

UNDERSTANDING DOCTOR DECISION MAKING: THE CASE OF DEPRESSION TREATMENT

JANET M CURRIE AND W BENTLEY MACLEOD¹

Abstract

Treatment for depression is complex, requiring decisions that may involve trade-offs between exploiting treatments with the highest expected value or experimenting with treatments with higher possible payoffs. Using patient claims data, we show that among skilled doctors, using a broader portfolio of drugs predicts better patient outcomes, except in cases where doctor's decisions violate loose professional guidelines. We introduce a behavioral model of decision making guided by our empirical observations. The model's novel feature is that the tradeoff between exploitation and experimentation depends on the doctor's diagnostic skill. The model predicts that higher diagnostic skill leads to greater diversity in drug choice and better matching of drugs to patients even among doctors with the same initial beliefs regarding drug effectiveness. Consistent with the finding that guideline violations predict poorer patient outcomes, simulations of the model suggest that increasing the number of possible drug choices can lower performance.

KEYWORDS: decision theory, bounded rationality, expert decision making, learning, human capital.

1. INTRODUCTION

Understanding doctor decision making is central to improving health and reducing health care costs. This paper examines variations in doctor 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 drug treatments. In the U.S., 13% of the adult population had taken an anti-depressant in the past 30 days in 2011/12, up from 6.8% in 1999/2000 (Kantor et al. (2015)). 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 2010.¹ Depression has been blamed for rising suicide rates in the U.S., with suicide ranking as the 10th leading cause of death in 2016.²

Depression is an especially interesting context for studying doctor decision making. First, there are 33 separate anti-depressant drug molecules available over our study period. Since every patient

Princeton University and NBER Email: jcurrie@princeton.edu

Columbia University and NBER Email: bentley.macleod@columbia.edu

1

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, University of Virginia, Dartmouth, Tel Aviv University, University College London, Universidad Carlos III Madrid, New York University, Pompeu Fabra, CEMFI, CEPRA/NBER, and the International Association of Applied Economists. The authors thank Hendrik Jürges, Aviv Nevo and three anonymous referees for helpful comments and Allen Campbell for working with us to help us access the IQVIA data. Xuan Li, Sara Shahanaghi, Haowen (Alice) Wu, Emily Cuddy and Utkarsh Kumar provided outstanding research assistance. We thank the NIA for support under P30-AG024928-14. Princeton University participates in the BCBS Alliance for Health Research. The Blue Cross Blue Shield Association (BCBSA) established the BCBS Alliance for Health Research to engage leading U.S. healthcare researchers in collaborative efforts to use a limited data set drawn from BCBS companies to explore critical health care issues in order to improve the health of Americans. The BCBS Alliance for Health Research provides researchers with use of a secure data portal to access a limited data set from BCBS Axis®, the largest collection of commercial insurance claims, medical professional and cost of care information. BCBSA is an association of independent BCBS companies. The authors are solely responsible for the content of the paper.

¹(See http://www.oecd.org/els/health-systems/Item10_Trends-in-antidepressant-consumption_NIPH.pdf)

²(See <https://www.nimh.nih.gov/health/statistics/suicide.shtml>).

responds differently to each drug, and there is no one drug that dominates all others for all patients, finding the best option necessarily requires experimentation. However, experimentation is costly in the sense of subjecting patients to sub-optimal drug choices, so expanding the drug choice set does not necessarily improve patient welfare within a fixed treatment window. Hence, guidelines that limit choice may be welfare-improving. Second, as Frank and McGuire (2000) observe, the assessment of patient condition is often more difficult in the case of mental illness than for many physical illnesses, suggesting that doctor diagnostic skill plays an important role. We show that there is a link between doctor skill and the payoff to experimentation, which implies that the best choice of treatment for a patient at a point in time depends in part on the doctor's skill level. Third, prices (and "detailing", i.e. marketing of drugs directly to doctors) are relatively unimportant in decision making about anti-depressants today since most anti-depressants are available as generics and many patients face the same cost (a small co-pay) for many drug choices. Anti-depressants belong to the large and interesting category of markets that clear largely without the aid of price signals (Roth (2018)). Hence, the usual economic approach of seeking to set prices "correctly" is unlikely to improve doctor decision making in this setting.

The goal of this paper is to better understand doctor practice style and to ask how decision making is related to patient outcomes. A closely related question is whether patient outcomes can be improved by limiting physician decision-making through the imposition of guidelines. We begin with an exploratory analysis of the relationship between doctor practice style, guidelines, and patient outcomes using novel data formed by merging information on all anti-depressant prescriptions for each doctor to national administrative claims data for hundreds of thousands of patients treated with anti-depressants. One advantage of using claims data (and a contribution of our paper) is that we can examine emergency room (ER) visits and hospitalizations as patient outcomes, rather than relying on suicides as the outcome measure as in previous studies (Berndt et al. (2015), Ludwig et al. (2009)). Suicides are thankfully relatively rare, making it difficult to compare doctors in terms of this patient outcome.

We characterize practice style for anti-depressant prescribing in terms of how dispersed it is (i.e. how many different drugs doctors use), and in terms of whether the doctor violates national and international guidelines for prescribing. Doctors vary widely on these two measures. Our main empirical findings are as follows: 1) Among patients seeing skilled doctors, more dispersed prescribing is associated with better patient outcomes; 2) Among patients seeing less skilled physicians, more dispersed prescribing is not associated with better patient outcomes; and 3) Violating guidelines about drug transitions is associated with worse outcomes in all patients. These results demonstrate that there is significant variation in these two measures of doctor practice style, and that they are related to patient outcomes.

We then turn to theory in an attempt to understand what dispersion in prescribing means in terms of an underlying model of physician behavior. Thompson (1933) recognized that drug choice can be viewed as a statistical decision problem involving estimation of the means and variances of treatments and updating them over time. Since the doctor can only observe one treatment at a time, the problem takes the form of a multi-armed bandit in which the patient is like a "slot machine," and the drug choice each period corresponds to the question of which arm of the machine to pull (Robbins

(1952)). The practitioner does not know the best drug for a particular patient in advance, and learns by trial and error. At any point in time there is a trade-off between choosing the treatment with the highest expected value, and experimenting to learn more about treatments that may be better for a particular individual.

Finding the optimal decision rule for a multi-armed bandit is computationally challenging (Kendrick et al. (2014)). Within computer science, a rapidly developing area of research builds on the adaptive learning model of Lai and Robbins (1985) in order to develop practical solution methods. These approaches are easier to implement than optimal dynamic programming solutions and achieve very good results (see Bubeck and Cesa-Bianchi (2012); Sutton and Barto (2018)). In these models, drug choice depends on both the expected value of a choice and on beliefs about the variance in outcomes for that choice. For example, if there are two drugs A and B, and A has a higher expected value but the variance of B is believed to be higher, then a doctor may choose drug B because there is more to learn regarding its effectiveness. When the doctor chooses B this reduces her uncertainty about B, and depending on the outcome she will either stick to B or go back to A. We adopt this general approach to modeling learning, and assume that doctors update their beliefs using Bayes rule.

The novel feature of our model is that the trade-off between experimentation and choosing the drug with the highest expected value depends on the physician’s diagnostic skill. Intuitively, if a physician is not able to correctly assess a patient’s condition, then there is little to be learned from experimenting and the patient is likely to be better off without it. In contrast, 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 potential gains from experimentation 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. Physicians who are more skilled will 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.

Finding the optimal treatment entails experimentation and experimentation will involve some choices that turn out to have been suboptimal ex poste. 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. Hence, guidelines that rule out some choices may improve patient welfare over a fixed treatment window. Simulations of our model demonstrate exactly this point, consistently indicating that restrictions on physician’s choice sets (guidelines) can be welfare improving over a finite treatment horizon.

Our empirical findings (1 and 2) are consistent with the model’s insight that experimentation is more beneficial when physicians have more skill. The third empirical finding suggests however, that there is a limit to the benefits of experimentation: Violating guidelines is associated with worse patient outcomes even in patients seeing skilled physicians.

The agenda of the paper is as follows: Section 2 provides some background and further discussion of the literature. Section 3 discusses the data, while the empirical results are discussed in section 4. Section 5 introduces the model details, and how one can connect doctor diagnostic skill and learning with observed performance. We conclude with Section 6.

2. BACKGROUND: PREVIOUS WORK ON DOCTOR DECISION MAKING AND PRACTICE STYLE

While differences between patients are obviously an important driver of heterogeneity in treatment, there is increasing evidence that variation in doctor decision making is important. For example, Cutler et al. (2019) document large differences in doctor beliefs about appropriate treatment. Similarly, previous studies have found little evidence that patient demand drives the large differences in C-section rates across U.S. providers (McCourt et al. (2007)). Finkelstein et al. (2015) suggest that at least half of the observed variation in procedure use among the elderly is due to supply-side factors.

Doctor peer effects are one potential source of area-level variation in treatment choices. Models of peer effects such as Chandra and Staiger (2007) imply that the practice styles of doctor peers should converge over time, though the evidence on that front is mixed perhaps because of the difficulty of defining peers (see Molitor (2016), Donohue et al. (2018), Epstein and Nicholson (2009), Dranove et al. (2011) and Chan (2015)). The mixed results suggest that proximity alone is not enough to drive spillovers in practice style, and that variation in practice style is likely to depend upon sub-specialty and detailed features of the environment in which doctors work.

Fear of litigation is another frequently cited reason for variations in doctor decision making. The idea that doctor decisions are shaped primarily by fear of litigation remains popular, even though it has been repeatedly de-bunked. For example, Baicker and Chandra (2005) find no evidence that treatment responds to changes in legal liability, except for some screening procedures.

Our focus is on the prescribing behavior of physicians. Starting with Stern and Trajtenberg (1998), several previous studies use concentration in a doctor's prescribing as a measure of practice style. Frank and Zeckhauser (2007) cite survey data showing that the most prescribed medication for a specific condition was responsible for about 60% of a doctors' prescriptions for that condition while patient demographics had little explanatory power. Berndt et al. (2015) use data on prescriptions of anti-psychotics and show that most doctors have a favorite drug and that on average 66% of their prescriptions are for this drug.

Other studies treat prescription choice as an application of consumer demand theory. Crawford and Shum (2005) study the market for ulcer drugs and estimate a structural model of drug choice. In their model all of the differences in treatment are driven by patient needs or preferences. 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." Dickstein (2015) uses a similar methodology to model how drug choices are affected by differences in patient co-payments 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, co-pays only varied from ~\$10.00 to ~\$20.00 per month, and drugs with a wide range of prices had similar co-pays, suggesting that price differences could only be a partial explanation of drug choice. Neither study allows differences in doctor practice style per se to explain the choice of prescription drugs.

Since we have doctor identifiers we are able to document variation in physician practice style, and connect it to patient outcomes. Differences in prescribing practices have led to calls for stronger

practice guidelines, especially in psychiatry. Meehl (1954), Grove et al. (2000) and Kahneman and Klein (2009) argue that algorithms could do as least as well as a psychiatrist in the treatment of mental illness, and the development of algorithms for treatment is an active area of ongoing research (see Adli et al. (2017)). On the other hand, Frank and Zeckhauser (2007) express concern that guidelines could prevent doctors from providing care that is sufficiently individualized, while Cosgrove et al. (2018) worry that guidelines may be influenced by pharmaceutical companies. 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 doctors is an important one (see Bloom et al. (2015), Tsai et al. (2015) and K et al. (2013)).

The challenge is to provide ways to measure practice style that can be linked to patient outcomes in order to assess doctor performance. Currie and MacLeod (2017) and Currie et al. (2016) develop a framework for studying individual doctor decision making in contexts where doctors make zero/one choices and there is little scope for learning about individual needs over time. They show that there is a great deal of variation in both the responsiveness and aggressiveness of treatment across doctors, and that these characteristics of doctors are fairly stable over time. This paper extends this work to a more complex case where learning is important. In the model we develop below, practice style is still determined by two characteristics of doctors, diagnostic skill (which corresponds to how responsive doctors are to patient condition) and reservation probabilities (which correspond to aggressiveness and provide a summary measure of underlying doctor characteristics such as risk preferences). This paper builds on our previous work by showing how diagnostic skill helps the doctor to learn about the effectiveness of treatment over time when there are many potential treatments and the best choice can only be determined by trial and error. Even given a trial, it may not be obvious whether a particular drug is working or not. Randomized controlled trials of drug treatments for depression can yield very noisy results: In most cases the placebo effect of giving a drug accounts for about 80% of the total measured effect.

