressant drug, what drug it was, who prescribed it, claims for drugs, outpatient visits, emergency room (ER) visits, and inpatient visits, and total health care costs generated by summing all claims across inpatient, outpatient, and pharmacy data bases.

One contribution of our study is to focus on ER visits as an indicator of adverse patient outcomes. In the US, patients with mental health crises are uniformly advised to proceed to the nearest ER for assessment even if the nearest institution does not have a psychiatric facility. The patient will

then be assessed and could be transferred elsewhere for an inpatient hospital stay if necessary. ER visits (and subsequent hospitalizations) for mental health indications are much more numerous than suicides and have a substantial impact on health care costs.[24] Our own calculations using hospital records data from the Health Care Utilization Project (HCUP) suggest that in 2014 there were 379.9 ER visits annually per 1,000 individuals.[25] Of these, 50.9 listed a mental health diagnosis on the hospital record; of these, 19.2 listed mood disorders as an indication and 18.6 listed anxiety. Anti-depressants are frequently prescribed for both of these indications.

A difficulty with using the BCBS claims data to examine practice style is that most doctors are likely to see patients with many types of insurance. Hence, it may not be meaningful to construct measures of practice style using only the BCBS data as we could be omitting many of a doctor's other patients. We remedy this problem by computing the entropy measure of physician practice style using a second data base from IQVIA, and then matching in this data using the doctor's name and state of practice.[26] We are able to find a match in the IQVIA data for 74.0 percent of the doctors in our BCBS sample. We also match in physician characteristics such as specialty from the National Plan and Provider Enumeration System (NPPES).

We follow Berndt et al. (2015) and limit our analysis to physicians who wrote 12 or more prescriptions for anti-depressants in the IQVIA data in at least one year of the sample, and who were not missing physician characteristics.[27] We allow the doctor's entropy score to vary over time, calculating a separate entropy score for each doctor for each year of data. We match 2013 BCBS data to 2012 doctor entropy data, and so on.[28] Since there are 32 different anti-depressant molecules in use over our sample period, the entropy score is computed using m=32. However, the top eleven molecules accounted for 94.78% of prescriptions in 2014, and when we focus on transitions from one drug to another we use these eleven molecules plus "all others" and drug combinations (which we dub "cocktails.")

[24]A few studies focus on anti-depressant prescription and suicide using country-level data. Ludwig et al. (2009) study the relationship between the use of Selective Serotonin Reuptake Inhibitors (SSRIs) and suicide, relying on institutional differences in, for example, when SSRIs were approved to explain variation in their use over time within countries. They find that an increase in prescribing of SSRIs of one pill per capita reduced suicide by five percent. Berndt et al. (2015) study a reduction in the use of SSRIs for youths after labels warning that SSRIs could increase suicide risk in young people were mandated and show that suicides went up following the mandate. Suicide is a sufficiently rare outcome that within-country studies tend to have low power. Yet cross-country comparisons may be contaminated by other factors that are changing differentially across countries.

[25]HCUP is a collaboration between state governments and the federal governments in which states allow their administrative hospital data to be made available through a central registry. However, only some states include ER data, and only a subset of those states identify a patient's county of residence. Our analysis of ER visits focused on six states that meet these criteria: Arizona, Florida, Kentucky, Maryland, New Jersey, and New York.

[26]IQVIA (formerly known as IMSQuintiles) is a public company specializing in pharmaceutical market intelligence. As of 2014, IQVIA directly surveyed 86% of retail pharmacies, with the remaining prescriptions imputed to add to industry totals using a patented projection method. The data includes information about each provider from the American Medical Association, including specialty.

[27]In principal, one can only get a good measure of practice style if one sees enough patients treated by a particular doctor. In addition to dropping doctors with fewer than 12 prescriptions in any year, we explored samples in which we kept only doctors with numbers of patients above the median in the data, and found that mean entropy scores (computed over patients in the sample) were quite similar. This follows from the fact that most patients are seen by doctors with many prescriptions.

[28]Due to data limitations, both 2015 and 2016 BCBS data are merged to entropy measures for 2014.

Figure (3a) uses the IQVIA data to illustrate the variation in favorite drugs prescribed across the U.S. In terms of our model, if all physicians were focused on getting the best short-run results for their patients, and there was one best drug for most patients, then they would all start with the same drug, the one shown to be "best" in clinical trials. However, one can clearly see differences in the favorite anti-depressant molecule prescribed, with Sertraline, Citalopram, and Fluoxetine all dominating in some areas.[29] Figure (3b) shows variation in entropy scores across the country. Scores range from about 0.5 to 0.7, indicating significant differences in the range of drugs prescribed across the country. These patterns are suggestive of important differences in practice style, though it is possible that patient needs could vary systematically across areas.

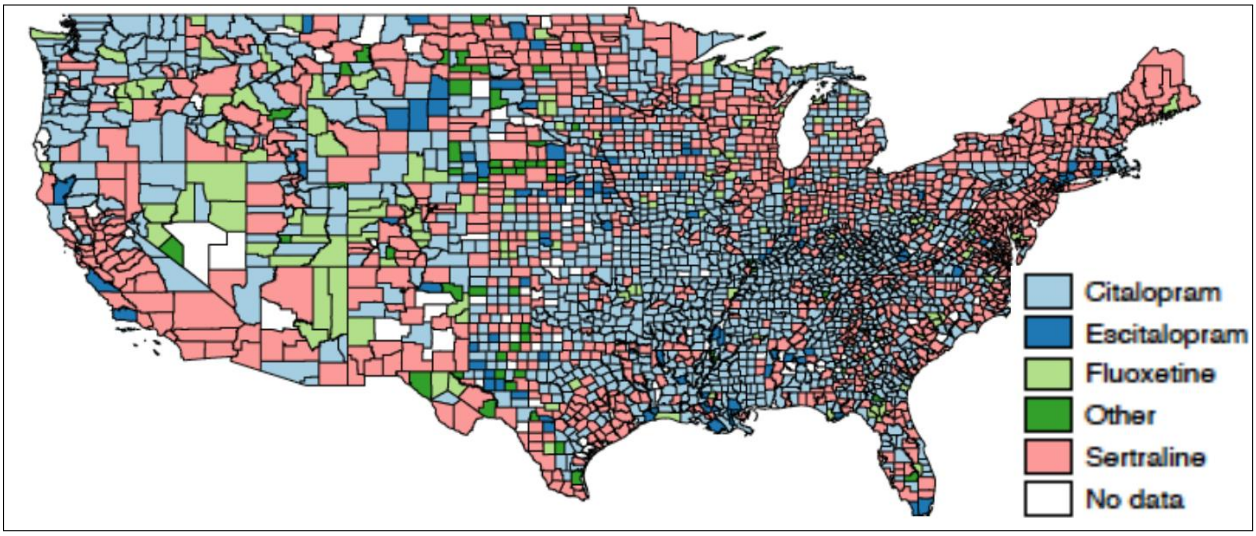

FIGURE 3A. Most popular anti-depressants, by county, by active ingredient, 2014

Another possibility is that provider expertise could vary systematically across the country, for example if some areas have more psychiatrists per capita. Table (2) provides further detail about the breakdown of anti-depressant prescriptions and entropy scores across types of providers in the IQVIA data. Table (2) shows that the average provider who wrote more than 12 prescriptions for any anti-depressant in any year of the IQVIA data, wrote 302 such prescriptions annually. Not surprisingly, psychiatrists wrote 1033 prescriptions annually compared to 451 for general practitioners. What may be more surprising is that many other MDs write anti-depressant prescriptions, with an annual mean in this group of 74 prescriptions per provider. Another surprising result is that most anti-depressant prescriptions are not written by psychiatrists, but by GPs and MDs who are likely to have very little specific training in the use of anti-depressants. The model discussed above predicts that providers with less expertise should do less experimentation. Table (2) shows that in keeping with this prediction, psychiatrists have the highest entropy scores, followed by GPs, with other MDs having substantially lower entropy scores. Entropy scores by physician cohort also follow this broad pattern, with "prime age" practitioners tending to have higher entropy scores than either very old or very young physicians.

[29]Sertraline is the generic for Zoloft; Citalopram is the generic for Celexa, and Fluoxetine is the generic for Prozac.

