Information and Innovation Diffusion: The Case of Pharmaceuticals in the United States*

Kenneth J. Arrow
L. Kamran Bilir
Alan Sorensen

July 2016
PRELIMINARY

Abstract

How does information affect the diffusion of innovations? This paper empirically assesses the influence of physicians' access to detailed drug information on their decisions about which products to prescribe. Combining data on prescriptions and electronic drug reference use for over 130,000 individual U.S. physicians, we find that physicians with access to a pharmaceutical reference database prescribe a significantly more diverse set of products than other doctors, and begin prescribing newly approved drugs sooner than other doctors. These latter effects are particularly pronounced for generic drugs, suggesting improvements to physician information access may have important implications not only for the diffusion of new medical technology, but also for the costs and efficiency of medical care.

*Arrow: Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305, arrow@stanford.edu. Bilir: Department of Economics, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, WI 53706, kbilir@ssc.wisc.edu. Sorensen: Department of Economics, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, WI 53706, sorensen@ssc.wisc.edu. We thank the Ewing Marion Kauffman Foundation for its financial support of this research; Bilir also thanks the Princeton University IES for its generous support. We thank Nick Bloom, Michael Dickstein, Jeff Gambino, Shane Greenstein, Bob Litan, Greg Rosston, Randall Stafford, Jeff Tangney, and the research team at IMS Health for guidance and very helpful conversations.

I. Introduction

Disparities in the cost and quality of medical care across U.S. regions have attracted the attention of economists and policymakers in recent years.[1] While many factors potentially contribute to these disparities, an important and intriguing possibility is that the observed variation in doctors' treatment decisions reflects, in part, a lack of uniformity in the information they have about available therapies—particularly in disease areas undergoing a rapid expansion in treatment options.

Consider for example, changes in the treatment of cardiovascular disease. When the link between cholesterol and heart attack risk was first established in the 1950s, priorities shifted toward disease prevention through cholesterol management, leading to the introduction of cholesterol drugs. The first such product to be made available for medical prescription in the United States was Mevacor, in 1987. By 2000, seven distinct products were available, including but not limited to statins; by 2010, physicians juggled at least 18 pharmaceutical products aimed at controlling lipids in the bloodstream (Table 1; Appendix A.1). From the perspective of the physician, differences between these products are often subtle: 'performance' depends on patient characteristics and preferences; prices depend on patient insurance coverage.[2] Complicating matters further, product information evolves continually as clinical trials reach new conclusions. Thus, while access to information about available therapies almost certainly shapes physicians' decisions about patient care, the magnitude and direction of this influence are unclear.

This paper empirically assesses the influence of physicians' access to information on their decisions to treat patients using new medical technology. We look specifically at doctors' decisions about which cholesterol-management drugs to prescribe, and how these decisions are affected by their adoption and use of an electronic drug reference database. This database allows physicians to look up detailed drug information—including, in some cases, whether a particular drug is covered by a patient's insurance plan—at the point of care, either on a computer or on a mobile device. Using detailed data for over 130,000 individual U.S. physicians between January 2000 and December 2010, we find that physicians with access to the electronic database prescribe a significantly more diverse set of products than other doctors, and prescribe new drugs sooner than other doctors. Such doctors are also, at any point in time, more likely to prescribe a new drug that is within 24 months of its initial release than are other doctors. These latter effects are particularly pronounced for generic products, suggesting improvements to physician information access may have important implications for both the diffusion of new medical technology and for the costs of healthcare.

Because access to the drug database is not randomly assigned—doctors choose whether

[1] See Wennberg et al (1996), Cooper et al (2015), Gawande (2009), and Chernew et al (2009).
[2] For example, Brooks et al (2014).

and when to subscribe—identifying the causal effects of database access is challenging. Prescription patterns of subscribing doctors may look different from those of non-subscribers not due to any effects of the database itself, but rather due to differences in the types of doctors who choose to subscribe. Indeed, we document that the observed characteristics of early adopters are meaningfully different from late adopters and non-adopters, suggesting that unobserved differences may also be important. With this challenge in mind, our analysis relies heavily on within-doctor variation: rather than estimating effects by comparing database users to non-users, we estimate by comparing a doctor's own prescriptions before versus after she begins using the database. In addition, to address concerns about selection, we estimate the impact of 'placebo' database use on the prescribing decisions of physicians who adopt the database but never use it to search for cholesterol drugs. Taken together, these specifications support a causal interpretation of our main results—i.e., that using the database caused a small but statistically significant change in prescribing patterns.

Our focus on the market for pharmaceutical drug products is motivated in part by the availability of unusually detailed data. Using a newly-assembled dataset from IMS Health and the provider of the aforementioned drug reference database,[3] we are able to evaluate the influence of information access on the physician-level prescription response to new drug introductions. Physician use of the electronic medical reference is extensive—over half of U.S. physicians are users—but requires registration that associates reference activity with the user's unique, time-invariant American Medical Association (AMA) identifier. Pharmacies similarly record each prescription drug purchase with the AMA identifier of the prescriber, enabling our link between information access and prescribing for each doctor, product, and time period. We are thereby able to compare prescriptions of new and existing products among users and non-users of the reference database, and across physicians that use the database with different intensities. In addition, the data we examine span 132 consecutive months, allowing us to observe responses to the entry of multiple new drugs. The panel structure of the data also allows us to better isolate the effects of database access by relying on within-doctor variation over time—that is, we can examine how a doctor's prescribing patterns change after she begins using the database.

Two features of the market for cholesterol control therapies make it especially suitable for our study. First, 12 new drugs were introduced to the U.S. market during our 11-year sample period. These innovations ranged from products based on entirely new molecules to new generic versions of existing molecules (Table 1). Second, as a class, lipid control drugs are very widely used. The Centers for Disease Control and Prevention estimate that approximately 71 million U.S. adults suffer from chronic hypercholesterolemia and dyslipidemia,

[3]The provider of the database is a leading U.S. point-of-care medical applications firm, but chose to remain unnamed in this study.

conditions in which abnormal levels of cholesterol or lipids are present in the bloodstream. These conditions are associated with heart disease, heart attack risk, and premature death; accordingly, sales of cholesterol therapies accounted for over $35 billion U.S. dollars in 2012 (IMS Health 2013). The rate of new product diffusion in this disease area thus has the potential to affect a large segment of the U.S. population, both economically and physically.

The data reveal substantial differences across U.S. locations and physicians in the prescription of cholesterol drugs and in their cost. In December 2010, for example, the share of generic products in overall prescribing spanned the full range (Figure 1, Panel A); physicians in 267 zipcodes prescribed only generics, while physicians in 177 others prescribed only branded products. Some of this variation may be explained by underlying patient heterogeneity, but several U.S. cities simultaneously host both locations that prescribe only generic products and locations that prescribe none; such cases strongly suggest factors beyond patient heterogeneity may influence prescribing.[4] Importantly, these differences affect costs. We summarize variation across U.S. locations in the average cost per pill dispensed in Figure 1, Panel C, using price data from the Centers from Medicare and Medicaid Services (CMS). These data imply an approximate interquartile range for the average cost per monthly prescription of $46.80 in December 2010.[5] Multiplied by the average number of prescriptions per year (916.7), this implies an approximate annual cost difference in excess of $42,000 between the 25th- and 75th-percentile prescribers. While this cost difference must reflect, to some extent, medically-appropriate responses to patient heterogeneity, it also indicates that a large potential exists for cost savings in the treatment of cardiovascular disease.