Like us, Frank and Zeckhauser (2007) observe that the treatment of depression can be viewed as a bandit problem. Rothschild (1974) shows that in a market setting, trial and error learning leads to a stable equilibrium with significant price heterogeneity. This heterogeneity arises because at some point it is no longer profitable to explore further choices, so that individuals with different consumption histories can settle on different choices. Aghion et al. (1991) study the optimal search strategy with learning and highlight the trade-off between exploitation (continuing with the same choice), or exploration (trying a new choice). They find that in a setting without discounting, parties will eventually learn the true demand functions facing firms. Bergemann and Välimäki (1996) extend these results to the case of endogenous investment into learning.

Cohen et al. (2007) suggests that humans solve this type of problem using an index rule - for each drug choice the physician assigns a value that is a combination of the expected treatment effect and the potential information to be gained by prescribing the drug. In contrast to dynamic programming which works backwards from the potential last period of treatment to the present, forward looking algorithms compute the index using currently available information. Sutton and Barto (2018) provide a review of the literature on such rules beginning with Gittins (1979) and Lai and Robbins (1985).³

³See also Erev and Roth (2014)'s of the earlier literature using reinforcement learning in economics.

One of the most successful algorithms is based on computing the “upper confidence bound” for a given arm (Bubeck and Cesa-Bianchi (2012); Srinivas et al. (2012); Srivastava et al. (2015)). These “UCB algorithms” are used to model human decision making in a variety of contexts, including laboratory experiments with bandit choice (Reverdy et al. (2014)), decision making in complex games (Wu et al. (2018)), and advertising (Schwartz et al. (2017)).

The model we introduce in section 5 uses the idea of an upper confidence bound from this literature to introduce the concept of a “reservation probability,” p , that defines an index q_d for choice d . We are inspired by Simon (1955)’s idea of satisficing, the basis of contemporary labor market models of search. In these models workers search until they receive a wage offer that exceeds their reservation wage. In our model, the probability that the effect of drug d is larger than q_d is p . A doctor’s practice style can be summarized by this reservation probability: The doctor will try new drugs until the probability that a new drug is better than the current drug falls below the reservation probability.

An implication of the model is that optimizing agents operating in the same environment can show a great deal of variation in their choices. In our case the doctor is solving a matching problem which involves trying to find the best treatment for each patient given unobserved patient characteristics. It is important that different patients should be given different treatments. Hence, better doctors will have more dispersed drug choices. To make this idea more precise we introduce a model of drug choice that incorporates diagnostic skill. We then discuss the implications of variations in both diagnostic skill and reservation probabilities for patient outcomes and the within-doctor dispersion of drug choices.

3. DATA

In order to examine the relationship between doctor skill, practice style, and patient outcomes, we use 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. The administrative claims data come from a limited data set made available through a secure data portal and are drawn from Blue Cross Blue Shield (BCBS) Axis®⁴, the largest source of commercial insurance claims data in the U.S. ⁴We 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 data 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 ever had a pharmacy claim, and of these, 723,818 members were ever prescribed antidepressants over the sample period. These members constitute our initial BCBS data sample. For each of these members, we generate a panel of data with a record for each month and year that they appeared in the claims data. 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. We focus on ER visits as an outcome because in the US, patients with mental health crises are uniformly

⁴Accessing insurance claims data often requires extended negotiations with individual insurance carriers, or with government entities. Further information about the BCBS Health of American Initiative, including information about their Axis® data base and contact information is available at: <https://www.bcbs.com/the-health-of-america/about>.

advised to proceed to the nearest ER for assessment. ER visits (and subsequent hospitalizations) for mental health indications are much more numerous than suicides and have a substantial impact on health care costs.⁵

A problem with using the BCBS data to examine practice style is that most doctors see patients with other types of insurance. Hence, measures of practice style constructed using only the BCBS data could omit many of a doctor's other patients. We remedy this problem by measuring dispersion in prescribing using a second data base from IQVIA which comes from their Xponent data base.⁶ 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. We are able to find a match in the IQVIA data for 74.0 percent of the doctors in our BCBS data sample. We also match in doctor characteristics such as specialty from the National Plan and Provider Enumeration System (NPES). We follow Berndt et al. (2015) and limit our analysis to doctors 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 doctor characteristics.

Following Theil (1967) and Shorrocks (1980) we use the Shannon entropy score as our main measure of dispersion in doctor prescribing, though we also show estimates using the share of a doctor's anti-depressant prescriptions that are for their favorite drug. Results are qualitatively similar if we use a Herfindahl index.

To define entropy, let $\mathbf{p} \in \Delta^n = \{\mathbf{p} \in [0, 1]^n \mid \sum_i p_i = 1\}$ be a probability vector. Let $n(\mathbf{p})$ denote the number of entries in the vector.

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

$$\Phi(\mathbf{p}) = \phi(\mathbf{p}) / \log(n) \in [0, 1],$$

where

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

is the Shannon entropy index, k is 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 $p_k = 1$.

Let n_{jt} be the number of patients that a doctor treats in period t , and let n_{djt} be the number of

⁵Our own calculations using hospital records from six states that participate in the Health Care Utilization Project (Arizona, Florida, Kentucky, Maryland, New Jersey, and New York) suggest that in 2014 there were 379.9 ER visits annually per 1,000 individuals. Of these, 50.9, or 13.4 percent listed a mental health diagnosis on the hospital record; of these, 5.1% listed mood disorders as an indication and 4.9% listed anxiety. Anti-depressants are frequently prescribed for both of these conditions.

⁶The IQVIA data is available for purchase to qualified researchers. For further information, contact Allen.Campbell@iqvia.com.

prescriptions of drug $d \in D = \{d^1, \dots, d^m\}$, where m is the number of drugs, in period t . Let

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

be the fraction of patients of doctor j who take drug d in period t . This m -dimensional vector $\mathbf{p}_{jt} \in [0, 1]^m$ is a measure of the doctor’s practice style at time t . It can be summarized using the normalized entropy score defined above. We allow a doctor’s entropy score to vary over time, calculating a separate entropy score for each doctor for each calendar year of data.⁷ We match 2013 BCBS data to 2012 doctor entropy data, and so on. Due to data limitations, both 2015 and 2016 BCBS data are merged to entropy measures for 2014; however, within doctor entropy changes only slowly over time. Once the IQVIA data are matched to the BCBS data, we use the mean entropy score for each doctor over the entire treatment interval with that patient. This decision eliminates small changes in entropy scores that would otherwise occur whenever the patient’s treatment interval happens to span more than one calendar year of data. It means that within-patient changes in entropy are due solely to changes in the prescribing doctor. Since there are 33 different anti-depressant molecules in use over our sample period, the entropy score is computed using $m=33$. However, the top eleven molecules accounted for the vast majority of the of prescriptions in 2014 (see Appendix Table 1), 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.”)

Table (I) provides information about the breakdown of anti-depressant prescriptions and entropy scores across types of providers in the IQVIA data. The vast majority of anti-depressant scripts are written not by psychiatrists but by general practitioners and other doctors who have little specific training in the use of anti-depressants. The model discussed below predicts that providers with less expertise should do less experimentation. Table (I) shows that in keeping with this prediction, psychiatrists have the highest entropy scores, followed by GPs, with other doctors having substantially lower entropy scores. Entropy scores increase with doctor cohort, except among the oldest doctors where they show a decline which is consistent with prime age doctors being the most skilled. A limitation of the version of the IQVIA data that we were able to obtain is that there is no patient identifier. Therefore, we cannot examine experimentation within-patient in these data, and the overall entropy scores in the second panel of Table (I) represent a mix of within patient and across patient prescribing.

⁷Since the entropy score is a non-linear function of treatment choice, one might worry about biased estimates for doctors with few patients. Simulations of the entropy index showed that the small sample bias is sufficiently small that we ignore.

TABLE I
VOLUME AND DISPERSION OF PHYSICIAN PRESCRIBING IN 2013, BY SPECIALTY

	All prescriptions	GPs	Psychiatrists	Other physicians
# 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	0.619	0.619	0.656	0.500
1976-1985	0.623	0.631	0.657	0.499
1986-1995	0.618	0.630	0.649	0.479
1996+	0.608	0.625	0.631	0.445
Average Share of Drugs to Physician's "Favorite" Drug				
<1975	0.257	0.256	0.223	0.364
1976-1985	0.253	0.245	0.218	0.366
1986-1995	0.254	0.242	0.220	0.387
1996+	0.258	0.242	0.229	0.415

Notes: Entropy calculations include only providers with ≥ 12 scripts in the year and are based on m-33 separate drug molecules. Data on prescriptions from IQVIA includes physician identifiers and is merged to American Medical Association files that provide the doctor's medical school graduation year and specialty.

In the bottom panel of the table, we show that among GPs, roughly a quarter of prescriptions are for the favorite drug, compared to about 22% among psychiatrists, and 36 to 42% among other specialties, depending on the cohort. We repeat our main estimates of the correlates of dispersion using this measure.

A second measure of practice style is the extent to which a doctor follows practice guidelines. We 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. They 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)).

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 drug 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)). This list excludes two drugs that together accounted for 17.58% of the market in 2014 (see Appendix Table 1). In addition, all guidelines urge caution in the use of "drug cocktails." At issue is that most combinations have not been evaluated in clinical trials, so the possible drug interactions or side effects are largely unknown.

Diffusion in prescribing and the violation of guidelines are related since following a guideline means

ruling out potential choices. Figure 1 illustrates this relationship by showing the actual distribution of entropy scores compared to the counter-factual distribution that would exist if all doctors adhered to the APA guidelines. In the counter-factual, prescriptions for medications that violate guidelines are distributed over the remaining options in proportion to actual prescription patterns. Figure 1 shows that if the APA guidelines were followed, the right tail of the entropy distribution would be compressed, and the whole distribution would shift to the left. Hence, the APA guidance, while loose, would still be binding on practice styles if it were followed.

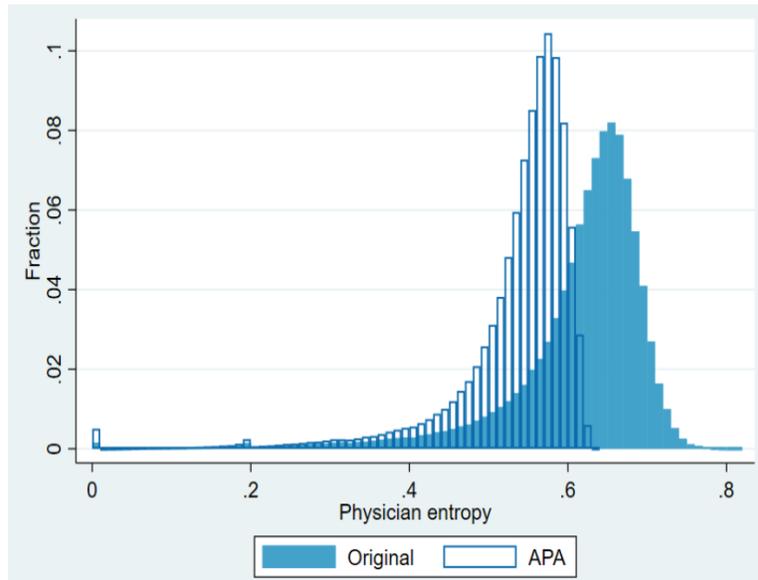


FIGURE 1.— Actual and Counterfactual Distribution of Physician Entropy Scores. This figure shows the distribution of actual entropy scores in 2013 (shaded bars) and the distribution that would have existed if doctors had followed the APA guidelines for prescribing (white bars). We assume prescriptions that are in violation of guidelines would have been distributed in the same way as a doctor’s remaining prescriptions had they followed the guidelines.

Table (II) provides an overview of the BCBS data. The table divides patients into those who ever saw a psychiatrist as an outpatient, and those who did not. We do not divide patients into whether they are currently seeing a psychiatrist or not, because the same patient is more likely to start seeing a psychiatrist if their condition deteriorates. Table (II) shows that most patients in our sample never see a psychiatrist over the 12 to 13 months that we follow them. Patients who ever see a psychiatrist during the window in which we observe them will be seeing more skilled providers on average. On average, patients experience 1.3 changes in entropy scores, which coincide with changes in the doctor treating them. When we look at drug transitions, we have a slightly smaller sample because we lose the first observation when we lag.