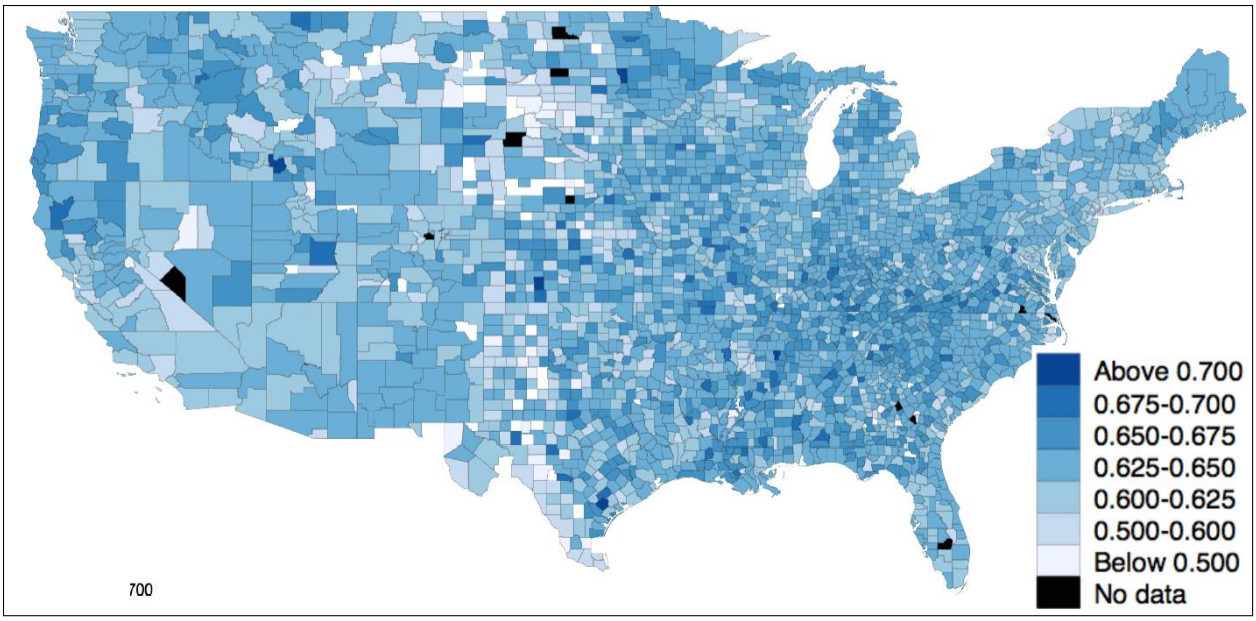

FIGURE 3B. Average Physician Entropy, by county, 2014

As discussed above, in addition to entropy, a second indicator of a physician's willingness to experiment is the extent to which they follow practice guidelines. We will consider guidelines provided by the American Psychiatric Association (Gelenberg et al. (2010)), the UK National Institute for Health and Care Excellence (NICE), and the Canadian government. The NICE guidelines suggest that clinicians should start with an SSRI, and if that does not work, then they should consider a drug in a different class (NICE, 2017). The Canadian guidelines point out that even within drug classes, some drugs are more efficacious and suggest that if the first drug does not work, clinicians should switch to a more effective drug. They provide rankings based on comparisons of the effectiveness of different drugs as first line treatments in clinical trials (Kennedy et al. (2016)).[30] The American Psychiatric Association's (APA) guidelines for treatment of major depressive disorder advise that if one drug is not effective, the patient should switch to another but they do not specify what that should be. They do however note that "the following medications are optimal for most patients: SSRIs, SNRIs, Mirtazapine, and Bupropion" (page 31, Gelenberg et al. (2010)) which excludes two drugs that together accounted for 17.49% of the market in 2014 (see Appendix Table C1). In addition, all guidelines urge caution in the use of "drug cocktails." At issue is that most possible combinations have not been evaluated in clinical trials, so the possible drug interactions or side effects are largely unknown.

Entropy and the violation of guidelines are closely related since following a guideline means ruling out choices that a doctor would otherwise have used. Figure 4 illustrates this relationship by showing the actual distribution of entropy scores in the data and comparing it to the distribution that would be obtained if all physicians adhered to the APA guidelines. In the counter-factual

[30]If one takes these rankings literally, then some drugs are completely dominated by other drugs and should never be prescribed.

distribution, the prescriptions for medications that violate these guidelines are distributed over the remaining allowed options in proportion to actual prescription patterns. Figure 4 shows that the right tail of the entropy distribution would be compressed, and that the whole distribution would shift to the left. The figure illustrates that the APA guidance, while quite loose, would still be binding on practice styles if it were followed uniformly.

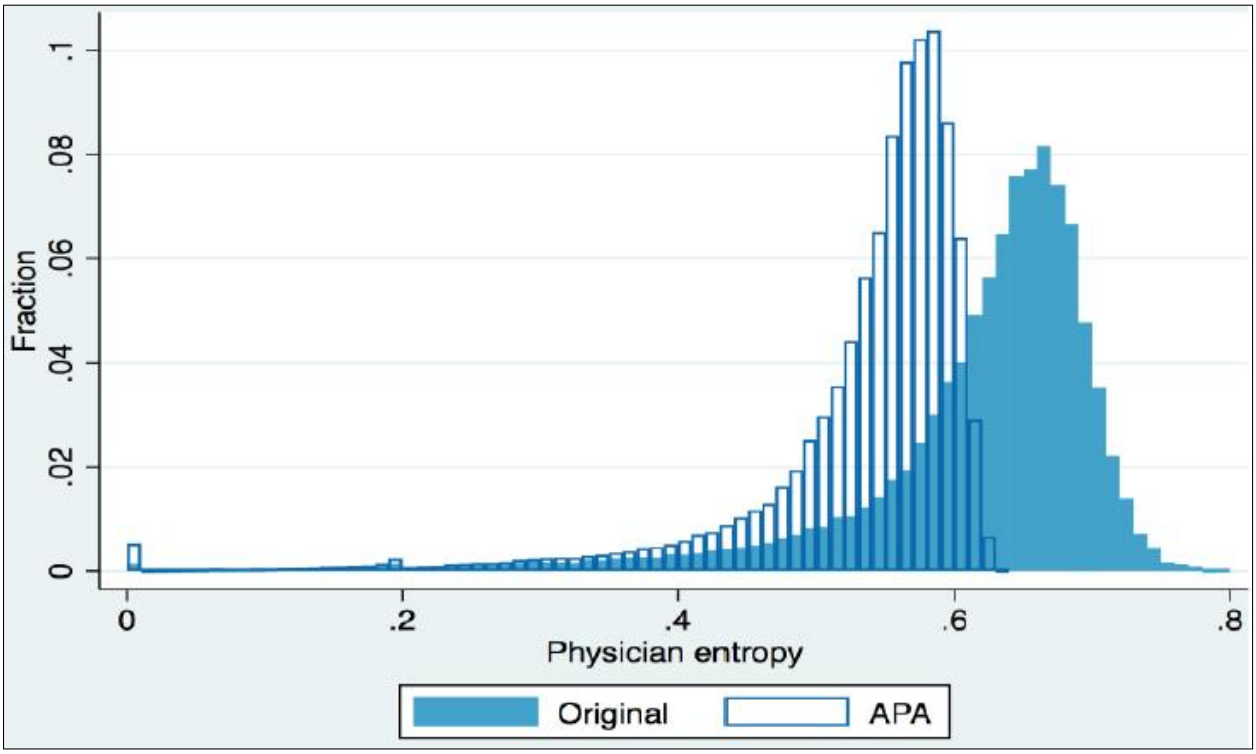

FIGURE 4. Following guidelines lowers entropy (data for 2013)

Table 3 provides an overview of the BCBS samples used in the estimations described below. As discussed above, it is possible and expected that patient heterogeneity drives some of the variation in use of anti-depressant drugs. Suppose sicker patients, or patients who are getting worse, are more likely to see psychiatrists as outpatients. Then, since psychiatrists have higher entropy scores on average, one might find that higher entropy was associated with worse patient outcomes. In order to account for this type of matching, Table 3 divides patients into those who ever saw a psychiatrist as an outpatient, and those who did not.[31] The first three columns are for the sample for whom we have non-missing entropy scores, while the last three columns are for the patients with a non-missing lagged transition (this sample is larger because it does not require matching to the IQVIA data). Table (3) shows that on average we follow depressed patients for about 11 to 12 months, and that they take anti-depressants for an average of 8 months during this time period.

[31]We do not divide patients into whether they are seeing a psychiatrist or not currently, because the same patient might start seeing a psychiatrist because they are getting worse. However, given the 11 or 12 month time period that we follow the average patient, it is safe to assume that patients who ever see a psychiatrist during this time interval will be seeing more skilled providers on average.

TABLE 1. Effect of Anti-Depression Drugs on Hamilton 17 Score for Depression Severity (mean score before treatment is 25.2)

Active Ingredient	Brand Name(S)	Class of Drug	Market Share 2014, Percent	Depression Reduction Effect	Standard Deviation Effect
Sertraline	Zoloft	SSRI	14.63	-9.90	7.78
Citalopram	Celexa	SSRI	12.83	-10.30	7.08
Fluoxetine	Prozac	SSRI	10.57	-9.40	6.13
Escitalopram	Lexapro	SSRI	9.68	-10.40	5.97
Paroxetine	Paxil	SSRI	5.32	-9.80	6.14
Trazodone	Oleptro	SARI	9.35	-15.70	9.00
Duloxetine	Cymbalta	SNRI	6.84	-10.70	7.00
Bupropion	Wellbutrin	NDRI	10.34	-12.00	8.70
Amitriptyline	Elavil	Tricyclic	5.18	-14.00	8.70
Venlafaxine	Effexor	SNRI	7.09	-12.10	8.71
Mirtazapine	Remeron	Tetracyclic	2.82	-14.00	7.70
Placebo	-			-8.00	6.67

Notes: These effects are culled from a number of meta-analyses of drug effects.
See the unpublished appendix for details

TABLE 2. Summary of Prescribing and Physician Entropy Scores in 2013, by Specialty Data on all prescriptions from Data IQVIA

	All physicians	GPs	Psychiatrists	Other medical
# Prescriptions (millions)	231.6	120.7	51.1	15.9
# Prescribers	767,985	267,898	49,523	214,928
Prescriptions/Provider	301.6	450.5	1032.5	73.8
Average Entropy Scores by Medical School Graduation Year				
<1975	0.624	0.624	0.661	0.503
1976-1985	0.628	0.636	0.662	0.503
1986-1995	0.623	0.635	0.655	0.482
1996+	0.613	0.630	0.637	0.448

Notes: Entropy calculations include only providers with ≥ 12 scripts in the year and are based on $m = 32$ separate molecules.

TABLE 3. Summary of BCBS Patient Data by Patient Outpatient Provider

	Non-missing Entropy Sample			Non-missing drug transition sample		
Patient Type:	All Patients	Ever Saw Psychiatrist	Never Saw Psychiatrist	All Patients	Ever Saw Psychiatrist	Never Saw Psychiatrist
# members	452,080	83,045	369,035	593,499	112,429	481,070
# member-months	5,413,368	1,117,687	4,295,681	6,556,938	1,424,983	5,131,955
# months/member	11.974	13.459	11.640	11.048	12.675	10.668
# months antidepressants/member	8.286	9.478	8.091	8.359	9.723	8.041
# changes in entropy/member	2.028	2.392	1.946			
# member-month with non-missing drug transitions				6,370,152	1,383,792	4,986,360
Percent drug transitions from t-2 to t-1 that violate each guideline (as a percent of row 5)						
UK				0.107	0.130	0.101
Canada				2.380	2.172	2.438
US				3.639	4.635	3.363
Cocktail				4.817	9.406	3.544
Costs (in Jan. 2013 dollars)						
total monthly cost: 50th p'tile	109.14	218.21	86.83	119.62	231.72	95.38
90th p'tile	1411.77	2025.54	1241.93	1490.91	2151.28	1296.71
pharmacy cost: 50th p'tile	23.95	48.36	20.15	26.54	53.06	22.29
90th p'tile	519.54	794.13	448.00	542.58	817.06	464.73
professionals cost: 50th p'tile	0.00	19.60	0.00	0.00	40.74	0.00
90th p'tile	504.02	699.73	451.21	523.73	729.02	464.25
facility cost: 50th p'tile	0.00	0.00	0.00	0.00	0.00	0.00
90th p'tile	108.34	198.30	88.63	132.70	253.13	105.76
99th p'tile	7825.66	9765.25	7228.97	8530.23	10886.29	7765.15
Facility Use						
1 if any ER/hospitalization	0.0237	0.0323	0.0214	0.025	0.035	0.022
1 if any ER/hosp. for mental health	0.0097	0.0170	0.0077	0.011	0.019	0.008

Notes: The treatment period is defined as up to 1 month before the first observed month with an anti-depressant script up till 3 months after the last observed month with an anti-depressant script. ER/Hosp. visits are considered to have been for mental health if that is one of the indications listed.