This paper is closely related to an extensive literature on the diffusion of new medical technology. Classic work by Coleman, Katz, and Menzel (1957, 1996) finds that new pharmaceutical products diffuse unevenly across medical practitioners: physicians that interact more frequently with other physicians are more likely to adopt early. We build on these results by examining the influence of a digital information database on physicians' drug adoption decisions and find that this, too, has an impact on adoption. Our work is also related to Skinner and Staiger (2007), who examine differential rates of adoption for technologies as varied as beta blockers and hybrid corn (Griliches 1957). While their focus is on explaining state-level variation in adoption rates, our focus is on individual (physician-level) differences. In this respect our work is similar to Crawford and Shum (2005), who analyze panel data on anti-ulcer prescriptions to estimate a model of physician learning; Agha and

[4]Three examples of this in December 2010 are Los Angeles, CA (90058 prescribes no generics, while 90056 and 90062 prescribe only generics), Birmingham, AL (35224 prescribes no generics, while 35223 prescribes only generics), and Indianapolis, IN (46224 prescribes no generics, while 46235 prescribes only generics). Heterogeneity is equally pronounced for Lipitor (Figure 1, Panel B), a branded statin.

[5]That is, we calculate the cost per prescription (a 30-day supply) written in December 2010 for each physician, and then calculate the interquartile range (across physicians) of these averages. Note that our use of CMS prices for these calculations implies they are only rough estimates of true prescription costs.

Molitor (2015), who examine the influence of physician investigators on the regional diffusion of new anti-cancer drugs; and Escarce (1996), who studies physicians' decisions to adopt a new surgical technology.

Our paper is also related to work investigating disparities in prescribing across U.S. physicians, including the Dartmouth Atlas and its analysis of prescription drug use among Medicare patients (Munson et al 2013), and Cooper et al (2015) for the privately-insured. The data we evaluate include prescriptions for all U.S. patients, enabling our assessment of prescribing determinants within a near-universal set of prescribers and patients. However, we do not observe individual patient characteristics, precluding a direct extension of Munson et al (2013) to non-Medicare patients, as well as a quantitative welfare analysis.[6]

We evaluate how physician access to an electronic drug database affects prescribing, and our paper thus contributes to work aimed at evaluating the impact of information technology on economic decisions and outcomes. Individual agents' ability to access electronically available information can affect productivity (Solow 1987) and has been specifically shown to improve performance in healthcare delivery (Athey and Stern 2002, Dranove et al. 2014).[7] Other studies have evaluated information technology as a cost-reducing process innovation.[8] In our setting, physicians' digital access to drug information could improve health outcomes by improving the match quality between patients and treatments, and could also reduce costs by accelerating the adoption of newly introduced generic drugs.

In this, our paper complements research on general theories of technology diffusion that feature agents with imperfect information. Such theories can be shown to explain the large existing differences in productivity across locations (Solow 1956, Arrow 1969, Parente and Prescott 1994, Comin and Hobijn 2004) as identified in Klenow and Rodriguez-Clare (1997) and Casselli and Coleman (2006), for example. In the medical context of our analysis, heterogeneity in the extent of new technology adoption may be natural, reflecting differences in patient composition; however, delay in adoption is not—particularly in the case of a new generic product, essentially a pure, cost-reducing innovation.[9]

The rest of the paper is organized as follows. Section II describes the data used in our analysis. Section III describes a simple model of prescription choice and our estimation framework. Section IV presents the empirical results, Sections V and VI discuss interpretation, and Section VII concludes.

[6]This aspect of our dataset further precludes estimating a model featuring prescription dynamics within each patient-physician pair, as in Crawford and Shum (2005) or Dickstein (2015).

[7]See also Bresnahan, Brynjolfsson, and Hitt (2002), Bloom, Sadun, and Van Reenen (2012).

[8]See, for example: Attewell 1992; Bresnahan and Greenstein 1996; Black and Lynch 2001; Brynjolfsson and Hitt 2003; Hubbard 2003; Forman, Goldfarb, and Greenstein 2005; Bloom et al. 2009; Agha 2012.

[9]The idea that underlying patient heterogeneity could influence the optimal prescription mix, and potentially also the diffusion of new technology, is related to David (1966); the study finds that in a neoclassical setting, labor-saving technology optimally diffuses more readily to labor-scarce locations, a pattern confirmed using U.S. data on the adoption of the mechanical cotton reaper.

II. Data and Descriptive Evidence

Evaluating the influence of information access on new pharmaceutical drug diffusion requires detailed measures of drug innovations and individual prescribers' treatment decisions, information usage, and characteristics. We introduce each of these measures below and go on to describe physicians' prescribing of new and existing pharmaceutical drugs.

A. *U.S. Innovations in Chronic Hypercholesterolemia and Dyslipidemia Therapy*

At the start of the sample period in January 2000, six pharmaceutical therapies were available to assist with patient cholesterol control, a clinical priority for the prevention and treatment of cardiovascular disease: Lescol, Lipitor, Mevacor, Niaspan, Pravachol, and Zocor.[10] Thereafter, twelve new cholesterol or lipid control therapies were introduced, including new formulations, combinations, and versions.[11] These include three new molecular entities, Crestor, Lovaza, and Zetia; three generic versions, lovastatin (Mevacor), pravastatin (Pravachol), and simvastatin (Zocor); two new formulations, Altoprev (extended-release Mevacor) and Lescol XL (extended-release Lescol); and four new drug combinations, Advicor (extended-release niacin and Mevacor), Pravigard PAC (aspirin and Pravachol), Vytorin (Zetia and Zocor), Simcor (extended-release niacin and Zocor). Each new therapy received nationwide approval by the U.S. Food and Drug Administration (FDA) on a known, drug-specific date (Table 1). All products are described in Appendix A.1.

While these 18 products are therapeutic substitutes, in that they aim at a similar clinical endpoint—cholesterol or trigliceride reduction—they are only imperfect substitutes: each product features distinctive characteristics relevant for the prescribing decision. First, many but not all cholesterol therapies are pure statins, which act to reduce cholesterol synthesis in the liver by inhibiting a specific coenzyme; these include Lescol (fluvastatin), Lipitor (atorvastatin), Mevacor (lovastatin), Pravachol (pravastatin), Zocor (simvastatin), Crestor (rosuvastatin), Altoprev (extended-release lovastatin), and Lescol XL (extended-release fluvastatin). Other products rely on different mechanisms of action: Zetia (ezetimibe), for example, is distinct in that it achieves cholesterol reduction by reducing intestinal absorption of cholesterol. A second distinction involves therapeutic intensity. High doses of Lipitor and Crestor are more effective at lowering low-density lipoprotein (LDL) cholesterol than alternatives (Law et al 2003). Side effects are also relevant; evidence suggests, for example, that high doses of Lipitor and Crestor may also increase the incidence of negative side effects,

[10] Cannon et al (2004).