TABLE II
SUMMARY OF BCBS PATIENT DATA BY OUTPATIENT PROVIDER

Patient Type:	Ever Saw		Never Saw
	All Patients	Psychiatrist	Psychiatrist
# members	450,802	82,810	367,992
# member-months	5,409,124	1,117,032	4,292,092
# months/member	11.999	13.489	11.664
# months antidepressants/member	8.303	9.499	8.034
# changes in entropy/member	1.331	1.621	1.265
# member-month with drug transitions	4,716,167	976,768	3,739,399
Percent monthly drug transitions from t-2 to t-1 that violate each guideline (as % of row 8)			
UK	0.102	0.117	0.098
Canada	2.406	2.177	2.466
US	3.601	4.623	3.334
Cocktail	4.491	8.917	3.335
Average Monthly Costs (in Jan. 2013 dollars)			
total monthly cost: 50th p'tile	109.17	218.27	86.87
90th p'tile	1,412.10	2,025.83	1,242.10
pharmacy cost: 50th p'tile	23.95	48.38	20.16
90th p'tile	519.57	794.23	447.98
professionals cost: 50th p'tile	0.00	19.75	0.00
90th p'tile	504.15	699.94	451.30
facility cost: 50th p'tile	0.00	0.00	0.00
90th p'tile	108.51	198.42	88.72
99th p'tile	7,828.53	9,767.20	7,232.10
Share with Emergency Room or Hospitalization each Month			
1 if any ER/hospitalization	0.0237	0.0323	0.0214
1 if any ER/hosp. for mental health	0.0097	0.0170	0.0077

Notes: Authors' calculations from the BCBS data. 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. Transitions can be from the same drug to the same drug or from a drug to no drug, for example.

The second panel of Table (II) provides information on how often patients changed medications in a way that violated one of the prescription guidelines. The informal APA guideline is most likely to be violated. Relatively large numbers of patients receive a cocktail of drugs, especially from psychiatrists. To get some understanding of the magnitudes, consider that a patient who is in the data for 11 months, has 10 month-to-month transitions observed, where “transitions” also include staying on the same medication from month to month. Hence, if 4.5 percent of all transitions violate the “no drug cocktails” guideline, the average patient has 0.45 violations over the period that we see them.

The third and fourth panels of Table (II) present data on the costs of care and outcomes. 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 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 monthly costs. These figures represent the actual amount paid to the various providers. Patients saw a psychiatrist at some point have higher costs in every category suggesting that they are sicker on average. We will include patient fixed effects to control for differences in the patient’s average severity level. Comparing models with and without patient fixed effects also provides information about which patients switch to higher skilled physicians, as discussed further below. While costs are important, our primary outcomes are indicators for whether the patient used the ER or was hospitalized with a mental health indication. Overall, approximately one percent of patients had an ER visit or hospitalization with a mental health indication each month.

Although these data are rich, they have some major limitations. The most significant limitation in the IQVIA data is the lack of a patient identifier. Arguably, high diffusion in drug choices within a patient represents a different sort of experimentation than diffusion across patients. High diffusion across patients could even be interpreted as an additional measure of skill since it could mean that a doctor is sorting patients into different categories before prescribing to them. Two limitations of the BCBS data are that we do not follow most patients very long, and that we only see a subset of each doctor’s patients (those who have BCBS coverage). A third limitation of the BCBS data is that many people who are treated with anti-depressant drugs lack any indication about why they are being treated. For example, if a patient is prescribed anti-depressants in the course of their annual physical, or in a followup visit to an obstetrician, the only indication on the claim may be for a general visit. Hence, it is possible that some people in our sample are being treated “off-label” for another condition such as anxiety. Thus, the most accurate description of our sample is that it includes people who are treated with anti-depressants at some point in our sample period.

4. EMPIRICAL FINDINGS

In order to examine the relationship between dispersion in a doctor’s prescribing and patient outcomes, we estimate models of the form:

$$(4.1) \quad Y_{ijt} = a_0 + b_1\phi_{jt-1} + b_2x_i + b_3county_i + b_4y_t + e_{ijt},$$

or alternatively:

$$(4.2) \quad Y_{ijt} = a_i + b_1\phi_{jt-1} + b_2y_t + e_{ijt},$$

where Y is one of the outcomes discussed above, x are the observable fixed 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, subsumes the observable patient characteristics and county fixed effects (since most patients do not move in the short interval that we observe them). By including a patient fixed effect, we control for important 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 entropy to be different for patients who see providers with different average skill levels. The doctor’s entropy

score is measured at $t-1$ for an outcome measured at time t , so that the measure of practice style always precedes the outcome.

In order to examine the relationship between violations of treatment guidelines and outcomes, we estimate models of the following form:

$$(4.3) \quad Y_{ijt} = a_0 + b_1 V_{ijt-1} + b_2 x_i + b_3 \text{county}_i + b_4 y_t + e_{ijt},$$

or alternatively:

$$(4.4) \quad Y_{ijt} = a_i + b_1 V_{ijt-1} + b_2 y_t + e_{ijt},$$

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, 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 diffusion in prescribing and the violation of guidelines are related to patient outcomes. Although all of these patients are being treated in the U.S., we can consider the hypothetical effects of violating guidelines from Canada and the U.K. and thus gain a rough idea of the efficacy of the different guidelines.

Including patient fixed effects offers a powerful way to control for patient heterogeneity in order to isolate the effects of doctor practice style. There may however, still be time-varying, unobserved characteristics of patients that are important. In particular, patients who are getting worse may transition to more skilled doctors (e.g. from general practitioners to psychiatrists). If higher entropy scores are associated with greater skill but sicker patients match with more skilled doctors, then the estimated effects of diffusion in prescribing will be biased towards zero so that our estimates will tend to understate any true positive effects of entropy. All standard errors are clustered on the doctor's ID in order to allow for correlations in treatment within doctors, both for the same patient over time and between patients seeing the same doctor.

Regressions of patient outcomes on provider entropy scores are shown in Table III for the full sample, as well as for the two subsamples defined by whether the patient ever saw 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 row of the first two columns suggests 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 significantly negative relationship with costs. What this tells us is that the sicker patients tend to match with the highest entropy providers, leading to a positive bias in the relationship between costs and entropy in OLS. When that selection is controlled for by looking within patients, we see that more diffuse prescribing is actually associated with better patient outcomes.

TABLE III
 PATIENT OUTCOMES AT T AND PROVIDER ENTROPY AT T-1

	(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
<i>Panel A: All patient observations</i>								
Entropy (t-1)	0.309 (0.034)	-0.085 (0.039)	-0.106 (0.035)	-0.160 (0.038)	-0.004 (0.001)	-0.005 (0.002)	0.008 (0.001)	-0.001 (0.001)
Constant	3.806 (0.033)	5.003 (0.025)	2.510 (0.035)	3.159 (0.025)	0.024 (0.001)	0.030 (0.001)	0.004 (0.001)	0.011 (0.001)
Mean Dep. Variable	3.976	3.976	2.512	2.511	0.024	0.024	0.010	0.010
Adj. R2	0.044	0.433	0.027	0.303	0.004	0.092	0.002	0.073
# Obs. (millions)	5.409	5.409	5.409	5.409	5.409	5.409	5.409	5.409
# Members	450,802	450,802	450,802	450,802	450,802	450,802	450,802	450,802
<i>Panel B: Patients who ever saw a psychiatrist as an outpatient</i>								
Entropy (t-1)	0.040 (0.066)	0.044 (0.073)	-0.214 (0.083)	-0.013 (0.073)	-0.013 (0.003)	-0.013 (0.004)	0.005 (0.002)	-0.010 (0.003)
Constant	4.379 (0.078)	5.500 (0.049)	3.004 (0.090)	3.661 (0.049)	0.035 (0.003)	0.046 (0.003)	0.014 (0.002)	0.024 (0.002)
Mean Dep. Variable	4.557	4.557	3.024	3.024	0.032	0.032	0.017	0.017
Adj. R2	0.059	0.424	0.051	0.353	0.012	0.110	0.008	0.098
# Obs. (millions)	1.117	1.117	1.117	1.117	1.117	1.117	1.117	1.117
# Members	82,801	82,801	82,801	82,801	82,801	82,801	82,801	82,801
<i>Panel C: Patients who never saw a psychiatrist as an outpatient</i>								
Entropy (t-1)	-0.153 (0.034)	-0.180 (0.045)	-0.564 (0.035)	-0.222 (0.045)	-0.008 (0.001)	0.000 (0.002)	0.003 (0.001)	0.003 (0.001)
Constant	3.986 (0.035)	4.907 (0.028)	2.685 (0.037)	3.036 (0.029)	0.026 (0.001)	0.024 (0.001)	0.005 (0.001)	0.006 (0.001)
Mean Dep. Variable	3.825	3.825	2.378	2.378	0.021	0.021	0.008	0.008
Adj. R2	0.048	0.374	0.028	0.282	0.004	0.086	0.002	0.060
# Obs. (millions)	4.292	4.292	4.292	4.292	4.292	4.292	4.292	4.292
# Members	367,992	367,992	367,992	367,992	367,992	367,992	367,992	367,992

Notes: Each column of each panel shows coefficients from a single regression model. These regressions show the relationship between the physician's entropy score at t-1 and patient outcomes at time t. 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.

Since the dependent variable is log costs, the coefficient of -0.085 in column 2 can be interpreted as an elasticity: A one unit change in entropy would lead to a 8.5% decrease in total costs. Entropy varies from around 0.4 to 0.8 as shown in Figure 1, with most mass between 0.5 and 0.75. Hence an entropy measure 0.25 higher is correlated with a 2.1% decrease in costs. The columns with fixed effects show that an increase in entropy is associated with a large decrease in non-drug costs (a 0.25 increase in entropy is estimated to reduce these costs by 4.0%). A small part of this reduction comes from reductions in the probability of ER visits and hospitalizations. An increase of 0.25 in entropy is associated with a 0.001 point decline in the probability of such visits on a baseline of 2.4%.

These overall patterns mask differences in the estimates by patient group. Patients who saw a psychiatrist at some point say more skilled practitioners on average. The patient fixed effects models show the differential effects of entropy net of differences in the patients' average conditions. Panel B shows that among patients who ever saw a psychiatrist, more diffuse prescribing is associated with

a reduction in ER visits and hospitalizations, both overall and specifically for mental health. The corresponding fixed effects models indicate that an increase of 0.25 in the provider entropy score is associated with a reduction of 10.2% in the probability of any ER visit or hospitalization. The probability of a visit specifically for a mental health diagnosis falls by 14.7%. Comparing the OLS and fixed effects estimates of the effect of entropy on the probability of an ER/Hospital visit with a mental health indication, we see that the sign changes from positive to negative. This change indicates that patients who are more likely to go to the ER or be hospitalized are also more likely to switch to a higher entropy provider (recall that psychiatrists have more diffuse prescribing on average). Hence, OLS estimates of the effect of entropy are biased upwards. When this bias is mitigated by including patient fixed effects, we find that among patients seeing more skilled doctors, switching to a higher entropy doctor is associated with better patient outcomes, our empirical finding #1.

The group of patients who never saw a psychiatrist as an outpatient are seeing less skilled providers on average, mainly GPs. The patient fixed effects models show that in this group switching to a higher entropy doctor appears to lower costs, but this is not driven by a significant reduction in ER visits or hospitalizations, suggesting that these high-entropy GPs are just lower cost providers on average. Moreover, although there is no change in the propensity to use the ER or hospital, there is a 9.7% increase in the probability of an ER/hospital visit with a mental health indication. It is possible that this increase reflects a higher likelihood of reporting a mental health problem for a given ER/hospital visit. The overall take away is that in patients seeing this group of relatively less skilled doctors, switching to a higher entropy provider does not improve patient outcomes, our empirical finding #2.

Our story that psychiatrists have both more diffuse prescribing and greater skill in treating mental health patients is consistent with both sets of results. An alternative story is one in which a patient switches doctors following a crisis, but then regresses back towards their mean level of illness. However, if this were the case, then we might expect to see a similar pattern in the two groups of patients. That is, it might be reasonable to assume that patients who experience a crisis are more likely to switch to a doctor willing to prescribe a broader portfolio of drugs even if they do not switch to a psychiatrist, and that regression to the mean would also operate similarly in both groups of patients. But to explain our results, it would have to be the case that regression to the mean was important in one group (those ever seeing psychiatrists) but not in the other.