The second panel of Table (3) provides information on how often patients changed medications in a way that violated one of the prescription guidelines discussed above. The table indicates that relatively small fractions of transitions actually violated a guideline, though the informal APA guideline recommending against trazadone and the older tricyclic anti-depressant amitriptyline is most likely to be violated, especially in patients who ever saw a psychiatrist as an outpatient. Also, relatively large numbers of patients receive a cocktail of drugs, especially from psychiatrists. See the Appendix Table C2 for a transition matrix that shows the frequency of each transition between drugs and an example of how we implement the APA guidelines in our data.

The third and fourth panels of Table (3) present data on the costs of care. These include total monthly costs, and costs broken into pharmacy, professional (e.g. doctor visits), and facility (e.g. hospital) claims. All of these costs are extremely right skewed: The modal patient is not very expensive, while the 90th percentile (or in the case of facilities costs, the 99th percentile) patient incurs considerable costs, especially when one considers that these are monthly numbers. Since this is BCBS claims data, these figures represent the actual amount paid by BCBS to the various providers. One can also see that the division of patients into those who saw psychiatrists as outpatients and those who did not is meaningful. The former have higher costs in every category suggesting that they are in fact sicker. These patients will also, by construction, be seeing higher skilled practitioners, on average.

In addition to costs, we also look at indicators for whether the patient used the ER or was hospitalized. We focus special attention on these outcomes because they represent a severe crisis for a patient, which may be extremely disruptive to their lives, resulting in stigma, lost work or school, and so on. Overall, about 2% of these depressed patients visit an ER or are hospitalized in any given month. Hospital personnel can list many diagnosis codes on the claim forms, and have some financial incentive to include all diagnoses that will indicate a more complex case. On the other hand, it is not clear whether an ER visit for something like a self-inflicted wound will necessarily receive a mental health code. Therefore, we look at all ER visits, and also at those that have any mental health diagnosis listed. About 1% of these patients have an ER or hospital visit that lists a mental health diagnosis in any given month. The risk is more than twice as large for patients who have seen a psychiatrist as an outpatient compared to other patients.

4.1. **Estimation:** In order to examine the effects of entropy on patient outcomes, we estimate models of the form:

$$Y_{ijt} = a_0 + b_1\phi_{jt-1} + b_2 x_i + b_3 county_i + b_4 y_t + e_{ijt}, \tag{15}$$

or alternatively:

$$Y_{ijt} = a_i + b_1\phi_{jt-1} + b_2 y_t + e_{ijt}, \tag{16}$$

where Y is one of the outcomes discussed above, x are the observable patient characteristics (age category and gender), *county* indicates county fixed effects, and y indicates year fixed effects. The second specification, which includes patient fixed effects, does not include the observable patient

characteristics or county fixed effects. By including a patient fixed effect, we control for unobserved characteristics of the patient including their mean overall severity, history prior to appearing in the claims data, taste for medication, and so on. We estimate the model separately for patients who ever saw a psychiatrist as an outpatient and for those who did not, in order to allow for the effects of experimentation to be different for patients who see providers with different skill levels, on average. Also, it is important to note that the doctor's entropy score is measure at *t-1* for an outcome measured at time *t,* so that the measure of practice style always preceeds the outcome.

In order to examine the effects of violations of treatment guidelines, we estimate models using the same outcomes of the following form:

$$Y_{ijt} = a_0 + b_1 V_{ijt-1} + b_2 x_i + b_3 county_i + b_4 y_t + e_{ijt}, \tag{17}$$

or alternatively:

$$Y_{ijt} = a_i + b_1 V_{ijt-1} + b_2 y_t + e_{ijt}, \tag{18}$$

where V is a vector of four indicators each equal to one if a drug transition between *t-2* and *t-1* violated one of the three guidelines discussed above, or if it involved the prescription of a drug cocktail. Hence, in this formulation, we look at the outcome one period after a change in the drug regime, that is, at period t.

These regression models will show how our proxies for experimentation are related to patient outcomes. Including patient fixed effects offers a powerful way to control for patient heterogeneity in order to isolate the effects of physician practice style. All standard errors are clustered on the physician's ID in order to allow for correlations in treatment between patients seeing the same physician.

5. Empirical Results

Regressions of patient outcomes on provider entropy scores are shown in Table 4 for the full sample, as well as for the two subsamples defined by whether the patient had ever seen a psychiatrist. Odd numbered columns control for county fixed effects, broad patient age categories and gender. Even numbered columns control for patient fixed effects. The first two columns in the first row indicate that patient heterogeneity is important: In the regressions without patient fixed effects, it appears that provider entropy increases costs, whereas once patient fixed effects are included in the model, entropy is shown to have a significant negative effect on costs. Since the dependent variable is in terms of log costs, the coefficient can be interpreted as an elasticity: A one unit change in entropy would lead to a 12.7% decrease in total costs. In practice entropy varies from around 0.4 to 0.8 as shown in Figure 4, with most of the mass between 0.5 and 0.75. Considering an increase in entropy of 0.25 then, one could expect to see a decrease in costs of 3.2%. The remaining columns show that an increase in entropy is associated with a large decrease in non-drug costs (a .25 increase in entropy would reduce these costs by 5.4%). Some of this reduction is coming from reductions in the probability of ER visits and hospitalizations. An increase of 0.25 in physician entropy is associated with a 0.125 point decline in the probability of such visits on a baseline of 2.4%, or a

5.2% reduction in the probability of any ER visit or hospitalization. Similarly, the probability of an ER visit or hospitalization specifically for a mental health diagnosis also declines by about 5.0% with an increase of 0.25 in physician entropy.

These overall patterns reflect some differences in the effects of physician entropy by patient group. We can think of the group who ever saw a psychiatrist as an outpatient over the sample period as both sicker on average, and seeing a more skilled practitioner on average. Including a patient fixed effect controls for patient heterogeneity so that in the fixed effects models we can interpret differential effects of entropy as reflecting differences in average physician skill net of differences in the patient's average condition. Among patients who are seeing these higher skilled practitioners, higher entropy is associated with a reduction in ER visits and hospitalizations, both overall and specifically for mental health. In this population, an increase of 0.25 in the provider entropy score reduces the probability of any ER visit or hospitalization by 10.2% whereas the probability of a visit specifically for a mental health diagnosis is reduced by 13.2%. However, there is no significant effect of entropy on costs in this group once patient fixed effects are included, suggesting that the savings in terms of facilities charges are offset by increases in other costs. For example, a more skilled provider is likely to be more expensive, and may require more visits.

The group of patients who never saw a psychiatrist as an outpatient are less sick on average and are seeing less skilled providers. In models that include patient fixed effects to deal with patient heterogeneity we do not see significant effects of provider entropy scores on ER visits or hospitalizations, though we do see a reduction in total costs and in all non-drug costs. The coefficients indicate that a 0.25 increase in provider entropy at month *t-1* would be associated with a 4.5% reduction in total patient costs at month *t*.

Table (5) shows the estimated effects of having a drug transition from month *t-2* to *t-1* that violated treatment guidelines on patient outcomes in month *t*. While transitions that violated the UK guidelines increase costs without usually increasing ER visits or hospitalizations, transitions that violated the other treatment guidelines appear to have uniformly harmful effects, increasing costs, and ER visits and hospitalizations. Panel A, which shows the results for all patients, indicates that the coefficient estimates are generally smaller when patient fixed effects are included in the model (except for the Canadian guidelines), indicating that physicians are more likely to violate guidelines or to prescribe drug cocktails for sicker patients. The estimates suggest, for example, that violating the U.S. guidelines would increase total costs by 28.8%, increase the probability of any ER visit or hospitalization by 16.0%, and increase the probability of an ER visit or hospitalization specifically for mental health by 18.2%. Prescribing drug cocktails has a similar effect on ER visits and hospitalizations, but sharply increases total costs, by 50.4%. Some of the higher cost is mechanical in the sense that taking more drugs will usually cost more than taking less drugs. But column (4) shows that total non-drug costs also rise by 35.6%.