[11] To ensure adequate coverage in the data, we consider all cholesterol therapies introduced by December 2008 but not those introduced after this date. For the same reason, our analysis excludes Baycol, a drug that was available in January 2000 but withdrawn from the market in August 2001.

while combination therapies such as Vytorin may in certain cases be more appropriate care for patients with severe cholesterol abnormalities (Kastelein et al 2008).

More subtly, clinical evidence suggests the benefits and risks associated with statins are heterogeneous across patients; randomized-controlled trials (RCTs) indicate, for example, that the benefits of statin use are higher for patients with diabetes, negligible among those with prior heart failure, and vary with age; risks and side-effects also vary with statin intensity, age, weight, comorbidities, and so on (Brooks et al 2014). Adding to this, patients with 'complex' attributes are often underrepresented in RCTs, raising clinical uncertainty and, accordingly, the likelihood that patient preferences—including willingness to suffer side effects and to pay for medications—may influence the prescribing choice (Brooks et al 2014).

Physicians' decisions about which drugs to prescribe are further affected by the evolution of clinical information as new trials are completed—particularly head-to-head studies aimed at establishing the relative efficacy of one drug therapy over another.[12] These ongoing changes in clinical evidence, combined with an expanding set of available products, suggest that physicians may turn to drug references that help to ensure prescription decisions are based on current information.

B. *Prescriptions by U.S. Physicians*

To measure physicians' prescribing of new and existing therapies aimed at cholesterol and lipid control over time, we use physician-level prescription data for the 18 drugs described above from the IMS Health Xponent database. These data are provided at a monthly frequency by drug during the period January 2000 through December 2010, and cover each of the 280,622 U.S. physicians associated with at least ten cholesterol-drug prescriptions during January to December 2010; this low threshold for inclusion implies that our dataset captures essentially the universe of U.S. cholesterol drug prescriptions during this period. For each product and month, we observe the number of prescriptions written by each physician and filled through a U.S. pharmacy. Beginning in January 2006, the data also include information on the method of payment used to fill each prescription (Medicaid, Medicare Part D, Cash, or Commercial Third-Party Insurance). Importantly, each physician in the dataset is identified by a unique medical education number, name (first name, last name, middle name), and location (a five-digit U.S. zipcode). These identifiers enable us to match individual prescribers with their observed pharmaceutical information technology use.

[12]For example, an RCT completed in 2004 demonstrated that for patients with severe cholesterol abnormality, the incrementally larger reductions achieved by Lipitor resulted in fewer deaths and major coronary events relative to patients taking Pravachol (Cannon et al 2004). Another such study released in 2008 found that, while Vytorin achieved larger cholesterol reductions than simvastatin, the two drugs were observably identical when it came to the thickness of arterial plaque buildup (atherosclerosis); adding to this, a second study in 2008 found a positive association between Vytorin and cancer (Rossebo 2008) that was later reversed (Cannon et al 2015).

To ensure that our sample includes only those physicians actively prescribing cholesterol drugs during the entire sample period, we restrict attention to the 131,323 physicians that prescribe ten or more statins both during January to December 2000, and during January to December 2010; this allows us to abstract from potential differences in prescribing that may surround a physician's entry into or exit from medical practice, and also ensures that we have adequate data on database adopters' pre-adoption and post-adoption prescribing patterns. The final prescription dataset includes over 200 million observations (132 months $\times$ 131,323 physicians $\times$ up to 18 drugs). Summary statistics appear in Table 2, and additional details regarding data assembly and the Xponent database appear in Appendix A.2.

C. *Drug Information Access by U.S. Physicians*

To construct an index for the extent of physicians' pharmaceutical information access, we use physician-level data from the private firm that owns and operates a prominent electronic reference for pharmaceutical products. The data include registration status and information on use of the drug reference for the same set of 280,622 physicians, 18 cholesterol control products, and 132 months described above. Specifically, we observe a monthly physician-specific indicator for whether a U.S. physician is a registered user of the pharmaceutical reference. This variable indicates the database is widely used by physicians in the United States: by December 2010, 44.6 percent of the 131,323 sample physicians use the reference. The sample thus includes physicians that never use the database, but for whom we nevertheless observe prescribing decisions. The data also include information about registered physicians' actual use of the database. We summarize this information with a variable indicating whether a physician used the database to look up a cholesterol drug at least once during the sample period. In the average month, this variable indicates 24.2 percent of physicians are registered users, and 13.1 percent of physicians used the database to look up one of the cholesterol drugs considered in our study. The data thus include 'placebo' physicians that adopt the database but never use it to search for information about cholesterol medications. There is no reason to expect the database to affect these 'placebo' physicians' prescribing; we examine this hypothesis carefully in Section V.

The drug reference contains information that is, in principle, relevant for improving the match between patient characteristics and available pharmaceutical products. At any point in time, the drug reference contains detailed information about each available U.S. FDA-approved medication. This information is obtained from the medical literature, specialist recommendations, clinical guidelines, manufacturer labeling, standard medical references, and FDA drug safety alerts and is updated continually; the results of this ongoing research are condensed into drug-specific monographs that may be accessed through the electronic database interface. Beyond standard clinical information such as contraindications, cautions,

adverse reactions, safety, monitoring, and pharmacology, the reference monographs also include a set of additional variables for each product that may affect prescribing decisions. Specifically, the monographs include retail pricing and formulary status information for each drug, drug interaction information, FDA warnings, and off-label and pediatric usage guidelines. The database includes separate entries for each branded product and each generic (if available), based on product-specific information such as available formulations, dosing, indications, manufacturer, and pricing, as these differ even between products containing the same active ingredient. The information contained in the drug database is updated over time and therefore reflects both the current set of products and the current state of knowledge regarding drug characteristics and clinical practice. Importantly, information about new drugs approved by the FDA becomes available through the database at around the time the drug becomes commercially available.

Because the drug reference combines existing information into a single, current monograph rather than contributing novel drug information, it is best viewed as a tool that makes it convenient for physicians to quickly access accurate and condensed clinical, insurance, and pricing information about a drug. Doctors commonly use the reference to check dosages and contraindications, for example, but rely onother sources, such as medical journals or other more encyclopedic references, for information such as a drug's results in clinical trials.

Figure 2 indicates that use of the reference database during the sample period is not random, but instead differs across U.S. physicians according to observable characteristics.[13] Physicians graduating from top-ranked U.S. medical schools, as well as recent graduates, tend to begin use of the database sooner; a greater share of such individuals thus use the database at any point in time during the sample period (Panels A, B). Figure 2 further indicates male physicians adopt the database with a greater propensity than do female physicians (Panel C). While there is little relationship between reference use and monthly statin prescribing volumes (Panel D), doctors specializing in obstetrics and gynecology, and those practicing in the U.S. South are slower to adopt the database, compared with others (Panels E, F).