Table (IV) shows the estimated relationship between patient outcomes in month t , and having had a drug transition from month $t-2$ to $t-1$ that violated treatment guidelines. Transitions that violated the UK guidelines are associated with higher costs and an increase in ER visits or hospitalizations in OLS, but not in patient fixed effects models. This change in significance indicates that patients are more likely to make such transitions when they are sicker. Transitions that violate the Canadian or U.S. treatment guidelines are associated with uniformly worse patient outcomes including higher costs, and more ER visits and hospitalizations with mental health indications. Panel A, which shows estimates for all patients, indicates that the coefficient estimates are reduced when patient fixed effects are included in the model, showing that doctors are more likely to violate guidelines or to prescribe drug cocktails for sicker patients. But they remain statistically significant: Even within patient, drug transitions that violate guidelines are associated with increases in ER visits and

hospitalizations with mental health indications. The fixed effects estimates suggest that violations of the U.S. guidelines are associated with a 28.2% increase in total costs, an increase in the probability of any ER visit or hospitalization of 13.0%, and an increase in the probability of an ER visit or hospitalization with a mental health indication of 20.0%.

Prescribing drug cocktails is also associated with increases in ER visits and hospitalizations, as well as with a 51.4% increase in total costs. Some of the higher cost is mechanical in the sense that taking more drugs will usually involve higher costs than taking less drugs. But column (4) shows that total non-drug costs also rise by 36.3% which may reflect the 20% higher probability of having an ER visit or hospitalization with a mental health indication.

The next two panels of Table (IV) examine the relationship between violations of the three different treatment guidelines, prescribing drug cocktails, and outcomes in the two subsamples of patients defined by whether or not they ever saw a psychiatrist as an outpatient. The relationships are quite similar across the two groups indicating that violations of treatment guidelines are associated with bad outcomes whether they are prescribed by more or less skilled practitioners, which is our empirical result #3. 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. In the patient fixed effects models, a violation of U.S. guidelines is associated with an increase in the probability of any ER visit or hospitalization of 15.6%, and with a 17.6% increase in the probability of an ER visit or hospitalization for a mental health condition. Hence, patients who experience a drug transition that violates treatment guidelines have a significantly higher probability of a subsequent ER visit or hospitalization, and much of this increase is accounted for by visits with mental health indications.

TABLE IV
 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
<i>Panel A: All patient observations</i>								
Violation UK Guidelines	0.564 (0.037)	0.214 (0.034)	0.703 (0.045)	0.216 (0.041)	0.012 (0.003)	0.003 (0.003)	0.006 (0.002)	-0.001 (0.002)
Violation US Guidelines	0.403 (0.008)	0.282 (0.007)	0.519 (0.009)	0.291 (0.008)	0.010 (0.0005)	0.003 (0.0005)	0.004 (0.0003)	0.002 (0.0003)
Violation Can. Guidelines	0.388 (0.009)	0.488 (0.008)	0.423 (0.010)	0.416 (0.009)	0.007 (0.001)	0.004 (0.001)	0.003 (0.0003)	0.002 (0.0003)
Cocktail	1.009 (0.011)	0.514 (0.009)	0.800 (0.013)	0.363 (0.010)	0.017 (0.001)	0.005 (0.001)	0.011 (0.0003)	0.002 (0.0004)
Mean Dep. Variable	4.010	4.010	2.514	2.514	0.023	0.023	0.010	0.010
Adj. R2	0.054	0.389	0.034	0.308	0.005	0.093	0.003	0.074
# Obs. (millions)	4.716	4.716	4.716	4.716	4.716	4.716	4.716	4.716
# Members	450,802	450,802	450,802	450,802	450,802	450,802	450,802	450,802
<i>Panel B: Patients who ever saw a psychiatrist as an outpatient</i>								
Violation UK Guidelines	0.444 (0.071)	0.249 (0.066)	0.726 (0.090)	0.315 (0.078)	0.012 (0.007)	-0.001 (0.006)	0.010 (0.005)	-0.001 (0.005)
Violation US Guidelines	0.256 (0.016)	0.241 (0.013)	0.378 (0.018)	0.271 (0.015)	0.011 (0.001)	0.005 (0.001)	0.005 (0.001)	0.003 (0.001)
Violation Can. Guidelines	0.498 (0.020)	0.513 (0.017)	0.486 (0.023)	0.416 (0.020)	0.008 (0.001)	0.002 (0.001)	0.005 (0.001)	0.002 (0.001)
Cocktail	0.774 (0.017)	0.581 (0.014)	0.638 (0.022)	0.444 (0.015)	0.016 (0.001)	0.005 (0.001)	0.012 (0.001)	0.002 (0.001)
Mean Dep. Variable	4.573	4.573	3.017	3.017	0.032	0.032	0.017	0.017
Adj. R2	0.070	0.391	0.059	0.362	0.012	0.111	0.009	0.101
# Obs.	976768	976768	976768	976768	976768	976768	976768	976768
# Members	82,810	82,810	82,810	82,810	82,810	82,810	82,810	82,810
<i>Panel C: Patients who never saw a psychiatrist as an outpatient</i>								
Violation UK Guidelines	0.551 (0.042)	0.200 (0.039)	0.647 (0.051)	0.183 (0.047)	0.011 (0.003)	0.004 (0.003)	0.005 (0.002)	0.000 (0.002)
Violation US Guidelines	0.387 (0.009)	0.298 (0.008)	0.504 (0.010)	0.299 (0.009)	0.009 (0.001)	0.003 (0.001)	0.003 (0.0003)	0.001 (0.0003)
Violation Can. Guidelines	0.372 (0.010)	0.483 (0.009)	0.416 (0.011)	0.418 (0.010)	0.007 (0.001)	0.005 (0.001)	0.003 (0.0003)	0.002 (0.0004)
Cocktail	0.877 (0.014)	0.468 (0.012)	0.640 (0.016)	0.307 (0.013)	0.014 (0.001)	0.004 (0.001)	0.007 (0.0004)	0.002 (0.0004)
Mean Dep. Variable	3.863	3.863	2.383	2.383	0.021	0.021	0.008	0.008
Adj. R2	0.055	0.380	0.033	0.286	0.005	0.086	0.002	0.060
# Obs. (millions)	3.739	3.739	3.739	3.739	3.739	3.739	3.739	3.739
# Members	367,992	367,992	367,992	367,992	367,992	367,992	367,992	367,992

Notes: Each column of each panel represents a single regression model. These regressions show the relationship between a transition (between t-2 and t-1) that violated one of the guidelines, and the patient's outcomes at time t. 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.

In summary, the estimates in Table (III) indicate that seeing a doctor with more diffuse prescribing

is associated with better patient outcomes among patients who see more skilled practitioners, but not among patients who see less skilled practitioners. Table (IV) suggests, however, that there may be a limit on the extent to which experimentation is beneficial even among skilled practitioners, and that therefore there may be a useful role for guidelines that restrict some prescribing practices.

5. UNDERSTANDING VARIATION IN DOCTOR PRACTICE STYLE

The empirical evidence shows that diffusion in prescribing and the violation of guidelines are two dimensions of doctor practice style that show a great deal of variation in the data. Patients are better off with doctors who do not violate treatment guidelines and with psychiatrists who use a greater variety of drugs in treatment. In this section we outline a simple model of doctor treatment choice that can help us understand these facts. In the first sub-section we illustrate the main ideas for a two drug, two period example where we can explicitly compute the Bayesian optimal choice. The example shows how increasing the variance of a choice increases the return from experimentation in period 1. We show that given the preferences and skill of the doctor, there exists a reservation probability that implements the optimal choice. Moreover, the reservation probability increases with the doctor's skill and decreases with the discount rate so that doctors who put more weight on long term outcomes have a correspondingly higher reservation probability.

Currie and MacLeod (2020) introduce the notion of a reservation probability and show that in the multi-drug case the reservation probability rule converges to the optimal long run treatment at the same rate as the optimal choice rule. Moreover, higher skilled doctors converge towards the optimum faster and exhibit both more experimentation and more diffuse prescribing in the short run. However, the short run properties of the rule depend on small sample statistics, and are difficult to understand analytically. The most efficient way to try to understand how the algorithm behaves in the short run is through simulation. Hence, in the second subsection below we use data from randomized controlled trials to calibrate the reservation probability model. We show that doctors with low skill have better outcomes when they have a relatively low reservation probability, while higher skilled doctors achieve better outcomes with higher reservation probabilities. In the long run, experimentation can lead to better outcomes, but since experimentation can involve exposing patients to sub-optimal treatments, it also leads to short-term costs. Finally, we observe that restricting choice via guidelines can enhance the performance of all doctors when patients are being seen over shorter treatment intervals. To keep the analysis as clear as possible, we suppose that there is no risk aversion and that doctors and patients have the same preferences over outcomes. Since we use a normal-linear framework the analysis can be extended to deal with risk aversion, and such an extension does not affect the main results.

5.1. *A Two Drug, Two Period Example of Drug Choice*

Suppose that doctors follow a procedure that is consistent with optimal choice, but that observed behavior may deviate depending on doctor characteristics, especially skill. Our approach to thinking about the problem follows from Simon (1955). Simon's shows that a reservation wage (or in his terminology, a "satisfactory" wage) provides a parsimonious description of behavior. One may not

observe worker discount rates or search costs, but knowing the reservation wage enables us to predict a worker's choice. Our context is similar: Since the outcome from drug treatment is uncertain, a single observation is not sufficient to determine the best choice of drug. The doctor must use repeated trials of different drugs in order to find the best choice. At each decision point, there is a trade-off between sticking to a treatment or experimenting with other treatments that may have higher benefit. In the multi-armed bandit problem this tradeoff is framed as the choice between exploitation - continuing with the same arm - or exploration - trying a new arm.

The essential building block of our model is a simple index rule for assigning value to each choice. Suppose that the doctor is deciding between two drugs, $d \in \{A, B\}$, where the doctor's beliefs regarding the effect of these drugs are characterized by a Normal distribution: $e_d \sim N(\mu_{td}, \sigma_{td}^2)$, where μ_{td} is the *ex ante* mean effect of the drug d in period t and σ_{td}^2 is the doctor's belief regarding the variance of the effect for a patient in this period.

Now suppose that doctor j 's taste for experimentation at date t can be summarized by a *reservation probability* $p_{jt} \in [0.5, 1]$. The index rule we use to assign a value to drug $d \in \{A, B\}$ is:

$$(5.1) \quad q_{jtd} = \mu_{td} + F^{-1}(p_{jt}) \sigma_{td},$$

where $F(\cdot)$ is the cumulative distribution function for the standard Normal distribution. Since $F^{-1}(p_{jt})$ is increasing in p_{jt} then a higher reservation probability leads to a higher weight on the variance. Since $e_{td} \sim N(\mu_{td}, \sigma_{td}^2)$, then in period t for $d \in \{A, B\}$:

$$(5.2) \quad Pr[e_{td} \leq q_{jtd}] = F\left(\frac{q_{jtd} - \mu_{td}}{\sigma_{td}}\right) = p_{jt}.$$

Observe that when the reservation probability is 50% then $F^{-1}(0.5) = 0$. In that case:

$$Pr[e_{td} \leq q_{jtd} = \mu_{td}] = 0.5.$$

In other words in this case the index is simply the estimated effect of treatment.

Suppose the mean effect of drug A is greater than drug B , but the variance of B is higher than the variance of A . When the reservation probability is 50% the doctor behaves like a short term optimizer and selects the arm with the highest expected effect. Hence drug A would be chosen. Since the variance of treatment B is higher than that of treatment A , from (5.1) it follows that for some some probability $\bar{p}_{AB} > 0.5$ we have for all $p_{jt} > \bar{p}_{AB}$:

$$q_{jtB} > q_{jtA}.$$

These possibilities are illustrated in Figure (2). When the reservation probability is 0.5 the index is equal to the mean of the treatment effect, $q_{jtd}^{0.5} = \mu_{jtd}$, $d \in \{A, B\}$, and $q_{jtA}^{0.5} > q_{jtB}^{0.5}$. One "behavioral" interpretation of this rule is that if the probability that drug B is better than drug A is less than 50%, then the doctor sticks with drug A . When the reservation probability is 95%, then $q_{jtB}^{0.95} > q_{jtA}^{0.95}$, and the doctor will try drug B as long as there is a 5% or better chance that it will have a better effect than drug A . We can show that this rule corresponds to an *optimal* choice for the two period

decision.