TABLE 4. Regressions of Patient Outcomes on Lagged Provider Entropy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(6)
Outcome:	**ln(total cost)**	**ln(total cost)**	**ln(non-drug costs)**	**ln(non-drug costs)**	**ER or Hospital**	**ER or Hospital**	**ER/Hosp for Mental Health**	**ER/Hosp for Mental Health**
Patient FE	no	yes	no	yes	no	yes	no	yes
Panel A: All patient observations								
Provider Entropy	0.284	-0.127	-0.122	-0.215	-0.004	-0.005	0.007	-0.002
	(0.032)	(0.039)	(0.034)	(0.037)	(0.001)	(0.002)	(0.001)	(0.001)
Constant	3.822	5.028	2.520	3.191	0.025	0.030	0.004	0.011
	(0.032)	(0.025)	(0.035)	(0.024)	(0.001)	(0.001)	(0.001)	(0.001)
Mean Dep. Variable	3.976	3.976	2.511	2.511	0.024	0.024	0.010	0.010
Adj. R2	0.043	0.382	0.027	0.303	0.004	0.092	0.002	0.073
# Obs. (millions)	5.413	5.413	5.413	5.413	5.413	5.413	5.413	5.413
# Members	452,080	452,080	452,080	452,080	452,080	452,080	452,080	452,080
Panel B: Patients who ever saw a psychiatrist as an outpatient								
Provider Entropy	0.022	0.020	-0.232	-0.074	-0.013	-0.013	0.005	-0.009
	(0.064)	(0.075)	(0.081)	(0.072)	(0.003)	(0.004)	(0.002)	(0.003)
Constant	4.391	5.516	3.016	3.700	0.035	0.046	0.014	0.024
	(0.076)	(0.050)	(0.089)	(0.048)	(0.003)	(0.003)	(0.002)	(0.002)
Mean Dep. Variable	4.557	4.557	3.024	3.024	0.032	0.032	0.017	0.017
Adj. R2	0.057	0.378	0.051	0.353	0.012	0.109	0.008	0.098
# Obs. (millions)	1.118	1.118	1.118	1.118	1.118	1.118	1.118	1.118
# Members	83,045	83,045	83,045	83,045	83,045	83,045	83,045	83,045
Panel C: Patients who never saw a psychiatrist as an outpatient								
Provider Entropy	-0.161	-0.180	-0.562	-0.265	-0.008	-0.002	0.003	0.002
	(0.033)	(0.045)	(0.034)	(0.042)	(0.001)	(0.002)	(0.001)	(0.001)
Constant	3.991	4.907	2.682	3.061	0.026	0.025	0.005	0.007
	(0.035)	(0.028)	(0.036)	(0.027)	(0.001)	(0.001)	(0.001)	(0.001)
Mean Dep. Variable	3.825	3.825	2.378	2.378	0.021	0.021	0.008	0.008
Adj. R2	0.047	0.374	0.028	0.282	0.004	0.086	0.002	0.060
# Obs. (millions)	4.296	4.296	4.296	4.296	4.296	4.296	4.296	4.296
# Members	369,035	369,035	369,035	369,035	369,035	369,035	369,035	369,035

Notes: All models include year and month fixed effects. Regressions in odd numbered columns include county fixed effects, patient age, and gender. These controls are dropped when patient fixed effects are added. Standard errors are clustered on provider ID.

TABLE 5. Outcomes at t when Drug Transition from (t-2) to (t-1) Violated Prescribing Guidelines

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**ln(total cost)**	**ln(total cost)**	**ln(non-drug costs)**	**ln(non-drug costs)**	**ER or Hospital**	**ER or Hospital**	**ER/Hosp for Mental Health**	**ER/Hosp for Mental Health**
Patient FE	no	yes	no	yes	no	yes	no	yes
Panel A: All patient observations								
Violation UK Guidelines	0.636	0.255	0.776	0.265	0.013	0.004	0.006	-0.001
	(0.031)	(0.028)	(0.038)	(0.034)	(0.002)	(0.002)	(0.002)	(0.002)
Violation US Guidelines	0.431	0.288	0.556	0.303	0.012	0.004	0.006	0.002
	(0.007)	(0.006)	(0.008)	(0.007)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Violation Can. Guidelines	0.391	0.482	0.425	0.407	0.007	0.004	0.004	0.002
	(0.008)	(0.007)	(0.009)	(0.008)	(0.005)	(0.0005)	(0.0003)	(0.0003)
Cocktail	1.028	0.504	0.829	0.356	0.020	0.005	0.013	0.002
	(0.009)	(0.008)	(0.011)	(0.008)	(0.0005)	(0.0005)	(0.0003)	(0.0004)
Mean Dep. Variable	4.057	4.057	2.561	2.561	0.025	0.025	0.011	0.011
Adj. R2	0.054	0.389	0.035	0.314	0.006	0.101	0.003	0.082
# Obs. (millions)	6.370	6.370	6.370	6.370	6.370	6.370	6.370	6.370
# Members	592,147	592,147	592,147	592,147	592,147	592,147	592,147	592,147
Panel B: Patients who ever saw a psychiatrist as an outpatient								
Violation UK Guidelines	0.608	0.367	0.872	0.431	0.015	0.001	0.010	-0.003
	(0.056)	(0.052)	(0.072)	(0.062)	(0.006)	(0.005)	(0.004)	(0.004)
Violation US Guidelines	0.291	0.256	0.420	0.292	0.014	0.006	0.008	0.004
	(0.013)	(0.011)	(0.015)	(0.012)	(0.001)	(0.001)	(0.001)	(0.001)
Violation Can. Guidelines	0.493	0.498	0.479	0.403	0.008	0.003	0.006	0.003
	(0.016)	(0.015)	(0.019)	(0.016)	(0.001)	(0.001)	(0.001)	(0.001)
Cocktail	0.776	0.554	0.640	0.422	0.019	0.005	0.014	0.003
	(0.013)	(0.012)	(0.018)	(0.012)	(0.001)	(0.001)	(0.001)	(0.001)
Mean Dep. Variable	4.735	4.735	3.083	3.083	0.035	0.035	0.019	0.019
Adj. R2	0.070	0.387	0.061	0.366	0.012	0.121	0.009	0.108
# Obs. (millions)	1.384	1.384	1.384	1.384	1.384	1.384	1.384	1.384
# Members	112,175	112,175	112,175	112,175	112,175	112,175	112,175	112,175
Panel C: Patients who never saw a psychiatrist as an outpatient								
Violation UK Guidelines	0.580	0.213	0.677	0.204	0.012	0.005	0.004	-0.001
	(0.036)	(0.033)	(0.043)	(0.040)	(0.003)	(0.003)	(0.002)	(0.002)
Violation US Guidelines	0.408	0.301	0.535	0.308	0.010	0.004	0.003	0.002
	(0.008)	(0.007)	(0.009)	(0.008)	(0.0005)	(0.0005)	(0.0003)	(0.0003)
Violation Can. Guidelines	0.374	0.479	0.418	0.410	0.007	0.004	0.003	0.002
	(0.009)	(0.008)	(0.009)	(0.009)	(0.0005)	(0.0005)	(0.0003)	(0.0003)
Cocktail	0.895	0.467	0.668	0.308	0.015	0.005	0.008	0.002
	(0.012)	(0.010)	(0.013)	(0.011)	(0.001)	(0.001)	(0.0004)	(0.0004)
Mean Dep. Variable	3.896	3.896	2.416	2.416	0.022	0.022	0.008	0.008
Adj. R2	0.056	0.381	0.033	0.291	0.005	0.093	0.002	0.066
# Obs. (millions)	4.986	4.986	4.986	4.986	4.986	4.986	4.986	4.986
# Members	479,972	479,972	479,972	479,972	479,972	479,972	479,972	479,972

Notes: Each column of each panel represents a single regression model. All models include year and month fixed effects. Regressions in odd numbered columns include county fixed effects, patient age, and gender. These controls are dropped when patient fixed effects are added. Standard errors are clustered on provider ID.

The next two panels of Table (5) examine the effects of violations of the three different treatment guidelines on the two subsamples of patients defined by whether or not they ever saw a psychiatrist as an outpatient. As discussed above, patients who have seen a psychiatrist are seeing providers who are more skilled on average. The effects are quite similar across the two groups indicating that violations of treatment guidelines are undesirable whether they are committed by more or less

skilled practitioners. For instance, although the point estimates on ER visits and hospitalizations are higher in the patients who ever saw a psychiatrist as an outpatient, this is because the baseline risk of this outcome is higher in this group. The percentage effects are similar to those just described: For example, a violation of U.S. guidelines would increase the probability of any ER visit or hospitalization by 17.1%, and the probability of an ER visit or hospitalization for mental health by 21.0%.

Overall, the results in Table (4) indicate that higher entropy is associated with better patient outcomes and, as predicted, higher entropy has more positive effects among more skilled practitioners. Table (5) suggests, however, that there is a limit on experimentation and a useful role for guidelines that restrict some prescribing practices since violating guidelines leads to worse patient outcomes regardless of the skill level of the practitioner. Hence, empirically there seems to be a "sweet spot" in terms of physician experimentation.

6. Conclusions

We think of the sequence of drug choices as a multi-armed bandit problem, which involves a trade off between experimenting to learn more about what works best for a particular patient, and systematically choosing the alternative with the highest expected payoff. Experimentation will necessarily involve mistakes. Hence, some medical "errors" may be unavoidable pitfalls in the search to find the best treatment for a given patient. The novel feature of our model is that the payoff to experimentation depends on physician diagnostic skill as well as the time horizon (or discount factor) and beliefs. One insight of the model is that there will not necessarily be one correct treatment, even for identical patients since the best treatment will also depend on the doctor who is treating the patient. Other things equal, experimentation will be most beneficial when the physician is better able to draw the correct inference from the experiment, which is how we think of diagnostic skill. This insight is likely to be applicable in other markets with expert decision makers. For example, other things being equal, a skilled surgeon should perform more surgeries on marginal patients than a less skilled surgeon.

A second contribution of our paper is to suggest using the Upper Confidence Bound (UCB) algorithm from the machine learning literature both as a tractable way to solve the bandit problem, and as a model that approximates how decision makers may actually behave (we know that they do not typically calculate the Gitten's index). We show that the UCB will attain the optimal payoff as the number of time periods becomes large and that it includes short-run (myopic) Bayesian optimal choice as a special case. Using simulations based on data from clinical drug trials, we illustrate the differences between doctors who are optimally experimental (i.e. follow the UCB algorithm with a fixed tuning parameter) and those who choose the drug with the highest expected value in the current period at each time period. These simulations suggest that whether it is good (from the patient point of view) to have an "experimental" doctor or not is an empirical matter which depends on the skill of the physician as well as the treatment time horizon.