D. *Descriptive Evidence*

The data provide suggestive indications that incomplete information may affect physicians' prescribing. Consider the statistics presented in Table 3, which quantify differences in prescribing across U.S. physicians for the class of cholesterol medications evaluated. The statistics in Panel A provide evidence for the December 2010 cross section, restricting attention to physicians prescribing at least 30 cholesterol medications during that month. Within this group, it is apparent that the pronounced variation in cholesterol-drug prescribing pre-

[13]Physician characteristics were obtained from the Centers for Medicare & Medicaid Services Physician Compare database, and were matched based on their first name, last name, and five-digit zipcode.

viously found among Medicare patients (e.g. Munson et al 2013, Brooks et al 2014) is also present within the overall population (columns 6 and 7). The share of prescriptions accounted for by Lipitor, for example, ranges from zero to one in column 6; moreover, while the average physician prescribes Lipitor in 19.2 percent of cases, the standard deviation is also large (11.0 percent). Importantly, column 7 indicates these wide observed differences in prescribing are even larger along the branded-generic margin.

Even if physicians were perfectly informed, variation in prescribing could result from an uneven distribution of patient characteristics. For example, Lipitor is a high-intensity statin that may be preferable for patients with a severe cholesterol abnormality, the incidence of which may cluster geographically. Columns 4 and 5 thus evaluate prescribing among physicians practicing within the same five-digit U.S. zipcode. Even though the scope for patient heterogeneity should be limited within such a narrow geography, the statistics reveal significant variation even at this level. In the case of Lipitor, the average U.S. zipcode observes a 15.8 percentage point gap in its prescription share between the highest- and lowest-intensity Lipitor prescribers, and a 24.6 percentage point gap in the share of prescriptions accounted for by generics.

Unobserved patient heterogeneity likely explains some of this variation in prescribing, but columns 1, 2, and 3 indicate that additional factors are also likely present. Specifically, these columns assess within-zipcode variation in the diffusion of new generic products. The advantage of this approach is that it is possible to compare prescribing of both a branded product and its molecularly-equivalent generic, two distinct drugs that have no relevant clinical differences. And, by examining changes over time in the generic share of molecule-specific prescriptions, it is possible to determine whether stable patient heterogeneity is likely to be the only explanation for variations in care. For each of the three generic drug introductions (lovastatin, pravastatin, and simvastatin), the data indicate that physicians differ in their use of generics in the short run, six months after generic entry, and that substitution toward generics is initially incomplete at this point (Panel B). By contrast, in the long run, physicians differ substantially less: nearly complete substitution toward generics is observed for each of the three products (Panel C).[14] This pattern of delayed substitution strongly suggests factors other than patient heterogeneity contribute to prescribing differences among cholesterol drugs, and is consistent with information frictions.

Beyond cost implications, these same factors may impede the diffusion of new non-generic therapies, with consequences for health outcomes. The data indicate that the physicians are

[14]By December 2010, physicians had broadly switched away from prescribing Mevacor, Pravachol, and Zocor: each has an average within-molecule prescription share of essentially zero among those prescribing at least 30 cholesterol drugs in that month. However, six months after each respective patent expired, generic prescribing was far more varied across physicians even though the generic version was in each case already substantially less expensive. In November 2006, the gap in generic prescription shares was over 20 percent for each product in the average zipcode, but fell to less than four percent by December 2010.

slow to begin prescribing new molecular entities, new drug combinations, and new dosage forms—branded products not facing generic competition. Figure 3 plots the gradual diffusion of Crestor across U.S. zipcodes; Table 1 describes how the time lag in months between a drug's approval and its initial prescription varies across U.S. physicians for each drug introduction. The average physician delays prescribing a new drug for 19.1 months among the new products considered in our analysis; the standard deviation is even larger (22.0 months), and this adoption lag ranges between zero and 122 months, indicating some physicians adopt immediately and others had yet to adopt the first new drug of the sample by the final period, December 2010 (Table 2). With these motivating facts in hand, we now turn to evaluate the influence of physicians' electronic pharmaceutical reference use on prescribing for new drugs.

III. Empirical Strategy

The drug reference database we evaluate may be viewed as a technology that reduces physicians' costs of acquiring information relevant to matching patients with treatments. In this section we provide a framework indicating how we expect this cost reduction to affect database users' prescribing patterns, and describe our approach to measuring these effects.

A. *Conceptual Framework*

Consider a baseline model in which physician i faces a period-t choice over which drug to prescribe for each of her patients $n = 1, 2, ..., N_{it}$. Like other economic studies of prescribing decisions, suppose that physician i makes this decision for each patient by selecting the single drug $j \in \{1, 2, ..., J_t\}$ available at t that maximizes patient utility according to physician-i information.[15] Specifically, suppose that the true utility derived by patient n from drug j at t is $u_{njt} \equiv \theta_{jt} + V_{njt}$, which combines the quality of drug j that is both known at t and common across patients (θ_{jt}) with the quality of j that is unknown and partially specific to patient n (V_{njt}). The first of these terms (θ_{jt}) thus captures the accepted wisdom at t about the efficacy, costs, side effects, and so on of drug j for the average patient, while the second reflects novel information that may, in part, be relevant to the match between j and patient n. In particular, suppose that V_{njt} combines two terms: $V_{njt} \equiv v_{jt} + \epsilon_{njt}$, where v_{jt} is a drug-specific value—a revision to accepted wisdom about the quality of drug j—and where ϵ_{njt} reflects the quality of the match between patient n and drug j. While the precise value of V_{njt} is not immediately known to physician i, suppose that she may exert effort to learn about its value. In particular, assume that by exerting effort $\phi_{it} \in [0, \infty)$, physician i bases

[15] See, for example, Dickstein (2015), Crawford and Shum (2006).

her prescribing on a partial observation of u_{njt}, given by

$$\hat{u}_{njt} \equiv \theta_{jt} + (1 - e^{-\phi_{it}})V_{njt} = \theta_{jt} + (1 - e^{-\phi_{it}})\,(v_{jt} + \epsilon_{njt}). \tag{1}$$

This ϕ may include, among other things, effort spent researching drug side effects, interactions, and efficacy based on current clinical trials in standard medical references; it may include effort put toward asking patients about symptoms or medical history; or it could include acquiring other patient-drug specific information such as dosage, retail pricing, and insurance formulary status.[16] A higher level of effort ϕ in (1) thus increases physician-i sensitivity to existing but novel information about drug quality (v_{jt}) and about the patient-specific match (ϵ_{njt}). In particular, (1) implies physicians exerting no effort $\phi = 0$ are insensitive to V_{njt} and thus prescribe the same drug—that with the highest θ_{jt}—for all patients, while physicians exerting infinite effort respond to V_{njt} perfectly.