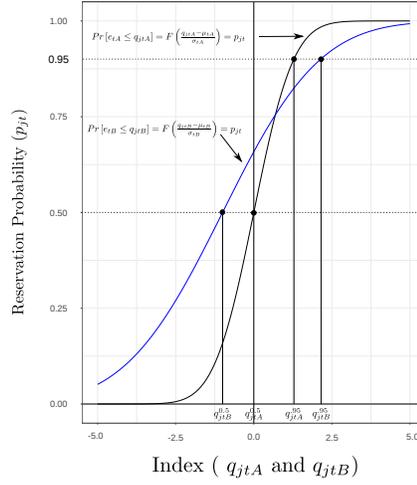


FIGURE 2.— Effect of Reservation Probability on Drug Choice

Suppose doctor j treats patient i . The doctor has two drugs available $d \in \{A, B\}$, with independent effects over two periods, $t \in \{0, 1\}$. Let $\gamma_{i0} \in \mathfrak{R}$ denote the patient's true condition, where $\gamma_{i0} < 0$ indicates a severe depression that should be treated. A patient taking drug d has condition $\gamma_{id} = \gamma_{i0} + e_{id} \in \mathfrak{R}$, where e_{id} is the *treatment effect* of the drug.

The doctor meets the patient twice. In period 1 the doctor observes a signal of patient condition, y_{ij1} . The doctor then prescribes a drug $d_1 \in \{A, B\}$. The patient takes the drug and realizes condition $\gamma_{i1d_1} = \gamma_{i0} + e_{i1d_1}$ at the end of period 1. At the beginning of period 2 the doctor meets again with the patient, observes her condition, y_{ij2} , and prescribes a drug $d_2 \in \{A, B\}$. The patient's condition is given by: $\gamma_{i2d_2} = \gamma_{i0} + e_{i2d_2}$.

We assume that the doctor's only goal is to improve the patient's condition. The doctor's payoff is given by:

$$\begin{aligned} U_j(d_1, d_2) &= E \{(1 - \zeta_j) \gamma_{i1d_1} + \zeta_j \gamma_{i2d_2}\}, \\ &= E \{\gamma_{i0}\} + E \{(1 - \zeta_j) e_{i1d_1} + \zeta_j e_{i2d_2}\} \end{aligned}$$

where $\zeta_j \in (0, 1)$ is the weight placed on payoffs in periods 1 and 2, and d_t is the drug prescribed in period $t \in \{1, 2\}$. When ζ_j approaches zero then the doctor only cares about outcomes in the first period. When ζ approaches 1 then the doctor is concerned with the long run well being of the patient, and may use the first period only to learn more about which drug is most effective for the patient. A doctor might care more about outcomes in the first period, if, for example, they are afraid that the patient will not continue to comply with a medication regime in future.

Given that the initial condition γ_0 of the patient is not affected by the doctor's choice in this model, without loss of generality we can set

$$(5.3) \quad U_j(d_1, d_2) = E \{(1 - \zeta_j) e_{i1d_1} + \zeta_j e_{i2d_2}\}.$$

Thus the doctor's goal is to find the drug that has the largest positive treatment effect.

5.1.1. Beliefs

Doctor choice is modeled as a Bayesian decision problem. The Bayesian approach requires the explicit modeling of beliefs for all parameters that are relevant for a decision:

1. *Patient Condition.* The doctor is assumed to have *ex ante* beliefs regarding the condition of each new patient. Let prior beliefs regarding patient condition given by:

$$(5.4) \quad \gamma_i \sim N\left(\mu_{j0}, \sigma_{j0}^2 = \frac{1}{\rho_{j0}}\right).$$

The *ex ante* precision of this belief is given by $\rho_{j0} > 0$, the inverse of the variance of the distribution. We use precision and variance interchangeably, depending upon which is more convenient: $\rho_{index} = \frac{1}{\sigma_{index}^2}$. On average, drugs that have been approved for treatment of depression are more effective than a placebo but there is a great deal of variance in the effect of drugs.

2. *Drug Efficacy.* The *ex ante* beliefs regarding the treatment effect of drug $d \in \{A, B\}$ for patient i by doctor j is given by:

$$(5.5) \quad e_{id} \sim N\left(\mu_{j0d}, \frac{1}{\rho_{j0d}}\right),$$

where $\mu_{j0d} > 0$ is the mean effect, and ρ_{j0d} is the precision of the doctor's beliefs regarding the benefit of drug d before meeting the patient. After treatment with d the patient has condition $\gamma_{id} = \gamma_{i0} + e_{id}$, where e_{id} is the true, unobserved, realized effect of drug d . The true effect is revealed over time during the course of treatment.

3. *Diagnostic skill.* The new ingredient we introduce is the ability of the doctor to observe a patient's underlying condition. For each patient i , there is a true effect of drug d given by e_{id} . The true effect is not observed by the doctor, but inferred through observation of the patient. At the beginning of period 1 the patient arrives and the doctor observes patient condition, y_{ij1} :

$$(5.6) \quad y_{ij1} = \gamma_i + \epsilon_{ij1},$$

where the error term satisfies $\epsilon_{ij1} \sim N\left(0, \frac{1}{2\xi_j}\right)$ (the factor 2 is included to simplify the subsequent expressions in this two drug case). Here we follow Currie and MacLeod (2017), and view the precision ρ_j as a measure of a doctor's diagnostic skill - the extent to which they are able to assess a patient's true underlying condition. A higher ξ_j corresponds to a more skilled doctor. Similarly, at the beginning of period 2, the patient again visits the doctor, who receives further information:

$$y_{ij2} = \gamma_i + e_{id_1} + \epsilon_{ij2},$$

where ϵ_{ij2} has the same distribution as ϵ_{ij1} .

5.1.2. *Decisions*

Given these beliefs and the doctor's skill level, the doctor then makes choices to maximize her payoff (5.3). In this simple two period model it is assumed $\mu_{j0d} > 0$, and hence it is always optimal to prescribe a drug. Assume that *ex ante* drug A is preferred to drug B ($\mu_{j1A} \geq \mu_{j1B} > 0$), and hence in the absence of new information the doctor (and patient) prefer drug A over drug B , and prefer B over no treatment.

At the beginning of period 1 the doctor's information is given by $I_1 = \{y_{ij1}\} \in \mathcal{I}_1$, the diagnosis at the end of the first meeting. Since the outcome of the interview provides no information about any of the drug treatments, there is no updating of beliefs regarding drug effects. Hence, beliefs in period 1 regarding the effects of the drugs remain unchanged, and for $d \in \{A, B\}$:

$$\mu_{ij1d} = \mu_{j0d},$$

$$\rho_{ij1d} = \rho_{j0d}.$$

At the beginning of period 2 the doctor's information includes the previous observation, the drug chosen, and a new observation of patient condition, $I_2 = \{y_{ij1}, d_1, y_{ij2}\} \in \mathcal{I}_2$. Given this information, the doctor's decisions can be summarized by the function $\vec{\delta}_j = \{\delta_{j1}, \delta_{j2}\}$, where the choice each period maps information to a drug choice:

$$\delta_{j1} : \mathcal{I}_1 \rightarrow \{A, B\},$$

$$\delta_{j2} : \mathcal{I}_2 \rightarrow \{A, B\}.$$

The Bayesian optimal decision problem is solved by backwards induction. One first solves for the optimal choice in period 2 for any possible information set, and then solves for the optimal choice in period 1 given its impact on period 2 payoffs.

Since $\zeta_j > 0$, the goal of the doctor in period two is to choose a drug that maximizes patient well being in period 2:

$$\delta_{j2}^*(I_2) = \operatorname{argmax}_{d_2 \in \{A, B\}} E \{e_{id_2} | I_2\}.$$

Consider first the case in which the doctor chooses $d_1 = A$, and let $I_2^A = \{y_{ij1}, d_1 = A, y_{ij2}\}$. We assume that the effects of the drugs are uncorrelated, and hence observing the effect of drug A has no effect on the doctor's beliefs about drug B and:

$$E \{e_{iB} | I_2^A\} = \mu_{ij2B} = \mu_{ij1B} = \mu_{j0B} > 0.$$

The doctor uses the outcome from the first period of treatment to update her assessment of the success of treatment with drug A . The doctor cannot directly observe true patient condition γ_{i0} , but can see whether the patient's condition improves between periods 2 and period 1:

$$y_{ij2} - y_{ij1} = e_{iA} + \epsilon_{ij2} - \epsilon_{ij1},$$

where the precision/quality of the signal ($y_{ij2} - y_{ij1}$) is $\xi_j = \frac{1}{\frac{1}{2\xi_j} + \frac{1}{2\xi_j}}$.

From this result it follows using the optimal updating rule for Normally distributed random variables that:

$$(5.7) \quad \begin{aligned} \mu_{j2A} &= E\{e_{iA}|I_2^A\} = \frac{\rho_{j1A}\mu_{j1A} + \xi_j(y_{j2} - y_{j1})}{\rho_{j1A} + \xi_j}, \\ &= \mu_{j1A} + \Delta_{ij2A}, \end{aligned}$$

where

$$\Delta_{ij2A} = \frac{\xi_j}{\rho_{j1A} + \xi_j} (y_{j2} - y_{j1} - \mu_{j1A}) \sim N(0, \hat{\sigma}_{j2A}^2),$$

is the new information regarding the effect of drug A . *Ex ante* this information has mean zero and variance:

$$(5.8) \quad \begin{aligned} \hat{\sigma}_{j2A}^2 &= \left(\frac{\xi_j}{\rho_{j1A} + \xi_j}\right)^2 (\sigma_{j1A}^2 + 2\sigma_j^2), \\ &= \left(\frac{\xi_j}{\rho_{j1A} + \xi_j}\right)^2 \left(\frac{1}{\rho_{j1A}} + \frac{1}{\xi_j}\right), \\ &= \frac{\xi_j}{\rho_{j1A}} \left(\frac{1}{\rho_{j1A} + \xi_j}\right), \\ &= \frac{1}{\frac{\rho_{j1A}^2}{\xi_j} + \rho_{j1A}}. \end{aligned}$$

The general formula for updating is given in the appendix. See also DeGroot (1972) for an explicit derivation of the one dimensional case used here.

Notice that the variance of the new information regarding the effect of drug A satisfies $\hat{\sigma}_{j2A}^2 \in [0, 1]$. The variance is increasing with diagnostic skill and decreasing with the precision of the doctor's beliefs regarding the effectiveness of drug A . When the doctor has no diagnostic skill, $\xi_j = 0$ and observing the outcome from period 1 provides no information so $\hat{\sigma}_{j1A}^2 = 0$. The variance of beliefs regarding the effect of the drug also remains unchanged, $\rho_{j2A} = \rho_{j1A}$. In contrast, a skilled doctor ($\xi_j > 0$) learns about the effect of a drug by observing the patient and updates her beliefs regarding the precision of her estimate of e_{id} :

$$\rho_{j2A} = \rho_{j1A} + \xi_j = \frac{1}{\text{var}\{e_{iA}|I_2^A\}} = \frac{1}{\sigma_{ij2A}^2}.$$

Period 2 is the last period of treatment, and hence the doctor prescribes the drug that she believes will provide the largest expected effect. She continues with drug A if and only if:

$$(5.9) \quad \begin{aligned} E\{e_{iA}|I_2^A\} &> E\{e_{iB}|I_2^A\}, \\ \mu_{j1A} + \Delta_{ij2A} &= \mu_{j2A} \geq \mu_{j2B} = \mu_{j1B}. \end{aligned}$$

This result determines the optimal choice at period 2 given that drug A was prescribed in period 1. In order to decide whether or not to try drug A , the doctor has to consider the expected benefit of prescribing drug A in period 1 on the improvement in period 2 decision making. Letting d_2^* denote the optimal choice in period 2, the expected benefit computed in period 1 is:

$$\begin{aligned} E \{e_{id_2^*} | y_{ij1}, d_1 = A\} &= E \{max \{\mu_{j1A} + \Delta_{ij2A}, \mu_{j1B}\}\} \\ &= \mu_{j1A} + E \{max \{\Delta_{ij2A}, \mu_{j1B} - \mu_{j1A}\}\} \\ &= \mu_{j1A} + \hat{\sigma}_{j2A} V \left(\frac{\mu_{j1B} - \mu_{j1A}}{\hat{\sigma}_{j2A}} \right), \end{aligned}$$

where $\hat{\sigma}_{j2A}$ is given by (5.8) and:

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

where $\gamma \sim N(0, 1)$ is the standard Normal distribution, and $L(x)$ is the unit-normal linear loss function $L(x)$:

$$L(x) = E \{max \{x, \gamma\}\} - x = (1 - F(x)) (\phi(x) - x) > 0.$$

$F(x)$ is the cumulative Normal distribution function and $\phi(x) = E \{\gamma | \gamma \geq x\}$ is the expected value of a lower truncated Normal distribution. Raiffa and Schlaifer (2000) provide an extensive discussion of the value of information. See their appendix for a table of values for the Normal unit loss function. This function is a measure of the value of information, where x is the expected return from a choice, and $L(x)$ is the increase in value due to having an option to take up values.