Armed with these insights, we turn to claims data on hundreds of thousands of patients who were treated with anti-depressants to evaluate the relationship between the doctor's propensity

to experiment and patient outcomes. Using Shannon's entropy score as an empirical proxy for a doctor's propensity to experiment we first show that as predicted, higher skilled providers are more experimental. Also, as an empirical matter and over the range of our data, seeing a more experimental provider improves a patient's outcomes measured using total costs, non-drug costs, emergency room visits and hospitalizations. These effects can be quite large, for example, among patients who ever saw a psychiatrist as an outpatient, an increase of 0.25 in the provider entropy score reduces the probability of an ER visit or hospitalization with a mental health diagnosis by 13.2%. We also look at a more dynamic measure of practice style by examining the effects of transitions from one drug to another. Although following guidelines mechanically lowers entropy if the guidelines are binding on behavior, these results suggest limits on the value of experimentation in that patients whose doctors violate prescribing guidelines have worse patient outcomes. For example, drug transitions between *t-2* and *t-1* that violate U.S. guidelines increase the probability of an ER visit or hospitalization with a mental health indication at time t by 21.0%. These findings suggest that optimal treatment guidelines may be loose enough to allow some experimentation, but tight enough to rule out bad practice.

Our results have number of implications for health policy. Much of the literature about variations in observed practice style begins with the assumption that conditional on price and patient characteristics there is a well defined optimal choice.[32] From this perspective, the goal of health policy is mainly to get the price "right" so that the optimal choice will be made. Our results show that in the presence of match specific treatment and learning there cannot be a single "optimal" choice. Rather, there is an optimal practice style that varies with physician skill and the uncertain information that the physician collects while treating the patient. Still, some treatments are ill advised, and patients are better off on average when physicians follow national standards that recommend against certain treatments. Overall, our results suggest that optimal policy should give physicians discretion within well defined boundaries; that is physicians need to be given the right to practice medicine with limits on behaviors that fall outside of accepted practice.

Appendix A. Proofs of Propositions

Proposition 1. *The value of information about drugs A and B is:*

$$V_d = \sigma_d L\left(\frac{\mu_A - \mu_B}{\sigma_d}\right),$$

where $\sigma_2^2 = \frac{\rho}{\rho_d(\rho+\rho_d)}$*, and, for* $x \geq 0$*,* $L(\cdot)$ *is the unit-loss function defined by:*

$$L(x) = (1 - F(x))(\phi(x) - x),$$

$F(x)$ *is the cumulative distribution function for the Normal distribution and* $\phi(x) = E\{\gamma|\gamma \geq x\}$ *is the expected value of a lower truncated Normal distribution,* $\gamma \sim N(0,1)$*.*

[32]For example Skinner (2012)'s nice review of the literature on regional variation is organized around the factors that lead physicians to choose one treatment versus another.

Proof. We begin with V_A, the value of information for the drug with the highest expected payoff in period 1. If there is no experimentation, then the choice in period 1 is A, with expected value $\mu_A > \mu_B$.

If the physician tries A in period 0, then the expected value given the signal is y_0 is:

$$E\left\{y_2|y_1, d_0 = A\right\} = E\left\{e_A|y_1, d_0 = A\right\} = \frac{\rho y_1 + \rho_A \mu_A}{\rho + \rho_A}.$$

Similar to (3) we can write $y_1 = \mu_A + s_A \gamma_1$, where $\gamma_1 \sim N(0,1)$ and $s_A^2 = \frac{1}{\rho} + \frac{1}{\rho_A} = \frac{\rho + \rho_A}{\rho \rho_A}$ is the variance of y_1. Let $\Delta = \mu_A - \mu_B > 0$, then the value of information in period 1 as:

$$\begin{aligned}
V_A &= E\left\{max\left\{\mu_B, \frac{\rho y_1 + \rho_A \mu_A}{\rho + \rho_A}\right\}\right\} - \mu_A \\
&= E\left\{max\left\{-\Delta, \frac{\rho\left(y_1 - \mu_A\right)}{\rho + \rho_A}\right\}\right\} \\
&= E\left\{max\left\{-\Delta, \frac{\rho}{\rho + \rho_A} s_A \gamma_1\right\}\right\} \\
&= E\left\{max\left\{-\Delta, \frac{\rho}{\rho + \rho_A}\sqrt{\frac{\rho + \rho_A}{\rho \rho_A}}\gamma_1\right\}\right\}, \\
&= \sigma_A E\left\{max\left\{-\Delta/\sigma_A, \gamma_1\right\}\right\},
\end{aligned}$$

where let $\sigma_A^2 = \frac{\rho}{\rho_A(\rho_A + \rho)}$. We can now write the value of information as:

$$V_A = \sigma_A E\left\{max\left\{-\frac{\Delta}{\sigma_A}, \gamma\right\}\right\},$$

where γ is the standard Normal distribution. This expression can be written in terms of the unit-Normal linear loss function (Raiffa and Schlaifer (2000)) for $x \leq 0$:

$$\begin{aligned}
L(x) &= E\left\{max\left\{x, \gamma\right\}\right\} \\
&= \int_{-\infty}^{x} x f(z)\, dz + \int_{x}^{\infty} z f(z)\, dz \\
&= F(x)\, x + \left(1 - F(x)\right)\phi(x),
\end{aligned}$$

where $f(), F()$, are the distribution and cumulative distribution functions for the standard Normal, while $\phi(x) = E\left\{\gamma|\gamma \geq x\right\}$ is the expected value from an upper truncated Normal distribution. Notice that:

$$lim_{x \to -\infty} L(x) = 0$$

and that $L_x(x) = F(x) > 0$. Hence this loss function is positive and increasing with $x \leq 0$. Thus we have:

$$V_A = \sigma_A L\left(-\Delta/\sigma_A\right) > 0.$$

When $x \geq 0$ the unit lost function is defined by:

$$E\left\{max\left\{x, \gamma\right\}\right\} = L(x) + x,$$

and hence:

$$\begin{aligned} L(x) &= E\{max\{x,\gamma\}\} - x \\ &= (1-F(x))(\phi(x)-x) \\ &= L(-x), \end{aligned} \tag{19}$$

from which we get:

$$V_A = \sigma_A L(\Delta/\sigma_A).$$

Next, consider V_B. The formula is similar, but differs due to the fact that the expression is compared to the same counterfactual:

$$\begin{aligned} V_B &= E\left\{max\left\{\mu_A, \frac{\rho y_1 + \rho_B \mu_B}{\rho+\rho_B}\right\}\right\} - \mu_A, \\ &= E\left\{max\left\{0, \frac{\rho}{\rho+\rho_B} s_B \gamma_1 - \Delta\right\}\right\}. \\ &= \sigma_B E\{max\{\Delta/\sigma_B, \gamma_1\}\} - \Delta. \end{aligned}$$

Now from (19) and the fact that $\Delta > 0$, we have:

$$V_B = \sigma_B L(\Delta/\sigma_B).$$

□

Corollary 1. *Experimenting with drug B has more value than drug A if and only if the uncertainty associated with drug B, σ_B, is higher than the uncertainty associated with drug A, σ_A. The value of such experimentation falls with the difference in expected effects ($\mu_A - \mu_B$).*

Proof. The function $L(x)$ is symmetric around zero, with its maximum value at 0, and decreases to zero as $x \to \pm\infty$. This immediately implies that the value of information falls with the absolute value of the differences in expected values, $|\mu_A - \mu_B|$. The effect of the uncertainty follows from:

$$\begin{aligned} \frac{\partial V_d}{\partial \sigma_d} &= L\left(\frac{\mu_B-\mu_A}{\sigma_d}\right) - \left(\frac{\mu_B-\mu_A}{\sigma_d}\right) L'\left(\frac{\mu_B-\mu_A}{\sigma_d}\right) \\ &= F\left(\frac{\mu_B-\mu_A}{\sigma_d}\right)\left(\frac{\mu_B-\mu_A}{\sigma_d}\right) + \left(1-F\left(\frac{\mu_B-\mu_A}{\sigma_d}\right)\right)\phi\left(\frac{\mu_B-\mu_A}{\sigma_d}\right) - \left(\frac{\mu_B-\mu_A}{\sigma_d}\right) F'\left(\frac{\mu_B-\mu_A}{\sigma_d}\right) \\ &= \left(1-F\left(\frac{\mu_B-\mu_A}{\sigma_d}\right)\right)\phi\left(\frac{\mu_B-\mu_A}{\sigma_d}\right) > 0 \end{aligned}$$

Finally, one can easily show that $\sigma_B \geq \sigma_A$ if and only if $\rho_A \geq \rho_B$, which completes the proof. □

Proposition 2. *For any period T of treatment, and any treatment strategy δ_j the physician's payoff satisfies:*

$$U_{ij}(T) \leq -log(T)\left\{2(n+1)\sigma_j^2 + O(1)\right\},$$

where $n+1$ is the number of possible treatments (including prescribing no drug).

Proof. Notice that we can rewrite:

$$\begin{aligned}
-U_{ij}(T) &= E\left\{\sum_{t=1}^{T}\sum_{d\in D} 1\left[d=d_{it}\right]\left(e_i^*-e_{id}\right)\right\},\\
&= \sum_{d\in D} E\left\{\sum_{t=1}^{T} 1\left[d=d_{it}\right]\left(e_i^*-e_{id}\right)\right\},\\
&= \sum_{d\in D} E\left\{n_{id}\right\}\left(e_i^*-e_{id}\right),\\
&\geq \sum_{d\in D}\left\{\frac{2}{\left(e_i^*-e_{id}\right)\rho_j}+O(1)\right\} log(T)\left(e_i^*-e_{id}\right),\\
&\geq log(T)\left\{2(n+1)\sigma_j^2+O(1)\right\},
\end{aligned}$$

where $1\left[d=d_{it}\right]$ is the indicator function - 1 if $d=d_{it}$ and zero otherwise, and $\rho_j=\frac{1}{\sigma_j^2}$ is the skill of the physician. □

The proof of proposition 3 follows immediately from the following, more precise version of the same result:

Proposition 3. *Suppose that the effects of the drugs are uncorrelated,* $K\geq\sqrt{2\pi e}$ *and* $\beta\geq 1.02$ *then with an uninformative prior:*

$$0\geq U_{ij}^{UCB}(T)\geq -\sum_{d\in D}\left(e_i^*-e_{id}\right)\left(\begin{array}{c}\left(\frac{4\beta^2\sigma_j^2}{\left(e_i^*-e_{id}\right)^2}\right)\left(\begin{array}{c}2log\left(\frac{K}{\sqrt{2\pi}}\right)+2log(T)\\ -log(2)-loglog\left(\frac{KT}{\sqrt{2\pi}}\right)\end{array}\right)\\ +\frac{2}{K}(1+logT)\end{array}\right)$$