If we assume that the ϵ_{njt} follow an i.i.d. Type-1 Extreme Value distribution, it is straightforward in this simple setup to show that the probability physician i prescribes drug j for patient n at t depends on the effort ϕ_{it} spent acquiring information as follows

$$p_{jt}(\phi_{it}) = \frac{\exp\left\{\frac{\theta_{jt}}{1-e^{-\phi_{it}}} + v_{jt}\right\}}{\sum_{k=1}^{J_t} \exp\left\{\frac{\theta_{kt}}{1-e^{-\phi_{it}}} + v_{kt}\right\}}$$

and that, accordingly, the probability P_{ijt} that drug j is prescribed by physician i at least once during period t is

$$P_{ijt}(\phi_{it}) \equiv P\{X_{ijt} > 0\} = 1 - P\{X_{ijt} = 0\} = 1 - (1 - p_{jt}(\phi_{it}))^{N_{it}} \tag{2}$$

where X_{ijt} is the number of physician-i prescriptions written for drug j at t.[17] Moreover, starting from an initial date t_0 (such as the introduction date of a new drug j), the expected number of periods T_{ij} that lapse before drug j is prescribed at least once by physician i is

$$\begin{aligned} E[T_{ij}] &= \sum_{t=t_0}^{\infty}(t - t_0)P_{ijt}(\phi_{it}) \prod_{s=t_0}^{t-1}(1 - P_{ijs}(\phi_{is})) \\ &= \sum_{t=t_0}^{\infty}(t - t_0)\left(1 - (1 - p_{jt}(\phi_{it}))^{N_{it}}\right) \prod_{s=t_0}^{t-1}(1 - p_{js}(\phi_{is}))^{N_{is}} \end{aligned} \tag{3}$$

[16] For simplicity, we abstract from the idea implicit in (1) that physician i could use $\hat{u}_{njt}$ and her knowledge of both ϕ_{it} and θ_{jt} to infer the value of V_{njt}. It would be straightforward to modify (1) in a way that rules out this possibility, for example by assuming that $\hat{u}_{njt} \equiv \theta_{jt} + (1 - e^{-\phi_{it}\eta_{it}})V_{njt}$ with η_{it} unobserved.

[17] Qualitatively identical results hold under more general assumptions regarding the distribution of ϵ_{njt}; the Type-1 Extreme Value assumption is thus imposed here only for expositional simplicity.

which also depends on ϕ_{it}, as does the expected number of unique drugs M_{it} prescribed by physician i during t,

$$E_t[M_{it}] \equiv E_t\left[\sum_{j=1}^{J_t} 1\{X_{ijt} > 0\}\right] = \sum_{j=1}^{J_t} P_{ijt}(\phi_{it}) = \sum_{j=1}^{J_t} \left(1 - (1 - p_{jt}(\phi_{it}))^{N_{it}}\right). \tag{4}$$

Suppose that each physician determines her optimal effort ϕ_{it} spent acquiring information by weighing the benefit of reduced prescribing errors against the cost of her effort. For ease, suppose that each physician faces a cost $C(\phi_{it})$ of prescribing errors that increases linearly in the sum of squared errors in beliefs across all patients n and drugs j during period t

$$C(\phi_{it}) \equiv c \sum_{n=1}^{N_{it}} \sum_{j=1}^{J_t} (u_{njt} - \hat{u}_{njt})^2 = c\ e^{-2\phi_{it}} \sum_{n=1}^{N_{it}} \sum_{j=1}^{J_t} (v_{jt} + \epsilon_{njt})^2 = c\ e^{-2\phi_{it}} \mathbf{V}_{it},$$

where c is a cost parameter, and the aggregate $\mathbf{V}_{it} \equiv \sum_{n=1}^{N_{it}} \sum_{j=1}^{J_t} (v_{jt} + \epsilon_{njt})^2$ is assumed known by physician i; this aggregate reflects the number of physician-i patients and the sample distributions (mean and variance) of v_{jt} and ϵ_{njt}, respectively, among these patients. Assume further an i-specific marginal cost of acquiring information $a_{it} < c$. That is, assume physician i faces a cost $A_{it}(\phi) = a_{it}\phi$ to exert effort ϕ. It is simple to show that, in this setting, physician i optimally seeks information with an intensity

$$\phi_{it}^* = \frac{1}{2} \ln\left(\frac{2c\,\mathbf{V}_{it}}{a_{it}}\right). \tag{5}$$

According to (5) above, the introduction of an improved drug search technology that reduces a_{it} across all physician users induces an increase in the optimal search intensity ϕ_{it}^*.[18,19] From (2), this increase in ϕ will result in a change in the probability drug j is prescribed: whether P_{ijt} increases or decreases for drug j depends on the distribution of v_{jt} across products. In general, P_{ijt} will increase for drugs with high values of v_{jt} relative to other drugs; alternatively, if all $v_{jt} = 0$, an increase in ϕ raises P_{ijt} for all drugs except for that with the highest θ_{jt}. Similarly, (3) implies the number of periods that pass before drug j is

[18] Although for simplicity we have not made this explicit, it is straightforward to introduce an adoption friction for use of the improved search technology such as a one-time fixed adoption cost. In this case, physicians for whom the marginal benefit of reduced medical errors is relatively high—that is, physicians with high V_{it} values (i.e. with many, relatively diverse patients)—are more likely to adopt than others. This implies that some physicians optimally choose not to adopt the database, while others find it optimal to delay adoption until V_{it} is sufficiently high. This also raises the important question of selection that is addressed in the empirical analysis below.

[19] By modeling an improvement in drug search technology as a reduction in a_{it}, notice that this simple model is consistent with the idea that information obtained by a physician from a new drug reference tool is useful, but only incremental relative to the knowledge and experience the physicians have already obtained about the drugs from other sources. An important empirical implication of this is that a decline in a_{it}, and the increase in ϕ_{it} that results, is unlikely to cause large shifts in prescribing.

prescribed declines in ϕ_{it} whenever P_{ijt} increases in ϕ_{it}. The impact of an increase in ϕ_{it} on the number of distinct drugs prescribed (4) depends on the distribution of V_{njt} across drugs j and patients n, but in general, a higher ϕ implies increased sensitivity to patient-specific match quality ϵ_{njt}, which tends to increase the diversity of prescribing.

It is important to note that doctors who regularly prescribe anti-cholesterol medications will be aware of most drugs' clinical attributes. But for newer, less familiar drugs, doctors may prefer to look up details like dosage and patient-specific economic details such as pricing and formulary status prior to writing a prescription. It is for these newer drugs that one may expect differences between u_{njt} and $\hat{u}_{njt}$ to be particularly relevant, and the empirical analysis thus considers the distinction between new and existing products in this respect.

B. *Estimating Equations*

One natural approach to evaluating the influence of reductions in a_{it} resulting from database adoption would be to directly estimate equations derived from the conceptual model above. However, even within the model above, it is important to control for unobserved covariates including drug quality θ_{jt} and physician search costs a_{it}. Given the size of the data sample, handling the nonlinearity implied by (2) in the presence of multiple sets of fixed effects is computationally infeasible. For this reason, we instead estimate the effects of database adoption on prescribing primarily through linear regression equations. An advantage of this approach is that it enables us to proceed without relying on a strong distributional assumption for the ϵ_{njt} terms.