The term

$$V_{j1A} \equiv \hat{\sigma}_{j2A} V \left(\frac{\mu_{j1B} - \mu_{j1A}}{\hat{\sigma}_{j2A}} \right)$$

represents the option value from choosing drug A in period 1. In period 2 one can continue with drug A and get μ_{j1A} , or change to drug B . The payoff from choosing drug A in period 1 given optimal decision making in period 2 is:

$$(5.10) \quad U(d_1 = A, d_{2A}^*) = (1 - \zeta_j) \mu_{j1A} + \zeta_j (\mu_{j1A} + V_{j1A}) = \mu_{j1A} + \zeta_j V_{j1A}.$$

There is a similar expression for drug B where d_{2B}^* represents the optimal choice in period 2 given that drug B was chosen in period 1:

$$(5.11) \quad U(d_1 = B, d_{2B}^*) = (1 - \zeta_j) \mu_{j1B} + \zeta_j (\mu_{j1B} + V_{j1B}) = \mu_{j1B} + \zeta_j V_{j1B}.$$

Putting this together we have:

PROPOSITION 2 *A rational Bayesian doctor chooses drug A over drug B in period 1 if and only if:*

$$(5.12) \quad \mu_{j1A} + \zeta_j V_{j1A} \geq \mu_{j1B} + \zeta_j V_{j1B},$$

an equality that is satisfied to the first order if and only if:

$$(5.13) \quad \mu_{j1A} + \tau_j \hat{\sigma}_{j2A} \geq \mu_{j1B} + \tau_j \hat{\sigma}_{j2B},$$

where

$$\tau_j = \frac{\zeta_j}{(1 - \zeta_j)} L(0) \simeq 0.4 \frac{\zeta_j}{(1 - \zeta_j)}.$$

Thus, to the first order, the optimal choice in period 1 is consistent with a reservation probability that is given by:

$$p_j^* = F(\tau_j) = F\left(\frac{\zeta_j}{(1 - \zeta_j)} L(0)\right) \in [0.5, 1].$$

This rule has two interesting features. First, when the doctor cares only about maximizing the patient's current condition ($\zeta_j = 0$), then the reservation probability is 50% and choice is determined by the static optimal choice rule - she chooses the drug with the highest expected value. When the doctor cares only about the long run, then $\zeta_j \rightarrow 1$ and $p_j^* \rightarrow 1$, and the choice is completely determined by the variance of the new information, $\hat{\sigma}_{j2D}^2$: The physician chooses the highest variance treatment in order to acquire the most information. Given that preferences are not observed, the reservation probability rule is equivalent to the optimal rule for some value of p_j^* .

The option value from choosing drug d , V_{j1d} , is a non-linear function of the mean effects for both drugs. This implies that the Gittins (1979) rule, in which the characteristics of the arms are additively separable, is not optimal for finite horizon problems as Gittins (1979) points out. See also Gittins et al. (2011) for an extensive discussion of the issue. Given that in practice choice is over a finite horizon, using a Gittins type index rule to model optimal choice is necessarily an approximation.

In this model, there is a clear link between diagnostic skill and experimentation. In particular, doctor diagnostic skill increases the value of information from experimentation. Expression (5.8) implies that increasing the doctor's diagnostic skill, ξ_j , increases the return to experimentation:

$$\frac{\partial \hat{\sigma}_{j2d}}{\partial \xi_j} > 0, d \in \{A, B\}.$$

If $\xi_j = 0$, then $\hat{\sigma}_{j2d}^2 = 0$ for $d \in \{A, B\}$, and the doctor will choose drug A for both periods because it has the highest *ex ante* value (since we have assumed $\mu_{j0A} > \mu_{j0B}$). Increasing doctor skill, ξ_j , increases the return from experimentation, and hence the set of situations where the doctor will choose B over A , even though A is known to be the treatment with the highest expected value *ex ante*. Thus, this example illustrates that a better doctor could choose a drug with a lower expected efficacy *ex ante*.

5.2. *Simulating Practice Style and Satisficing Choices*

Although the BCBS data are very rich, they will not allow us to fully explore the implications of our model for patient welfare. In particular, we typically do not follow patients in the data very long and so we cannot take account of what may be lengthy patient histories to directly measure time paths of drug switching. While we follow doctors for longer than patients, we cannot see how they might be grouping patients into types, so we also cannot implement the model by examining how doctors sequentially treat similar patients. In practice however, many treatment intervals are relatively short, suggesting that it would be useful to understand finite sample features of our models. However, in the finite sample case it is difficult to obtain meaningful closed form solutions. Hence, the short-run properties of the model are best assessed with simulations.

We conduct a simulation exercise based upon an extension of the two drug problem to an arbitrary number of drugs developed in Currie and MacLeod (2020). The simulation illustrates the way that doctor skill and reservation probabilities can affect provider behavior and patient utility over time, and we examine how both dispersion in prescribing and patient outcomes are affected by guidelines. An advantage of simulations is that we can fix the distribution of patients to be exactly the same for each doctor. Even though patients are *ex ante* identical, the simulation results illustrate that whether a patient prefers a doctor to experiment or not will depend on the diagnostic skill of the doctor, so there cannot be a unique optimal treatment protocol that applies to all patients.

Our simulations consider a 2x3 experiment with 6 doctor types. The doctor is either an expected utility maximizer with reservation probability of 50%, or an experimental type with a reservation probability of 95%. We also allow doctors to belong to low, medium, and high-skill groups in terms of diagnostic skill. Let $\xi_j \in \{10.0, 1.0, 0.1\}$, measure the accuracy with which the doctor assesses the patient's condition. Here $\xi_j = 10$ denotes high skill (H), and $\xi_j = 0.1$ denotes low skill (L). Suppose that treatment occurs monthly over a period of 36 months.

In our simulations we assume that doctor priors are given by data about the mean and variance of the effects of drugs from clinical trials. This implies that all doctors are equally well informed and hold similar beliefs, which is not likely to be true in practice but is a useful counter-factual to consider. Clinical trials for anti-depressants randomly assign depressed patients to treatment and control groups. 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 often measured using the Hamilton 17 score (See Hamilton (1960)) in which a score of $h > 7$ means that the person is depressed. We normalize this scale so that positive values indicate that the person is not depressed, and negative values indicate depression: The scale is given by $\log(7) - \log(h_i)$, where h_i is the patient's Hamilton 17 score. Meta-analysis of drug trials appear regularly in the medical literature so that in principal, all professionals have access to current results. See for example Linde et al. (2015) and Cipriani et al. (2016). The clinical trials data we use are described further in the Appendix. The assumptions we make about drug efficacy in order to model doctor beliefs are briefly summarized in Table V. The table details the efficacy of each of the top 11 anti-depressant drugs (ranked by market share in 2014) in terms of their effects on the improvement in the Hamilton 17 score. The placebo effect is quite large. In the case of Sertraline (the generic for Zoloft), which is 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% of a drug’s effect is one factor that can help explain why it is hard to find the most effective treatment. To address this large placebo effect, we renormalize the effect of treatment by subtracting the placebo effect. We also allow the effect of “no drug” to be correlated with each drug within patient.

TABLE V
EFFECT OF ANTI-DEPRESSION DRUGS ON HAMILTON 17 SCORE FOR DEPRESSION SEVERITY FROM CLINICAL TRIALS (MEAN SCORE BEFORE TREATMENT IS 25.2)

Active Ingredient	Brand Name(S)	Drug	Share 2014,	Reduction	Deviation
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	Olepto	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: Many anti-depressant drug trials use the Hamilton 17 score for assessing the severity of depression because it is a continuous measure. The mean depression reduction and the standard deviations presented here are culled from meta-analyses of drug effects as described further in an unpublished appendix that is included in the NBER working paper version of this paper. The market shares are computed by the authors using the IQVIA data set.

The simulations consider a doctor who has a constant load of 300 patients, all drawn from the same k -dimensional normal distribution of people with “true” drug effects taken from the clinical trials. Each patient will have a different optimal drug, but at first all patients will appear to be identical to the doctor. When doing the simulation, we reseed the random number generator for each of the 6 doctor types. This means that each doctor is facing exactly the same population of patients and exactly the same signals from the patients. Hence, any observed differences in outcomes are due only to differences in the doctor’s characteristics. We do 10 runs and take the average over the runs, and then plot doctor utility (which depends on the number of deviations from optimal treatment) and dispersion in doctor prescribing, characterized using the entropy score (we also tried 100 runs, but the results were very similar). Since beliefs are fixed by the clinical data, the simulations explore the effects of doctor diagnostic skill and the doctor’s taste for experimentation.

We consider one additional factor, which is whether there are treatment guidelines. As discussed above, following guidelines constrains the doctor’s treatment choices. In the long run, eliminating treatment options should make patients strictly worse off. However, we saw that empirically, doctors who violate guidelines have worse patient outcomes (empirical finding #3). In the short run, search is costly to patients since it may involve being subjected to sub-optimal treatments, so it is possible that eliminating choices that have a low probability of being optimal could make patients better off.

Poly-pharmacy is a particularly interesting example. Even if the best treatment for a patient involves using several drugs, having to search over all possible drug combinations would increase search costs and the length of time it takes to find the optimal treatment. Since there is little information from drug trials about the efficacy of drug combinations, we focused on two-drug combinations and assumed that their efficacy was given by a weighted sum of the efficacy of each drug, where the weight given to each drug was 0.6. This choice of weights means that in our simulations, the optimal treatment is actually more likely to be a two-drug combination than a single drug. The variance of each two-drug combination is assumed to be equal to the sum of the variances of each drug times the square of 0.6, plus a small amount of idiosyncratic noise. The simulations were done with Julia and the diagrams with ggplot. The code and data that we are allowed to share is available on dataverse (<https://dataverse.harvard.edu/dataverse/curriemacleod>). Table (VI) outlines the steps, with references to the corresponding derivations for the two period problem. The simulation is based upon a multi-period, multi-drug extension of the two period model. See Currie and MacLeod (2020) for details.

TABLE VI
SIMULATION STEPS

- 1.1. Read in parameters values for drug effects.
- 1.2. For runs $R = 1, \dots, 100$ do the following code:
 - 2.1 Set new seed(R) = next prime number after seed($R-1$).
 - 2.2 For each of the six doctor types j in J do the following:
 - 3.1. Seed random number generator with seed(R)
(each doctor sees exactly same set of 300 patients)
 - 3.2. For each patient i in I draw an unobserved vector of treatment effects (5.5) and then:
 - 4.1. Initialize doctor beliefs (5.4)
 - 4.2. Each month $t = 1, \dots, 36$ do:
 - 5.1. Draw an observation y_{ijt} using (5.6)
 - 5.2. Update beliefs regarding treatment effect of drugs (5.7-5.8)
 - 5.3. Compute value for each treatment given beliefs and reservation probability (5.10-5.11)
 - 5.4. Compute optimal choice (2) depending upon guideline case.
 - 5.5. Save results.
 - 3.3. For each month-doctor type compute mean patient utility and prescription diversity for the 300 patients and save results.
- 1.3. Average patient utility and prescription diversity over the runs by month-doctor type.
- 1.4 Plot the results.