Proof. The proof is a modification and strengthening of Reverdy et al. (2014), appendix B. To decrease the clutter in notation, the subscript j is omitted. Begin with:

$$\begin{aligned}
n_{id}^T &= \sum_{t=1}^{T} 1\left[d_t=d\right]\\
&\leq \sum_{t=1}^{T} 1\left[Q(d,z_{it})>Q\left(d_i^*,z_{it}\right)\right]\\
&\leq \eta+\sum_{t=1}^{T} 1\left[Q(d,z_{it})>Q\left(d_i^*,z_{it}\right),n_{id}^{t-1}\geq\eta\right], \qquad (20)
\end{aligned}$$

where η is a positive integer. At time t the physician picks d over d_i^* only if:

$$Q(d,z_{it})\geq Q\left(d_i^*,z_{it}\right).$$

This is true when at least one of the following equations holds:

$$\mu_{id^*}^t \leq e_i^* - C_{id^*}^t \tag{21}$$

$$\mu_{id}^t \geq e_{id} + C_{id}^t \tag{22}$$

$$e_i^* < e_{id} + 2C_{id}^t \tag{23}$$

where $C_{id}^t = \left(\frac{\sigma}{\sqrt{s_{id}^2+n_{id}^t}}\right) F^{-1}\left(1-\frac{1}{Kt}\right)$ and $s_{id}^2 = \frac{\sigma^2}{\Sigma_{i0}[d,d]}$ is the ratio of the variance of the physician's observation error over the prior variance in beliefs. Otherwise, if these 3 equations do not hold, then:

$$Q\left(d^*, z_{it}\right) = \mu_{id^*}^t + C_{id^*}^t > e_i^* \geq e_{id} + 2C_{id}^t > \mu_{id}^t + C_{id}^t = Q_i^t,$$

and option d^* is picked over d at time t.

The next step is to compute the probabilities of (21) and (22). Notice, that conditional upon n_{id}^t then

$$\mu_i^t[d] \sim N\left(\frac{s_{id}^2\mu_{id0} + n_{id}^t e_{id}}{s_{id}^2 + n_{id}^t}, \frac{n_{id}^t\sigma^2}{\left(s_{id}^2+n_{id}^t\right)^2}\right).$$

Equation (21) will hold if:

$$e_i^* \geq \mu_i^t[d^*] + \frac{\sigma}{\sqrt{s_{id^*}^2 + n_{id^*}^t}} F^{-1}\left(1-\alpha_t\right). \tag{24}$$

We can write

$$\mu_i^t[d^*] = \frac{s_{id}^2\mu_{id0} + n_{id}^t e_{id}}{s_{id}^2 + n_{id}^t} + \frac{\sigma\sqrt{n_{id^*}^t}}{s_{id^*}^2 + n_{id^*}^t} z,$$

where $z \sim N(0,1)$. Thus (24) holds iff:

$$z \leq -\sqrt{\frac{n_{id^*}^t + s_{id^*2}}{n_{id^*}^t}} F^{-1}\left(1-\alpha_t\right) + \frac{s_{id^*}^2\left(e_{id^*} - \mu_{id^*}^0\right)}{\sigma\sqrt{n_{id^*}^t}}.$$

For an uninformative prior $s_{id^*}^2 \to 0^+$ and hence (21) holds if and only if $z \leq -F^{-1}\left(1-\alpha_t\right)$, and therefore for an uninformative prior:

$$Pr\left[(21)\ holds\right] = \alpha_t = \frac{1}{Kt}.$$

Similarly, (22) with an uninformative prior holds if:

$$Pr\left[(22)\ holds\right] = \alpha_t = \frac{1}{Kt}.$$

The next step is to consider (23). It holds if:

$$e_{id^*} < e_{id} + \frac{2\sigma}{\sqrt{s_{id}^2 + n_{id}^t}} F^{-1}\left(1-\alpha_t\right).$$

Letting $\Delta_{id} = e_{id^*} - e_{id}$, this implies:

$$\Delta_{id} < \frac{2\sigma}{\sqrt{s_{id}^2 + n_{id}^t}} F^{-1}\left(1 - \alpha_t\right).$$

From Theorem 1 of Reverdy et al. (2014) we have for $\beta \geq 1.02$:

$$F^{-1}\left(1 - \alpha_t\right) < \beta\sqrt{-log\left(-\left(2\pi\alpha_t^2\right) log\left(2\pi\alpha_t^2\right)\right)},$$

which in turn implies:

$$\frac{\Delta_{id}^2}{4\beta^2\sigma^2}\left(s_{id}^2 + n_{id}^t\right) < -log\left(-\left(2\pi\alpha_t^2\right) log\left(2\pi\alpha_t^2\right)\right) = log\left(\frac{K^2t^2}{2\pi}\right) - loglog\left(\frac{K^2t^2}{2\pi}\right).$$

Notice that $log\left(x\right) - loglog\left(x\right)$ is increasing in x for $x \geq e$, and $K \geq \sqrt{2\pi e}$, and since $T \geq t \geq 1$ we have:

$$\begin{aligned}\frac{\Delta_{id}^2}{4\beta^2\sigma^2}\left(s_{id}^2 + n_{id}^t\right) &< log\left(\frac{K^2T^2}{2\pi}\right) - loglog\left(\frac{K^2T^2}{2\pi}\right) \\ &< 2log\left(\frac{K}{\sqrt{2\pi}}\right) + 2log\left(T\right) - log\left(2\right) - loglog\left(\frac{KT}{\sqrt{2\pi}}\right).\end{aligned}$$

Thus for an uninformative prior ($s_{id}^2 \to 0$) we have that (23) never holds if $n_{id}^t > \eta\left(T, K\right)$, where:

$$\eta\left(T, K\right) = \left(\frac{4\beta^2\sigma^2}{\Delta_{id}^2}\right)\left(2log\left(\frac{K}{\sqrt{2\pi}}\right) + 2log\left(T\right) - log\left(2\right) - loglog\left(\frac{KT}{\sqrt{2\pi}}\right)\right).$$

From (20) we have for $\eta = \eta\left(T, K\right)$:

$$\begin{aligned}E\left\{n_{id}^T\right\} &\leq \eta + \sum_{t=1}^{T} Pr\left[Q\left(d, z_{ijt}\right) > Q\left(d_i^*, z_{ijt}\right), n_{id}^{t-1} \geq \eta\right] \\ &= \eta + \sum_{t=1}^{T} Pr\left[equation\left(21\right), n_{id}^{t-1} \geq \eta\right] + \sum_{t=1}^{T} Pr\left[equation\left(22\right), n_{id}^{t-1} \geq \eta\right] \\ &< \eta + \frac{2}{K}\sum_{t=1}^{T}\frac{1}{t} \leq \eta + \frac{2}{K}\left(1 + logT\right) \qquad (25)\end{aligned}$$

From this we get that the payoff is:

$$U_i^{UCB}\left(T\right) = -\sum_{d \in D} \Delta_{id} E\left\{n_{id}^T\right\}$$

which with (25) completes the proposition. □

Appendix B. Drug effects and dropouts: How the data for Table 1 is constructed

B.1. **Overview.** This document will go drug-by-drug and show how the data used to model doctor tastes in the simulations were constructed. All cited papers are listed in the bibliography at the end. Each drug is listed by its pharmaceutical name, with its primary trade name included in parentheses. All effect means and standard deviations use the Hamilton-17 (HAMD-17) scale as their metric of improvement. Market shares were computed by the authors using the 2014 IQVIA data.

B.2. **Sertraline HCL (Zoloft).**

(1) **Effects:** All effect data were drawn from Hieronymus et al. (2015) table 2, which includes multiple sertraline studies. First, the average was taken over all sertraline studies to get average means and standard deviations of the HDRS-17 score both at baseline and endpoint. Mean effects were computed as the difference between average baseline score and average endpoint score. To compute standard deviations, we take advantage of the assumption that baseline scores and drug effects are independent. Under this assumption,

$$\sigma^2_{endpoint} = \sigma^2_{baseline} + \sigma^2_{effect}$$

Solving for σ_{effect}, we have:

$$\sigma_{effect} = \sqrt{\sigma^2_{endpoint} - \sigma^2_{baseline}}$$

B.3. **Citalopram HBR (Celexa).**

(1) **Effects:** All effect data were drawn from Hieronymus et al. (2015) table 2, which includes multiple citalopram studies. Means and standard deviations were computed using an identical procedure as used for sertraline.

B.4. **Fluoxetine HCL (Prozac).**

(1) **Effects:** All effect data were drawn from Hieronymus et al. (2015) table 2, which includes multiple fluoxetine studies. Means and standard deviations were computed using an identical procedure as used for sertraline.

B.5. **Escitalopram Oxal (Lexapro).**

(1) **Effects:** All effect data were drawn from Llorca et al. (2005), table 3. The mean effect was taken to be the difference in Hamilton-17 score between baseline and LOCF (Last Observation Carried Forward). Like for sertraline, we take advantage of the assumed independence between the baseline score and effect, and compute the standard deviation of the effect as:

$$\sigma_{effect} = \sqrt{\sigma^2_{LOCF} - \sigma^2_{baseline}}$$

B.6. **Trazodone HCL (Oleptro).**

(1) **Effects:** The mean effect was drawn from Kasper (1995) table 3, line 3 (Belgium). The effect is expressed as the mean change in HAMD-17 score for a single study. No data was found on the standard deviation of the effect for trazodone. However, Van Moffaert et al. (1995) claim that the standard deviation of mirtazapine's effect is about 20% lower than that for trazodone. Thus, we let $\sigma^{traz}_{effect} = \sigma^{mirt}_{effect}$, where σ^{mirt}_{effect} is defined below.

B.7. **Duloxetine HCL (Cymbalta).**

(1) **Effects:** The mean effect was drawn fom Detke et al. (2002), table 2. The effect is expressed as the mean change in HAMD-17 for a single study. The standard deviation of the effect was drawn from page 227 of Goldstein et al. (2002) which doesn't provide the standard deviation derived from their data but rather an "assumed" standard deviation of 7. We can hope that this standard deviation was informed by their data, but are not sure of this.