We thus consider three estimating equations corresponding to the three outcomes discussed above (T_{ij}, M_{it}, and P_{ijt}). We first assess the impact of information access on the duration of time T_{ij} that elapses between the initial market release of drug j and its first prescription by physician i as in Coleman, Katz, and Menzel (1957) using the following specification

$$\log T_{ij} = \eta_j + \eta_i + \beta_0 Z_{ij} + \epsilon_{ij}, \tag{6}$$

in which Z_{ij} is an indicator for database use at the time drug j is first introduced, and where η_j and η_i are product and physician fixed effects, respectively. Equation (6) is estimated on the subset of drugs first introduced during the sample period. Finding that the coefficient of interest β_0 is negative would indicate that when a physician obtains information access (a decline in a_{it}) she prescribes new drugs significantly more quickly than previously, relative to a physician without database access. Notice that the inclusion of physician fixed effects implies that the coefficients β are identified using within-doctor variation over time: physician i may be a database user at the time drug j is first introduced, but may not yet be a user at the time another product j' is introduced. These fixed effects are important if stable,

unobserved physician characteristics determine both physician-specific database use Z_{ij} and the rate of new drug adoption T_{ij} (e.g. early adopters).

Second, building from (4), we consider the possibility that information access could affect physician i's knowledge of the match quality between drug j and patient n, inducing better-informed physicians to prescribe a more diverse set of products than less-informed peers. To assess this possibility, we determine the number of unique drug products $M_{it} \equiv \sum_{j \in \mathcal{J}_t} 1\{\mathrm{X}_{ijt} > 0\}$, where $1\{\mathrm{X}_{ijt} > 0\}$ is an indicator for whether physician i writes at least one prescription for drug j during month t, and evaluate the following specification

$$M_{it} = \eta_t + \eta_i + \beta_0 Z_{it} + \epsilon_{it}, \tag{7}$$

where η_t is a month fixed effect, and all other variables are as defined above. M_{it} is low when the prescriptions of physician i are concentrated within a narrow subset of products during month t, and is high when prescribing is diverse; finding that β_0 is positive in (7) above would thus indicate that information access is associated with higher product diversity among physician i's prescriptions. We also estimate (7) replacing M_{it} with the Herfindahl-Hirschman index as an alternative dependent variable for each physician i-month t pair. Notice that (7) includes physician fixed effects η_i and thus controls for any stable, unobserved physician characteristics that affect prescription diversity and may also be correlated with information access Z_{it}. The coefficient β_0 is thus identified using within-physician variation over time in information access Z_{it}. To account for changes over time in unobserved, location-specific patient characteristics, we also estimate (7) with zipcode-month fixed effects; these are particularly important if patient characteristics or other local factors evolve in ways that affect prescribing and are correlated with measured physician technology adoption.

While (7) captures the scope of physician prescribing, it is also of interest to understand changes in the composition of prescriptions across drugs—and in particular, to understand how database users' P_{ijt} values across new and old drugs j differ after database adoption. Moreover, because new patent-protected products differ from new generics in both cost and novelty, the impact of information on prescribing may differ based on the patent status of a new product. We thus evaluate whether physicians using the electronic drug database are also more likely to prescribe specific product types using the following specification

$$\begin{aligned} 1\{\mathrm{X}_{ijt} > 0\} = \eta_{jt} + \eta_{ij} + \beta_0 Z_{it} \times New^{\tau}_{jt} \times Gen_j + \beta_1 Z_{it} \times New^{\tau}_{jt} \times (1 - Gen_j) \\ + \beta_2 Z_{it} \times (1 - New^{\tau}_{jt}) \times Gen_j + \beta_3 Z_{it} \times (1 - New^{\tau}_{jt}) \times (1 - Gen_j) + \epsilon_{ijt}, \end{aligned} \tag{8}$$

where $1\{\mathrm{X}_{ijt} > 0\}$ is an indicator for whether physician i writes at least one prescription for drug j during month t, Gen_i is an indicator that is equal to 1 if product i is a generic variety, and New^{τ}_{jt} indicates whether drug j is within 24 months of its initial approval for

U.S. sale. The main coefficients of interest β_0, β_1, β_2, and β_3 jointly capture the influence of database use Z_{it} on prescription choice for both new drugs ($\beta_0 + \beta_1$) and established products ($\beta_2 + \beta_3$), where finding $\beta_0 + \beta_1 > 0$ would indicate physicians with better access to information are more likely to prescribe a given drug j that is within τ months of initial market release. Finding that $\beta_0 > 0$ would indicate that physicians with better access to information ($Z_{it} = 1$) are more likely to prescribe a new, generic product j at t relative to other physicians; $\beta_1 > 0$ would indicate an analogous effect for other new drugs.

Equation (8) includes two sets of controls. Physician-drug fixed effects η_{ij} absorb any physician-specific characteristics that affect prescribing differentially across drugs j such as location, patient composition, age, education, and medical specialty. These effects are thus able to account for the possibility that patients in some locations are price-sensitive and tend to prefer a generic product when available, or that a physician prefers prescribing a particular drug based on the history of patient experience known to the physician. Importantly, these effects also capture underlying differences across physicians in the tendency to prescribe new products, first introduced during the sample period. The coefficients of interest $\boldsymbol{\beta}$ are thus identified primarily from within doctor-drug variation over time in information access Z_{it} and drug status New_{jt}^{τ}. Drug-month fixed effects η_{jt} further account for the average perceived quality of drug j by sample physicians at t, which may depend on factors such as drug potency and side effects known at t and the average pharmacy price faced by consumers at t. The error term ϵ_{ijt} combines any omitted factors that affect physicians' prescribing patterns.

IV. Main Results

A. *Time to First Prescription*

Estimates of equation (6) appear in Table 4. Columns 1–4 evaluate the relationship between physician-i information access and the time lapse T_{ij} between the market introduction of a new drug j and its initial prescription by i for the full sample of U.S. prescribers. Columns 5–6 replicate columns 3 and 4 but restrict the physician sample to include only those who register for the database and use it to look up at least one cholesterol drug during the sample period; physicians included in these latter specifications thus differ only in their respective database adoption dates. The independent variable Information (Z_{ij}) is specific to each physician-drug pair and takes a value of 1 if physician i has access to the electronic drug reference at the time drug j is approved for sale in the U.S. market, and is otherwise zero; Information x Generic interacts Z_{ij} with an indicator for generic products. All columns include drug fixed effects that account for differences in product characteristics such as quality, average physician reference access at the date the drug was approved for sale, and the set of competing products available for prescription at that approval date; physician fixed

effects are also included in columns 2, 4, and 6, which thus rely on within-physician changes in database access across drug introductions for identification. These latter specifications therefore account for the persistent component of a physician's prescription volume, tendency to adopt new technology, and patient characteristics, all of which may impact the timing of new drug adoption independently of Z_{ij}.