Figure (3) illustrates the evolution of dispersion in a doctor's prescribing from time 0 to 36 months as a function of doctor characteristics and whether or not the doctor follows a guideline. Dispersion tends to rise among doctors of all skill levels, though it is constrained by following guidelines as one would expect. However, the effect of following guidelines is much smaller for doctors with a high reservation probability, who place a high value on experimentation. Doctors with a low reservation probability end up with significantly lower dispersion if they follow guidelines.

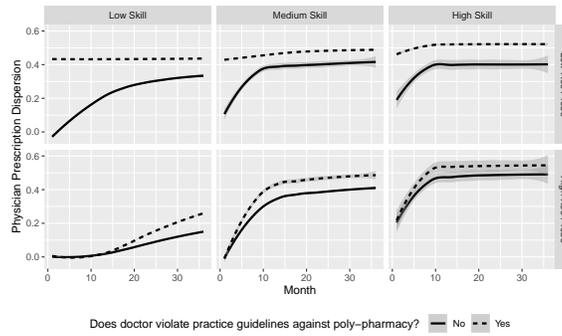


FIGURE 3.— Simulation Results on the Effects of Doctor Skill and Reservation Probabilities on Diffusion in Prescribing, With and Without Guidelines

In Figure (4) we plot the doctor’s utility as a fraction of the maximum possible payoff. Since in our model doctor utility depends only on patient outcomes, higher doctor utility is synonymous with better patient outcomes. In this simulation, low skilled doctors have very flat payoff functions, though they are slowly increasing over time. In contrast, high skilled doctors have payoffs that rise rapidly over the first year of treatment, and then rise more slowly over time. Having a higher reservation probability (i.e. being more experimental) is associated with higher payoffs, especially among high skilled doctors. Perhaps the most remarkable result from this section is that following guidelines is associated with better short-run patient outcomes, even though guidelines constrain treatment choice. This is consistent with our empirical finding #3, that violating guidelines is associated with worse outcomes. The effect of guidelines is greatest for low skilled doctors with high reservation probabilities and for high skilled doctors with low reservation probabilities. One interpretation is that guidelines constrain low skilled doctors who are inclined to experiment on their patients; Guidelines also make it less likely that a high skilled doctor who is inclined to stop searching early will get stuck on a very suboptimal treatment.

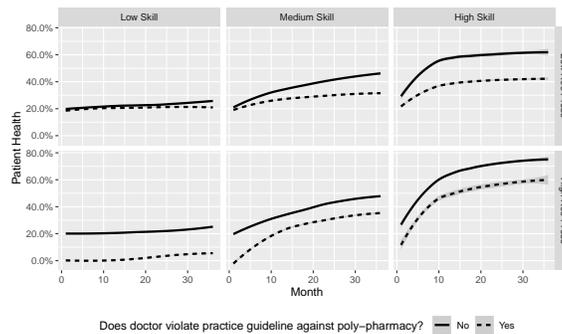


FIGURE 4.— Simulation of the Effects of Doctor Skill and Reservation Probabilities on Utility, With and Without Guidelines

The qualitative patterns we simulate for high skilled doctors match what we observe in the group of patients who ever see psychiatrists. In patients seeing high skilled doctors, higher entropy is as-

sociated with better patient outcomes (empirical finding #1), suggesting that doctors with more dispersed prescribing patterns are better able to match drugs to patients. However, as discussed above, in the patients who see only less-skilled doctors, higher dispersion in prescribing is not empirically associated with better patient outcomes (empirical finding #2). Our simulations suggest that it is difficult to generate this result in our model in which all doctors have similar well-informed priors about drug efficacy and use information efficiently. In real life, a doctor may prescribe badly either because of a lack of diagnostic skill, or because of faulty priors.

Given prescribing data that followed a doctor over a relatively long time period, included a patient identifier, allowed one to track patients who move between doctors, and had information about patient outcomes, one might be able to explore different models of physician belief formation. This might be achieved by estimating a structural model that examines the empirical relationship between doctors' beliefs, reservation probabilities, and skill levels.

The reservation probability could be measured using the number of times a doctor changed drugs within a patient, normalized by the length of time the patient was seen. Data on first prescriptions (ideally to patients who had not been taking any drug for some period of time) could be used to identify doctors with different priors. For example, a doctor who routinely prescribed one drug to new, medication-naive patients could be assumed to have different priors than one who routinely prescribed a different drug. Empirically, all of the information that can be obtained from the data about possible differences in priors can be summarized by differences in actual drug choices conditional on the number of drugs prescribed, so one could control for differences in priors by identifying them with doctor types based on these choices.

Given an indicator about priors, the reservation probabilities, and information about patient outcomes, and assuming that higher skilled doctors will have better outcomes on average as well as outcomes that improve with the length of treatment, the model could allow us to back out a distribution of diagnostic skill. It would then be possible to examine the empirical relationship between skill and reservation probabilities, conditional on priors. Given rapidly evolving access to claims data, and improvements in the quality of such data, it may be possible to estimate a model along these lines in the near future.

6. CONCLUSIONS

Using unique data formed by combining information about how doctors prescribe anti-depressants combined with individual claims data from a large health insurer, we characterize doctor practice style by measuring dispersion in prescribing and whether the doctor violates prescribing guidelines. In turn, differences in these measures of practice style are significantly associated with patient outcomes including total costs, non-drug costs, and emergency room visits, and hospitalizations.

The relationships between practice style and outcomes are quite different among patients who ever saw psychiatrists, and the larger group of patients who were treated only by non-specialists. Among patients seeing skilled doctors, seeing a doctor with more diffuse prescribing is associated with a reduction in the probability of a subsequent ER visit or hospitalization with a mental health indication. In patients seeing unskilled doctors, switching to a doctor with more diffuse prescribing is not associated with improvements in patient outcomes.

We also look at the effects of transitions from one drug to another. Patients whose doctors violate prescribing guidelines pertaining to these transitions have worse outcomes. Estimates from models with patient fixed effects suggest that violations of U.S. guidelines are associated with a 28.2% increase in total costs, a 13.0% increase in the probability of any ER visit or hospitalization, and a 20.0% increase in the probability of an ER visit or hospitalization with a mental health indication. These findings indicate that while some flexibility in treatment options is likely optimal for skilled providers, guidelines are useful to rule out bad practice.

Having shown that practice style is strongly related to patient outcomes, we turn to interpreting the estimates using a model of doctor decision making. We think of the sequence of drug choices as a classic multi-armed bandit problem. Such problems 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. In our model, the trade-off between experimentation and choosing the drug with the highest expected value depends on the physician's diagnostic skill. 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 potential gains from experimentation are large.

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. Physicians who are more skilled will 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. 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, and a more skillful lawyer should take more cases to trial.

Since finding the optimal treatment entails experimentation, it is likely that some choices will turn out to have been suboptimal *ex post*. 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. Using simulations based on data from clinical drug trials and doctors with different assumed levels of skill, we show that in the short run, guidelines that rule out some choices can improve patient welfare. Our empirical findings are consistent with the model's insight that experimentation is more beneficial when physicians have more skill, and that restrictions on physician's choice sets are associated with better patient outcomes over the treatment intervals in our data.

Our results have number of potential 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 of treatment.⁸ 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, the optimal practice style varies with doctor skill and with the uncertain information that the doctor collects while treating the patient. In short-run treatment settings, patients may be better off when guidelines take some treatments off the table, even when those treatments are

⁸For example Skinner (2012)'s nice review of the literature on regional variation is organized around the factors that lead doctors to choose one treatment versus another.

optimal for some patients. The cost of searching over a wide set of treatments may outweigh the benefit of finding a somewhat better treatment. Overall, our results suggest that optimal policy regarding guidelines does not involve a stark choice between following rules or allowing unlimited doctor discretion but should allow doctors to practice medicine within well defined boundaries.

REFERENCES

- 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, 721–730.
- AGHION, P., P. BOLTON, C. HARRIS, AND B. JULLIAN (1991): “Optimal Learning by Experimentation,” *Review of Economics Studies*, 58, 621–654.
- BAICKER, K. AND A. CHANDRA (2005): “The Effect of Malpractice Liability on the Delivery of Health Care,” in *Frontiers of Health Policy Research*, ed. by D. M. Cutler and A. M. Garber, Berkeley, CA: The Berkeley Electronic Press, vol. 8.
- BERGEMANN, D. AND J. VÄLIMÄKI (1996): “Learning and strategic pricing,” *Econometrica*, 1125–1149.
- 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, 457–489.
- BUBECK, S. AND N. CESA-BIANCHI (2012): “Regret analysis of stochastic and nonstochastic multi-armed bandit problems,” *Foundations and Trends® in Machine Learning*, 5, 1–122.
- CHAN, D. C. (2015): “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, 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.
- COHEN, J. D., S. M. MCCLURE, AND A. J. YU (2007): “Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration,” *Philosophical Transactions of the Royal Society B-Biological Sciences*, 362, 933–942.
- COSGROVE, L., A. SHAUGHNESSY, AND T. SHANEYFELT (2018): “When is a guideline not a guideline? The devil is in the details,” *BMJ Evidence-based Medicine*, 23, 33–36, cited By 2.
- CRAWFORD, G. S. AND M. SHUM (2005): “Uncertainty and Learning in Pharmaceutical Demand,” *Econometrica*, 73, 1137–1173.
- CURRIE, J., W. B. MACLEOD, AND J. V. PARYS (2016): “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): “Diagnosis and Unnecessary Procedure Use: Evidence from C-Section,” *Journal of Labor Economics*, 35, 1–42.
- (2020): “Learning, Reservation Probabilities and Boundedly Rational Choice in Bandit Problems,” .
- CUTLER, D., J. SKINNER, A. D. STERN, AND D. WENNBERG (2019): “Physician Beliefs And Patient Preferences: A New Look at Regional Variation In Health Care Spending,” *American Economic Journal: Economic Policy*, 11, 192–221.
- DEGROOT, M. H. (1972): *Optimal Statistical Decisions*, New York, NY: McGraw-Hill Book C.
- 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, 308–315.