B.8. **Bupropion HCL XL (Wellbutrin XL).**

(1) **Effects:** No papers were found measuring the direct effect and standard deviation for bupropion. However, Maneeton et al. (2013)claims that these would be approximately the same as those for venlafaxine. For this reason, the effect and standard deviation of the effect of bupropion was made identical to that for venlafaxine (see below).

B.9. **Amitriptyline HCL (Elavil).**

(1) **Effects:** All effect data was drawn from Kasper (1995) page 30 (within the text). These data came from a single study of both amitriptyline and mirtazapine. You will notice they provide data for both "mean change from baseline" and "reductions at the endpoint". The data pulled are those corresponding to reductions at the endpoint.

B.10. **Venlafaxine (Effexor).**

(1) **Effects:** All effect data was drawn from table 1 of Kirsch et al. (2008), which includes several different of venlafaxine. In order to obtain a single figure for the mean and standard deviation of change, the average was taken over the relevant studies presented in the table. Note that the d denotes the standard deviation.

B.11. **Mirtazapine (Remeron).**

(1) **Effects:** All effect data was drawn from Kasper (1995), page 27 (within the text). These data came from an analysis of pooled data of mirtazapine trials.

B.12. **Paroxetine (Paxil).**

(1) **Effects:** All effect data were drawn from Hieronymus et al. (2015) table 2, which includes multiple paroxetine studies. Means and standard deviations were computed using the same procedure as for sertraline. Note that these paroxetine studies include a variety of different dosages.

B.13. **Placebo.**

(1) **Effects:** Most of the studies we have come across provide data on the effect of placebos on patients with major depressive disorder. We have defined our "placebo" effects and standard deviations by taking the average over the data provided in Hieronymus et al., which provides data on 18 different placebo-controlled trials. To compute the mean and standard deviation of the effect, we employ the same procedure used for sertraline, for example (please see above).

Appendix C. Appendix Tables

Table C1. Share of all antidepressant prescriptions, 2006 and 2014

	Drug Class		Molecule		Product	
	2006	**2014**	**2006**	**2014**	**2006**	**2014**
SSRI	54.40	53.90				
Sertraline			14.02	14.63	4.69	14.56
(Zoloft)					9.33	0.07
Citalopram			6.17	12.83	5.98	12.81
(Celexa)					0.19	0.02
Fluoxetine			11.08	10.57	10.69	10.53
(Prozac)					0.39	0.04
Escitalopram			13.05	9.68	..	9.50
(Lexapro)					13.05	0.18
Paroxetine			1.69	5.33	..	5.28
(Paxil)					1.78	0.04
SNRI	13.50	15.04				
Venlafaxine			9.14	7.09	0.18	6.99
(Effexor)					8.96	0.10
Duloxetine			4.33	6.84	..	6.36
(Cymbalta)					4.33	0.48
NDRI	11.45	10.46				
Bupropion			11.45	10.46	2.88	10.29
(Wellbutrin)					6.58	0.15
SARI	7.12	9.35				
Trazadone			7.12	9.35	7.11	9.35
Tricyclic	10.69	8.14				
Amitriptyline			6.72	5.18	6.62	5.15
(Elavil)					0.00	0.00
Tetracyclic	2.44	2.84				
Mirtazapine			2.36	2.82	2.36	2.82
(Remeron)				0.02	0.08	0.02
Total	99.60	99.73	87.13	94.78	87.18	94.73

Note: Brand names in parentheses. E.g. Zoloft is the brand name for the generic Sertraline. Each molecule generally has a brand and a generic product.

TABLE C2. Illustration of drug transitions and violations of the American Psychiatric Association Guidelines

	Prior													
	no drug	**sertra-line**	**citalo-pram**	**fluoxe-tine**	**bupro-pion**	**escital-opram**	**trazo-done**	**venlafa-xine**	**duloxe-tine**	**paroxe-tine**	**amitri-ptyline**	**mirtaz-apine**	**others**	**cocktail**
no drug	7.42%	1.72%	0.85%	0.62%	0.10%	0.72%	0.99%	0.78%	0.62%	0.62%	0.61%	0.14%	0.52%	0.26%
sertraline	1.94%	15.12%	0.01%	0.00%	0.00%	0.00%	0.04%	0.01%	0.01%	0.01%	0.02%	0.00%	0.01%	0.12%
citalopram	0.97%	0.01%	6.72%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.05%
fluoxetine	0.79%	0.01%	0.00%	5.69%	0.00%	0.00%	0.02%	0.01%	0.00%	0.00%	0.01%	0.00%	0.01%	0.06%
bupropion	0.17%	0.00%	0.00%	0.00%	0.88%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%
escitalopram	0.89%	0.01%	0.00%	0.00%	0.00%	5.47%	0.02%	0.00%	0.00%	0.00%	0.01%	0.00%	0.01%	0.04%
trazodone	1.07%	0.04%	0.02%	0.02%	0.00%	0.02%	5.95%	0.03%	0.02%	0.01%	0.01%	0.00%	0.01%	0.23%
venlafaxine	0.82%	0.02%	0.01%	0.01%	0.00%	0.01%	0.03%	9.28%	0.01%	0.01%	0.01%	0.00%	0.01%	0.12%
duloxetine	0.63%	0.01%	0.00%	0.00%	0.00%	0.00%	0.02%	0.01%	4.69%	0.00%	0.01%	0.00%	0.01%	0.09%
paroxetine	0.71%	0.01%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	6.27%	0.01%	0.00%	0.00%	0.05%
amitriptyline	0.64%	0.01%	0.01%	0.01%	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	4.32%	0.00%	0.01%	0.09%
mirtazapine	0.13%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.81%	0.00%	0.03%
others	0.50%	0.01%	0.01%	0.01%	0.00%	0.01%	0.01%	0.01%	0.01%	0.00%	0.01%	0.00%	3.91%	0.08%
cocktail	0.24%	0.12%	0.05%	0.06%	0.02%	0.05%	0.22%	0.12%	0.09%	0.05%	0.10%	0.03%	0.08%	2.55%
Total	**16.93%**	**17.09%**	**7.68%**	**6.42%**	**1.01%**	**6.29%**	**7.36%**	**10.28%**	**5.47%**	**7.00%**	**5.11%**	**1.00%**	**4.59%**	**3.78%**
# 1000 Obs.	625	630	283	237	37	232	271	379	202	258	188	37	169	139

Notes: This transition matrix is created from the BCBS data. Across the top are drug choices in t-1, while down the side are drug choices at time t. For simplicity, we break out the top 11 drugs by market share and group the rest into "others." We also consider drug "cocktails," i.e. multiple drugs, separately. The boxes highlighted in green on the main diagonal represent cases in which the patient stayed with the same treatment between t-1 and t. This represents the majority of cases. Boxes highlighted in red represent our interpretation of transitions that violated APA guidelines. These specify that most patients can be successfully treated using the unhighlighted drugs and without drug cocktails. In cases where a patient was already taking a drug like trazodone, we give the doctor the benefit of the doubt by assuming that if the patient sticks with it between t-1 and t, then it was working for that patient.

References

Abernethy, J., P. Bartlett, and A. Rakhlin (2007). *Multitask learning with expert advice*, Volume 4539 of *Lecture Notes in Computer Science*, pp. 484–+.

Acuna, D. E. and P. Schrater (2010). Structure of learning in human sequential decision-making. *Plos Computational Biology 6*(12).

Adli, M., K. Wiethoff, T. C. Baghai, R. Fisher, F. Seemüller, G. Laakmann, P. Brieger, J. Cordes, J. Malevani, G. Laux, I. Hauth, H.-J. Möller, K.-T. Kronmüller, M. N. Smolka, P. Schlattmann, M. Berger, R. Ricken, T. J. Stamm, A. Heinz, and M. Bauer (2017). How effective is algorithm-guided treatment for depressed inpatients? results from the randomized controlled multicenter german algorithm project 3 trial. *International Journal of Neuropsychopharmacology 20*(9), 721–730.

Auer, P., N. Cesa-Bianchi, and P. Fischer (2002). Finite-time analysis of the multiarmed bandit problem. *Machine learning 47*(2-3), 235–256.

Baicker, K. and A. Chandra (2005). The effect of malpractice liability on the delivery of health care. In D. M. Cutler and A. M. Garber (Eds.), *Frontiers of Health Policy Research*, Volume 8. Berkeley, CA: The Berkeley Electronic Press.

Bergemann, D. and J. Valimaki (2006, January). Bandit problems. techreport 1551, Cowles Foundation.

Berndt, E. R., R. S. Gibbons, A. Kolotilin, and A. L. Taub (2015). The heterogeneity of concentrated prescribing behavior: Theory and evidence from antipsychotics. *Journal of Health Economics 40*, 26 – 39.

Bloom, N., C. Propper, S. Seiler, and J. Van Reenen (2015). The impact of competition on management quality: Evidence from public hospitals. *The Review of Economic Studies 82*(2), 457–489.

Brezzi, M. and T. L. Lai (2002, November). Optimal learning and experimentation in bandit problems. *Journal of Economic Dynamics and Control 27*(1), 87–108.

Cappé, O., A. Garivier, O.-A. Maillard, R. Munos, and G. Stoltz (2013). Kullback-leibler upper confidence bounds for optimal sequential allocation. *The Annals of Statistics 41*(3), 1516–1541.

Chan, D. C. (2015, April 19). Tacit learning and influence behind practice variation: Evidence from physicians in training. mimeo.

Chandra, A. and D. O. Staiger (2007). Productivity spillovers in health care: Evidence from the treatment of heart attacks. *Journal of Political Economy 115*(1), pp.103–140.

Cipriani, A., X. Zhou, C. Del Giovane, S. E. Hetrick, B. Qin, C. Whittington, D. Coghill, Y. Zhang, P. Hazell, S. Leucht, P. Cuijpers, J. Pu, D. Cohen, A. V. Ravindran, Y. Liu, K. D. Michael, L. Yang, L. Liu, and P. Xie (2016). Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network meta-analysis. *The Lancet*. June 8.