The estimated coefficients on Z_{ij} in columns 1 and 2 are negative and highly significant, indicating physicians using the electronic drug database are, on average, significantly faster to begin prescribing a newly-approved drug than physicians not using the database. The magnitude of this effect suggests physician users write initial prescriptions of a new drug seven percent sooner than non-users in column 1 and 2.5 percent sooner when controlling for fixed, unobserved physician characteristics in column 2. As the adoption delay T_{ij} is 19.14 months on average (Table 2), these estimates imply database users adopt a new drug 1.33 months (column 1) or 0.48 months (column 2) sooner, on average, than non-users.

The estimates in columns 3 and 4 indicate these effects are driven primarily by prescriptions for newly-introduced generic products. Specifically, with prescriber fixed effects in column 4, physicians using the database are significantly faster (six percent) to prescribe new generic products; by contrast, the coefficient on branded products is insignificant, suggesting users and non-users are indistinguishable in their adoption rate for new branded products, once controlling for physician fixed effects. The results in column 6, which reduce the possible influence of selection by including only eventual database users, confirm this result for generics, with physicians users adopting new drugs 3 percent more quickly, on average, than prior to database adoption. Given that generic drugs share identical clinical attributes with branded versions, an explanation for these effects is that the drug database includes both retail price and insurance formulary information, providing physicians with patient-specific cost information at the point of care, and potentially influencing the choice of drug prescribed in favor of generics. Physician users in column 6 also appear to delay prescribing new branded products, though the coefficient is small and only weakly significant.

B. *Prescription Diversity*

To assess the possible relationship between database use and prescription diversity, we evaluate specification (7); results appear in Table 5. Estimates for the complete sample of U.S. physicians appear in columns 1–3. These estimates indicate that database users prescribe, on average, a significantly larger number of distinct drugs per month than physician non-users. In terms of empirical magnitudes, the least-squares estimate of β_0 in column 1 indicates that a database user prescribes 0.25 additional drug varieties each month relative to a non-user, a result that is confirmed qualitatively in a Poisson specification (column 2). Column 3 considers the prescription HHI as an alternative dependent variable, and indicates

that the concentration of prescribing across drugs is significantly lower among database users.

Columns 4–6 restrict the physician sample to include only eventual database users, as in Table 4 above; physicians included in these latter specifications thus differ only in their respective drug reference adoption dates. Estimates in columns 4–6 are small in magnitude relative to columns 1–3, suggesting that the extent of within-physician increases in prescription diversity over time may be positively correlated with database adoption during the sample period. This possibility would be consistent with a version of the model described in Section 3 above featuring a database adoption friction such as a one-time fixed setup cost. Nevertheless, qualitatively identical results obtain across the two physician samples, with prescription diversity increasing significantly when physicians adopt the drug database.

C. *The Composition of Prescriptions*

Estimates of the propensity equation (8) above appear in Table 6; the main coefficients of interest β_0 and β_1 correspond to physician-i information access Z_{it} in month t, and its separate interactions with an indicator for new generic and new branded products, respectively. New drugs are defined as those products that are within $\tau = 24$ months of initial market approval in month t. The estimated coefficient corresponding to new generic products $\hat{\beta}_0$ is positive and highly significant in column 1, which includes the full sample of physicians. This is consistent with information technology encouraging the prescription of new generic drugs. However, though the coefficient is an order of magnitude smaller, the estimate of β_1 in column 1 indicates that physician database users are also somewhat less likely to prescribe a new branded product. This is consistent with the estimates in Table 4, which indicate that in the presence of physician fixed effects, physician database users are significantly faster to begin prescribing a new generic drug, but slightly slower for new branded products. Column 1 also indicates database users are significantly more likely to prescribe old generic products, and to a mild extent, old branded products; the estimated coefficients β_2 and β_3 for old drugs are both positive in specification 1.

As in Tables 4 and 5, column 2 restricts the sample of prescribers to those who eventually adopt the drug database. These estimates indicate that physicians using the database are considerably more likely to prescribe drugs in each category, consistent with the estimates in Table 5. However, while Table 5 indicates a tendency toward increased prescription diversity among database users relative to non-users, Table 6 reveals that this does not reflect a symmetric shift in prescribing across drug products. Physician users are substantially more likely to prescribe generic products, but are only moderately more likely to prescribe branded drugs as coefficients for new and old generics are substantially larger than those for branded products: $\hat{\beta}_2 > \hat{\beta}_0 >> \hat{\beta}_3 > \hat{\beta}_1$. Specifically, while a physician with access to the electronic drug reference is only 0.25 percent more likely to prescribe an branded product that is within

24 months of approval, she is more than six times as likely (1.61 percent) to prescribe a new generic product within the same timeframe, relative to a physician without information access. But she is even more likely (2.28 percent) to prescribe an existing generic product.

Importantly, all specifications in Table 6 include physician-drug fixed effects, which capture stable physician-specific differences in the tendency to prescribe certain drugs, whether new, generic, or both. In addition, the drug-month fixed effects included in each column account for drug-specific changes in pricing, perceived quality, available information, and advertising, Further including zipcode-month fixed effects to account for location-specific changes over time in drug advertising, internet connectivity, policies and pricing, insurance generosity, income, health, demographics and so on results in nearly identical estimates.[20]

V. Interpreting the Results

The results in Tables 4 through 6 above indicate that physicians using the database begin prescribing new drugs, particularly new generics, sooner than non-users, and also prescribe a more diverse set of drugs than non-users. However, because database adoption is not randomly assigned, it remains unclear whether a causal interpretation of these results is supported by the data. Our baseline strategy of including physician (or physician-drug) fixed effects to account for stable unobservables, correlated with both database adoption and prescribing, helps rule out certain alternative explanations including cases in which 'early-adopter' physicians begin both using the database and prescribing a new drug sooner than other physicians. A limitation of this approach, however, arises when adoption is either correlated with, or involves selection on, time-varying rather than stable physician characteristics that are relevant to the prescribing outcomes we consider. In Table 5, for example, it is possible that a physician's decision to adopt the database is partially determined by the rate of increase in her prescribing diversity: naturally, adopting the database today could be a more attractive option for a physician who anticipates prescribing a wider range of products in the future, than for a physician in the opposite situation.

One approach to handling such cases is to find an instrument that generates quasi-random variation in database adoption, and to estimate relying on variation in this instrument. We have considered two such instruments: 1) a measure of location-specific hospital I.T. use from Dranove et al (2014), and 2) a measure of location-specific high-speed internet penetration; both are factors that could influence doctors' database adoption decisions while being plausibly unrelated to choices over which anti-cholesterol drugs to prescribe. However, we find these instruments to be only weak predictors of database adoption, resulting in second-stage estimates highly sensitive to small specification changes. Thus lacking compelling instruments, we take an alternative approach: specifically, we consider placebo tests

[20]Results available on request.

and split-sample tests that help to determine whether non-causal interpretations of the results are consistent—or instead inconsistent—with the data.