- DICKSTEIN, M. J. (2015): "Physician vs. Patient Incentives in Prescription Drug Choice," Stanford University.
- DONOHUE, J. M., H. GUCLU, W. F. GELLAD, C.-C. H. CHANG, H. A. HUSKAMP, N. K. CHOUDHRY, R. ZHANG, W.-H. LO-CIGANIC, S. P. JUNKER, T. ANDERSON, AND S. RICHARDS-SHUBIK (2018): "Influence of peer networks on physician adoption of new drugs," *PLOS ONE*, 13.
- 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.
- 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.
- EREV, I. AND A. E. ROTH (2014): "Maximization, learning, and economic behavior," *Proceedings of the National Academy of Sciences of the United States of America*, 111, 10818–10825.
- 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, 1681–1726.
- FRANK, R. G. AND T. G. MCGUIRE (2000): "Economics and mental health," in *Handbook of Health Economics*, ed. by M. V. Pauly, T. G. McGuire, and P. P. Barros, Elsevier, vol. 1, Part B, chap. 16, 893 – 954.
- 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, 1101 – 1127.
- 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*, American Psychiatric Association, third ed.
- GITTINS, J., K. GLAZEBROOK, AND R. WEBER (2011): *Multi-Armed Bandit Allocation Indices*, John Wiley & Sons, 2nd ed.
- GITTINS, J. C. (1979): "Bandit Processes and Dynamic Allocation Indexes," *Journal of the Royal Statistical Society Series B-Methodological*, 41, 148–177.
- 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, 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, 56.
- HIERONYMUS, F., J. F. EMILSSON, S. NILSSON, AND E. ERIKSSON (2015): "Consistent superiority of selective serotonin reuptake inhibitors over placebo in reducing depressed mood in patients with major depression," *Molecular Psychiatry*, 21, 523–530.
- K, M., L. RC, W. DR, M. TM, AND B. N (2013): "Management practices and the quality of care in cardiac units," *JAMA Internal Medicine*, 173, 684–692.
- KAHNEMAN, D. AND G. KLEIN (2009): "Conditions for Intuitive Expertise: A Failure to Disagree," *American Psychologist*, 64, 515–526.
- KANTOR, E. D., C. D. REHM, J. S. HAAS, A. T. CHAN, AND E. L. GIOVANNUCCI (2015): "Trends in Prescription Drug Use Among Adults in the United States From 1999-2012," *JAMA - Journal of the American Medical Association*, 314, 1818–1831.
- KASPER, S. (1995): "Clinical efficacy of mirtazapine: a review of meta-analyses of pooled data." *International Clinical Psychopharmacology*, 10, 25–35.
- KAY, S. M. (1993): *Fundamentals of Statistical Signal Processing*, Saddle-Hill, NJ: Prentice Hall PTR.
- KENDRICK, D. A., H. M. AMMAN, AND M. P. TUCCI (2014): "Chapter 1 - Learning About Learning in Dynamic Economic Models," in *Handbook of Computational Economics Vol. 3*, ed. by K. Schmedders and K. L. Judd, Elsevier, vol. 3 of *Handbook of Computational Economics*, 1 – 35.
- 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): "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, 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, 45.
- LAI, T. L. AND H. ROBBINS (1985): “Asymptotically efficient adaptive allocation rules,” *Advances in applied mathematics*, 6, 4–22.
- 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, 69–79.
- LORCA, 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, 268–275.
- LUDWIG, J., D. E. MARCOTTE, AND K. NORBERG (2009): “Anti-depressants and suicide,” *Journal of Health Economics*, 28, 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.
- 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, 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): “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, 544–571.
- ROBBINS, H. (1952): “Some Aspects of the Sequential Design of Experiments,” *Bulletin of the American Mathematical Society*, 58, 527–535.
- ROTH, A. E. (2018): “Marketplaces, Markets, and Market Design,” *American Economic Review*, 108, 1609–58.
- ROTHSCHILD, M. (1974): “A two-armed bandit theory of market pricing,” *Journal of Economic Theory*, 9, 185 – 202.
- SCHWARTZ, E. M., E. T. BRADLOW, AND P. S. FADER (2017): “Customer Acquisition via Display Advertising Using Multi-Armed Bandit Experiments,” *Marketing Science*, 36, 500–522.
- SHORROCKS, A. F. (1980): “The Class of Additively Decomposable Inequality Measures,” *Econometrica*, 48, 613–625.
- 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 *Handbook of Health Economics*, ed. by M. V. Pauly, T. G. McGuire, and P. P. Barros, Elsevier B. V., vol. 2, chap. 2, 45–49.
- SRINIVAS, N., A. KRAUSE, S. M. KAKADE, AND M. W. SEEGER (2012): “Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting,” *IEEE Transactions on Information Theory*, 58, 3250–3265.
- SRIVASTAVA, V., P. REVERDY, AND N. E. LEONARD (2015): “Correlated multiarmed bandit problem: Bayesian algorithms and regret analysis,” *arXiv preprint arXiv:1507.01160*.
- STERN, S. AND M. TRAJTENBERG (1998): “Empirical Implications of Physician Authority in Pharmaceutical Decisionmaking,” Working Paper 6851, National Bureau of Economic Research.
- SUTTON, R. S. AND A. G. BARTO (2018): *Reinforcement Learning, 2nd Edition*, MIT Press.
- THEIL, H. (1967): *Information Theory and Economics*, North-Holland.
- THOMPSON, W. R. (1933): “On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples,” *Biometrika*, 25, 285–294.
- TSAI, T. C., A. K. JHA, A. A. GAWANDE, R. S. HUCKMAN, N. BLOOM, AND R. SADUN (2015): “Hospital Board And Management Practices Are Strongly Related To Hospital Performance On Clinical Quality Metrics,” *Health affairs*, 34, 1304–1311N.
- VAN MOFFAERT, M., J. DE WILDE, A. VERECKEN, 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*.
- WU, C. M., E. SCHULZ, M. SPEEKENBRINK, J. D. NELSON, AND B. MEDER (2018): “Generalization guides human exploration in vast decision spaces,” *Nature Human Behaviour*, 2, 915–924.

APPENDIX A: ON LINE APPENDIX

Linear Learning Rule

Both the two period example, and the satisficing drug choice rule use efficient updating of beliefs regarding drug effectiveness. Here we state the optimal rule for reference purposes. See page 488, Chapter 14, Kay (1993) for details. This result also follows from the formula for the Kalman filter. We suppose that we have the following setup:

1. The prior beliefs regarding drug effectiveness are given by:

$$\boldsymbol{\gamma} \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \in \mathbb{R}^m.$$

2. Each period, $t = 1, 2, \dots, T$, a signal $\mathbf{y} \in \mathbb{R}^n$ of $\boldsymbol{\gamma} \in \mathbb{R}^M$ is observed:

$$\mathbf{y}_t = \mathbf{W}_t \boldsymbol{\gamma} + \boldsymbol{\epsilon}_{jt},$$

where the error, $\boldsymbol{\epsilon}_{jt} \in \mathbb{R}^n$, is a multi-variate, iid mean $\mathbf{0}$, Normal distribution with variance $\boldsymbol{\Sigma}_\epsilon \in \mathbb{R}^{n \times n}$, and precision $\boldsymbol{\Lambda}_\epsilon = \boldsymbol{\Sigma}_\epsilon^{-1}$.

3. Let $\boldsymbol{\Lambda}_t = \boldsymbol{\Sigma}_t^{-1}$ denote the precision matrix for the estimate $\boldsymbol{\mu}_t = E[\boldsymbol{\gamma} | \mathbf{y}_1, \dots, \mathbf{y}_t]$ then:

$$\begin{aligned} \boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t + \boldsymbol{\Lambda}_{t+1}^{-1} \mathbf{W}_t^T \boldsymbol{\Lambda}_\epsilon (\mathbf{y}_{t+1} - \mathbf{W}_t \boldsymbol{\mu}_t), \\ \boldsymbol{\Lambda}_{t+1} &= \boldsymbol{\Lambda}_t + \mathbf{W}_t^T \boldsymbol{\Lambda}_\epsilon \mathbf{W}_t. \end{aligned}$$

Notice that $E\{\mathbf{y}_{t+1} | \mathcal{I}_t\} = \mathbf{W}_t \boldsymbol{\mu}_t$ and hence $(\mathbf{y}_{t+1} - \mathbf{W}_t \boldsymbol{\mu}_t)$ is the new information in period $t + 1$ that is used to update the prior as a function of the precision of the signal, $\boldsymbol{\Lambda}_\epsilon$.

Proof of Proposition

PROPOSITION (2) *A Bayesian rational doctor chooses drug A over drug B in period 1 if and only if:*

$$(A.1) \quad \mu_{j1A} + \zeta_j V_{j1A} \geq \mu_{j1B} + \zeta_j V_{j1B},$$

an equality that is satisfied to the first order if and only if:

$$(A.2) \quad \mu_{j1A} + \tau_j \hat{\sigma}_{j2A} \geq \mu_{j1B} + \tau_j \hat{\sigma}_{j2B},$$

where

$$\tau_j = \frac{\zeta_j}{1 - \zeta_i} L(0) \simeq 0.4 \frac{\zeta_j}{1 - \zeta_i}.$$

PROOF: The first inequality follows immediately from the definition of the payoffs. For the second inequality take a Taylor series expansion of $V(x)$ around $x = 0$.

$$V(x) = \int_x^{-\infty} t f(t) dt + (1 - F(x)) x.$$

$$\begin{aligned} V'(x) &= x f(x) - f(x) x + (1 - F(x)), \\ &= (1 - F(x)). \end{aligned}$$

Hence:

$$V(x) \simeq L(0) + \frac{x}{2}.$$

Thus we have:

$$\begin{aligned} U(d_1 = A, d_{2A}^*) &\simeq \mu_{j1A} + \zeta_j \hat{\sigma}_{j2A} \left(L(0) + \left(\frac{\mu_{j1B} - \mu_{j1A}}{2 \hat{\sigma}_{j2A}} \right) \right) \\ &= \mu_{j1A} + \zeta_j \hat{\sigma}_{j2A} L(0) + \zeta_j \frac{\mu_{j1B} - \mu_{j1A}}{2}. \end{aligned}$$

We have a similar expression for drug B:

$$U(d_1 = B, d_{2B}^*) = \mu_{j1B} + \zeta_j \hat{\sigma}_{j2B} L(0) + \zeta_j \frac{\mu_{j1A} - \mu_{j1B}}{2}$$

Re-arranging the mean terms we get that to the first order $U(d_1 = A, d_{2A}^*) \geq U(d_1 = B, d_{2B}^*)$ if and only if:

$$\mu_{j1A} + \frac{\zeta_j}{(1 - \zeta_j)} L(0) \hat{\sigma}_{j2A} \geq \mu_{j1B} + \frac{\zeta_j}{(1 - \zeta_j)} L(0) \hat{\sigma}_{j2B},$$

from which we get (A.2). The numerical approximation to $L(0) \simeq 0.4$ is to be found in the appendix to Raiffa and Schlaifer (2000). Q.E.D.

A.1. Drug Effects and Dropouts: How Data for Table 5 is Constructed

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.

1. **Sertraline (Zoloft):** 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,

$$\sigma_{endpoint}^2 = \sigma_{baseline}^2 + \sigma_{effect}^2$$

Solving for σ_{effect} , we have:

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

2. **Citalopram HBR (Celexa):** 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.
3. **Fluoxetine HCL (Prozac):** 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.
4. **Escitalopram Oxal (Lexapro):** 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:

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

5. **Trazodone HCL (Oleptro):** 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 $\sigma_{effect}^{traz} = \sigma_{effect}^{mirt}$, where σ_{effect}^{mirt} is defined below.
6. **Duloxetine HCL (Cymbalta):** The mean effect was drawn from 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.

7. **Wellbutrin XL:** 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).

8. **Amitriptyline HCL (Elavil):** 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.

9. **Venlafaxine (Effexor):** 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.

10. **Mirtazapine (Remeron):** 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.

11. **Paroxetine (Paxil):** 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.

12. **Placebo:** 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).

A.2. Market Shares of Drugs

TABLE VII
 MARKET SHARES OF ALL ANTIDEPRESSANT PRESCRIPTIONS, 2006 AND 2014
 (BRANDED PRODUCTS IN ITALICS UNDER THE EQUIVALENT GENERIC PRODUCT)

	Drug Class		Molecule		Specific Product	
	2006	2014	2006	2014	2006	2014
SSRI	54.40	54.03				
Sertraline			14.02	14.63	4.69	14.56
<i>(Zoloft)</i>					9.33	0.07
Citalopram			6.17	12.83	5.98	12.81
<i>(Celexa)</i>					0.19	0.02
Fluoxetine			11.08	10.57	10.69	10.53
<i>(Prozac)</i>					0.39	0.04
Escitalopram			13.05	9.68	..	9.50
<i>(Lexapro)</i>					13.05	0.18
Paroxetine			9.71	5.33	7.86	5.28
<i>(Paxil)</i>					1.85	0.05
Other			0.37	0.99	0.37	0.99
SNRI	13.50	15.04				
Venlafaxine			9.14	7.09	0.18	6.99
<i>(Effexor)</i>					8.96	0.10
Duloxetine			4.33	6.84	..	6.36
<i>(Cymbalta)</i>					4.33	0.48
Other			0.03	1.11	0.03	1.11
NDRI	11.45	10.46				
Bupropion			11.45	10.46	2.88	10.29
<i>(Wellbutrin)</i>					8.57	0.17
SARI	7.37	9.42				
Trazadone			7.12	9.35	7.12	9.35
Other			0.25	0.07	0.25	0.07
Tricyclic	10.77	8.16				
Amitriptyline			6.62	5.15	6.62	5.15
Other			4.15	3.01	4.15	3.01
Tetracyclic	2.44	2.84				
Mirtazapine			2.36	2.82	2.36	2.82
<i>(Remeron)</i>			0.08	0.02	0.08	0.02
MAOI	0.07	0.05	0.07	0.05	0.07	0.05
Total	100.00	100.00	100.00	100.00	100.00	100.00

Note: Authors' calculations from the IQVIA data. Different drug classes correspond to different hypothesized methods of action. Only the most commonly prescribed drugs in each class are listed. In most cases there is a generic and a branded drug. E.g. Zoloft is the brand name and the equivalent generic is Sertraline. We give the molecule the generic name. The table shows that 11 drugs account for most of the market though 33 drugs were sold. "Wellbutrin" includes the branded drugs Wellbutrin, Budeprion, Forfivo, and Aplenzin; "Paxil" includes both Paxil and Pexeva.