Coscelli, A. (2000). The importance of doctors' and patients' preferences in the prescription decision. *The Journal of Industrial Economics 48*(3), 349–369.

Crawford, G. S. and M. Shum (2005). Uncertainty and learning in pharmaceutical demand. *Econometrica 73*(4), 1137–1173.

Cuddy, E., J. M. Currie, and W. B. MacLeod (2018). Are small area variations due to spillovers in practice styles: The case of depression. Mimeo, Princeton University.

Currie, J., W. B. MacLeod, and J. V. Parys (2016, May). Physician practice style and patient health outcomes: The case of heart attacks. *Journal of Health Economics 47*, 64–80.

Currie, J. M. and W. B. MacLeod (2017, January). Diagnosis and unnecessary procedure use: Evidence from c-section. *Journal of Labor Economics 35*(1), 1–42.

Cutler, D., J. Skinner, A. D. Stern, and D. Wennberg (2013, August). Physician beliefs and patient preferences: A new look at regional variation in health care spending. Technical Report 19320, NBER.

Detke, M. J., Y. Lu, D. J. Goldstein, J. R. Hayes, and M. A. Demitrack (2002). Duloxetine, 60 mg once daily, for major depressive disorder: a randomized double-blind placebo-controlled trial. *The Journal of Clinical Psychiatry 63*(4), 308–315.

Diaconis, P. and D. Freedman (1986). On the consistency of Bayes estimates. *The Annals of Statistics 14*(1), 1–26.

Dickstein, M. J. (2014, April). Efficient provision of experience goods: Evidence from antidepressant choice. Stanford University.

Dickstein, M. J. (2015, April). Physician vs. patient incentives in prescription drug choice. Stanford University.

Dranove, D., S. Ramanarayanan, and A. Sfekas (2011). Does the market punish aggressive experts? Evidence from cesarean sections. *B E Journal of Economic Analysis & Policy 11*(2).

Epstein, A. J. and S. Nicholson (2009). The formation and evolution of physician treatment styles: An application to cesarean sections. *Journal of Health Economics 28*, 1126–1140.

Finkelstein, A., M. Gentzkow, and H. Williams (2015). Sources of geographic variation in health care: Evidence from patient migration. *The Quarterly Journal of Economics 131*(4), 1681–1726.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas, and E. L. Pinder (2003). The implications of regional variations in medicare spending. *Annals of Internal Medicine 138*(4), 273–298.

Frank, R. G. and T. G. McGuire (2000). Economics and mental health. In M. V. Pauly, T. G. McGuire, and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume 1, Part B, Chapter 16, pp. 893 – 954. Elsevier.

Frank, R. G. and R. J. Zeckhauser (2007). Custom-made versus ready-to-wear treatments: Behavioral propensities in physicians' choices. *Journal of Health Economics 26*(6), 1101 – 1127.

Fuchs, V. (2004, NOV-DEC). More variation in use of care, more flat-of-the-curve medicine. *Health Affairs 23*(6), 104–107.

Gelenberg, A. J., M. P. Freeman, J. C. Markowitz, J. F. Rosenbaum, M. E. Thase, M. H. Trivedi, and R. S. V. Rhoads (2010). *Practice Guideline for the Treatment of Patients with Major Depressive Disorder* (Third ed.). American Psychiatric Association.

Gittins, J., K. Glazebrook, and R. Weber (2011). *Multi-Armed Bandit Allocation Indices*. Wiley and Sons.

Goldstein, D. J., C. Mallinckrodt, Y. Lu, and M. A. Demitrack (2002). Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. *The Journal of Clinical Psychiatry 63*(3), 225–231.

Grove, W., D. H. Zald, B. S. Lebow, B. E. Snitz, and C. Nelson (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment 12*, 19–30.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry 23*(1), 56.

Hicks, J., J. Bishop, K. Sangkuhl, D. Müller, Y. Ji, S. Leckband, J. Leeder, R. Graham, D. Chiulli, A. LLerena, T. Skaar, S. Scott, J. Stingl, T. Klein, K. Caudle, and A. Gaedigk (2015, August). Clinical pharmacogenetics implementation consortium (cpic) guideline for cyp2d6 and cyp2c19 genotypes and dosing of selective serotonin reuptake inhibitors. *Clinical Pharmacology & Therapeutics 98*(2), 127–134.

Hieronymus, F., J. F. Emilsson, S. Nilsson, and E. Eriksson (2015, March). Consistent superiority of selective serotonin reuptake inhibitors over placebo in reducing depressed mood in patients with major depression. *Molecular Psychiatry 21*(4), 523–530.

Kahneman, D. and G. Klein (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist 64*(6), 515–526.

Kantor, E. D., C. D. Rehm, J. S. Haas, A. T. Chan, and E. L. Giovannucci (2015, NOV 3). Trends in prescription drug use among adults in the United States from 1999-2012. *JAMA - Journal of the American Medical Association 314*(17), 1818–1831.

Kasper, S. (1995). Clinical efficacy of mirtazapine: a review of meta-analyses of pooled data. *International Clinical Psychopharmacology 10*, 25–35.

Kendrick, D. A., H. M. Amman, and M. P. Tucci (2014). Chapter 1 - learning about learning in dynamic economic models. In K. Schmedders and K. L. Judd (Eds.), *Handbook of Computational Economics Vol. 3*, Volume 3 of *Handbook of Computational Economics*, pp. 1 – 35. Elsevier.

Kennedy, S. H., R. W. Lam, R. S. McIntyre, S. V. Tourjman, V. Bhat, P. Blier, M. Hasnain, F. Jollant, A. J. Levitt, G. M. MacQueen, S. J. McInerney, D. McIntosh, R. V. Milev, D. J. Müller, S. V. Parikh, N. L. Pearson, A. V. Ravindran, R. Uher, and the CANMAT Depression Work Group (2016, August). Canadian network for mood and anxiety treatments (canmat) 2016 clinical guidelines for the management of adults with major depressive disorder: Section 3. pharmacological treatments. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie 61*(9), 540–560.

Kirsch, I., B. J. Deacon, T. B. Huedo-Medina, A. Scoboria, T. J. Moore, and B. T. Johnson (2008). Initial severity and antidepressant benefits: a meta-analysis of data submitted to the food and drug administration. *PLoS Medicine 5*(2), e45.

Klein, N. and S. Rady (2011). Negatively correlated bandits. *The Review of Economic Studies 78*(2), 693–732.

Lai, T. L. and H. Robbins (1985). Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics 6*(1), 4–22.

Lee, M. D., S. Zhang, M. Munro, and M. Steyvers (2011). Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research 12*(2), 164–174.

Linde, K., L. Kriston, G. Rücker, S. Jamil, I. Schumann, K. Meissner, K. Sigterman, and A. Schneider (2015). Efficacy and acceptability of pharmacological treatments for depressive disorders in primary care: Systematic review and network meta-analysis. *The Annals of Family Medicine 13*(1), 69–79.

Llorca, P.-M., J.-M. Azorin, N. Despiegel, and P. Verpillat (2005). Efficacy of escitalopram in patients with severe depression: a pooled analysis. *International journal of clinical practice 59*(3), 268–275.

Ludwig, J., D. E. Marcotte, and K. Norberg (2009). Anti-depressants and suicide. *Journal of Health Economics 28*(3), 659–676.

Maneeton, N., B. Maneeton, K. Eurviriyanukul, and M. Srisurapanont (2013). Efficacy, tolerability, and acceptability of bupropion for major depressive disorder: a meta-analysis of randomized–controlled trials comparison with venlafaxine. *Drug Design, Development and Therapy 7*, 1053.

Mathys, C., J. Daunizeau, K. J. Friston, and K. E. Stephan (2011). A bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience 5*.

McConnell, K., R. Lindrooth, D. Wholey, T. Maddox, and N. Bloom (2013). Management practices and the quality of care in cardiac units. *JAMA Internal Medicine 173*(8), 684–692.

McCourt, C., J. Weaver, H. Statham, S. Beake, J. Gamble, and D. K. Creedy (2007). Elective cesarean section and decision making: A critical review of the literature. *Birth 34*(1), 65–79.

Meehl, P. E. (1954). *Clinical vs. Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*. University of Minnesota Press.

Molitor, D. (2016, August). The evolution of physician practice styles: Evidence from cardiologist migration. Working Paper 22478, National Bureau of Economic Research.

Raiffa, H. and R. Schlaifer (2000). *Applied Statistical Decision Theory*. Wiley.

Reverdy, P. B., V. Srivastava, and N. E. Leonard (2014). Modeling human decision making in generalized gaussian multiarmed bandits. *Proceedings of the IEEE 102*(4), 544–571.

Roth, A. E. (2018, July). Marketplaces, markets, and market design. *American Economic Review 108*(7), 1609–58.

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

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics 69*, 99–118.

Skinner, J. (2012). Causes and consequences of regional variations in health care. In M. V. Pauly, T. G. McGuire, and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume 2, Chapter 2, pp. 45–49. Elsevier B. V.

Srinivas, N., A. Krause, S. M. Kakade, and M. W. Seeger (2012, May). Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on*

Information Theory 58(5), 3250–3265.

Stern, S. and M. Trajtenberg (1998, December). Empirical implications of physician authority in pharmaceutical decisionmaking. Working Paper 6851, National Bureau of Economic Research.

Theil, H. (1980). *The System-Wide Approach to Microeconomics*. Chicago, Il.: University of Chicago Press.

Tsai, T. C., A. K. Jha, A. A. Gawande, R. S. Huckman, N. Bloom, and R. Sadun (2015, 08). Hospital board and management practices are strongly related to hospital performance on clinical quality metrics. *Health affairs 34*(8), 1304–1311N.

Van Moffaert, M., J. De Wilde, A. Vereecken, M. Dierick, J. Evrard, J. Wilmotte, and J. Mendlewicz (1995). Mirtazapine is more effective than trazodone: a double-blind controlled study in hospitalized patients with major depression. *International Clinical Psychopharmacology*.

Princeton University and NBER*

Columbia University and NBER**