A. *Placebo Tests*

An important feature of the data is that we observe not only a physician's database registration date, but also indicators for the extent of physicians' database use. In particular, we observe an indicator variable for whether a registered physician uses the reference for a cholesterol drug at least once during the sample period. Using this indicator, we find that the data include a large set of 'placebo' physicians that adopt the database, but never use it to search for information about the drugs we consider. There is no reason to expect the database to affect prescription outcomes for such physicians in Tables 4 through 6.

Tables 7, 8, and 9 thus report results from regressions analogous to those reported in Tables 4, 5, and 6, with effects estimated separately for actual and 'placebo' database users. In Table 7, the results for the placebo users are statistically indistinguishable from zero in all but one case, and in every case the point estimates of the effects are near zero. By contrast, the results for actual users closely resemble those reported in Table 4. This suggests that the estimated effects for actual users in Table 7 cannot be dismissed as reflections of endogenous selection: if the results are driven by a correlation between database adoption and a prescribing pattern or trend that affects prescribing broadly, in a way not specific to cholesterol drugs, that correlation would affect both actual and 'placebo' users. To ascribe the results in Table 7 to endogenous selection, one would, in this case, need to contend that selection effects are systematically related not only to a physician's database usage, but to a physician's database usage specifically for cholesterol drugs.

Table 8 considers the diversity of physicians' prescriptions. Although effects for placebo users are smaller than for actual users, they are nevertheless statistically significant. These non-zero estimates suggest database adoption may indeed by correlated with some factor related to changes in the diversity of a doctor's prescriptions. However, the strictest of these tests in columns 4–6 indicate that, while actual users' prescription diversity increases with adoption, placebo users' diversity falls. This distinction is consistent with the idea that adopters that actually use the database to look up cholesterol drugs increase the diversity of prescribing beyond what could potentially be explained by selection.

Similarly, the estimates in Table 9 for placebo users are not statistical zeroes, but in most cases are smaller in magnitude than the estimates for actual users. For example, the largest estimated effects for database users are the 1.64 and 2.38 percentage-point increases in the likelihood of prescribing new generics and old generics, respectively. For placebo users, the corresponding estimates are 1.34 and 1.29 percentage points—smaller by roughly a factor of 2 in the latter case. Thus, as in Table 8, the difference in estimated magnitudes suggests it

is plausible that actual users' database access has a causal impact on prescribing, beyond what could be explained by selection.

Taken together, these placebo tests provide sharper evidence about the impact of the database, relative to Section IV above. While the results in section IV are robust to concerns of selection on stable physician characteristics, they are subject to the caveat that the timing of database adoption may be correlated with, or even driven by, anticipated changes in a physician's prescribing. The placebo tests indicate that this is unlikely to be a concern, provided that the anticipated changes are general across drug products rather than specific to cholesterol drugs. In this case, all physicians adopting the database are affected by the same selection forces, but a causal effect should only be present for physicians who actually use the database to look up cholesterol drugs. The results in Table 7 are consistent with this. However, the placebo tests cannot rule out the possibility that database adoption is driven by anticipated changes that are specific to cholesterol drugs. Thus the placebo tests rule out some, but not all, non-causal alternative interpretations of our results.

B. *Mandatory Substitution Laws*

To encourage cost savings, many U.S. states impose regulations mandating generic substitution where available; in most cases, such laws have been in force since the 1970s (Grabowski and Vernon 1979) and were thus in effect during the sample period. This observation raises the possibility that, in the case of generics, a physician may write a prescription for the branded drug, yet her patient may receive its generic form. If the implementation of mandatory substitution laws differs across states and over time in a manner correlated with physicians' adoption of the drug database, it could affect the interpretation of our results. We address this possibility in two ways. First, we re-evaluate results on prescription diversity (Table 5) and drug-specific prescribing propensity (Table 6) including zipcode-month fixed effects that control for the possibility of differential substitution regulations across observations. Even without these fixed effects, notice that mandatory substitution laws are likely to affect prescribing similarly for users and non-users of the drug database, and our baseline specifications therefore account for this consideration. Second, to capture this possibility in the duration regressions (Table 4), we adopt a split-sample approach, re-estimating Table 4 separately for states with, and states without, mandatory substitution laws in for at the start of the sample period in January 2000. The estimates appear in Table 10 and reveal no substantive differences across the two samples; it therefore appears unlikely that mandatory substitution laws are able to explain the results.

C. *Academic Medicine*

Physicians practicing in locations known for pharmaceutical innovation may have access to frontier knowledge regarding pharmaceutical development, limiting the scope for use of the condensed drug information database we observe to influence prescribing decisions; within the conceptual framework outlined in section 3, such physicians may have near-zero information costs a_{it}, and database adoption has, correspondingly, little potential to impact prescribing outcomes. As innovation tends to cluster geographically (e.g. Sohn 2016), adding zipcode-month fixed effects to the specifications in Tables 5 and 6 should help to account for this. To more directly evaluate the potential influence of location-specific differences in innovativeness, we rely on patent data from the NBER U.S. Patent Citations Data File (Hall, Jaffe, and Trajtenberg 2001). These data specify, for each U.S. patent, the location of the inventor as well as detailed information about the patented technology. We use information on the number of pharmaceutical patents granted between 1975 and 1999 by zipcode to identify U.S. locations in the top and bottom five percent by drug patenting, and then re-estimate the duration analysis in Table 4 separately for these two samples. The estimates appear in Table 11, and suggest that physicians plausibly located near the knowledge frontier—that is, physicians in zipcodes among the top five percent by drug patenting—indeed respond to drug information differently than their more distant peers. Specifically, the estimates in column 4 indicate that while use of the database in non-innovative locations is associated only with faster adoption of new generics, the database is only weakly influential among innovative frontier locations and affects impact branded and generics similarly (column 8).

D. *Other robustness checks*

To account for the possibility of spatially-correlated measurement error, which could arise due to the pharmacy sampling methodology used by IMS Health to build the prescription data, we re-estimate the results in Tables 4 and 5 with standard errors clustered by zipcode. We also estimate the specifications in Tables 4 and 5 using Poisson or Negative Binomial regressions, and we estimate the specifications of Table 6 with logistic fixed effects regressions. These robustness checks reveal qualitatively similar results.[21]

As noted above, we also estimate the regressions in Tables 5 and 6 with zipcode-month fixed effects, which help rule out concerns that pharmaceutical detailing could be influencing doctors' prescriptions (e.g. Datta and Dave 2013). If the influence of detailing differs across drugs but is similar across physicians at a point in time, then the drug-month dummies included in Table 6 would account for it. If, by contrast, the intensity of drug detailing differs across physician-drug pairs but is relatively stable over time, then the physician-drug fixed effects in Table 6 would account for it. Zipcode-month fixed effects would account for broad differences across locations over time. Our estimates indicate that that the results

[21] Detailed results available on request.

are robust to including zipcode-month fixed effects, which have essentially no impact on the coefficients of interest. This is not surprising: our results our unlikely to be driven by detailing, because there is little reason to expect detailing to be correlated with database adoption. Also, the consistent result emerging from Tables 4 and 6 is that the impact of database use appears largest for generic products, which tend not to benefit from detailing.

The fact that the clearest effects are for generic drugs raises a separate possibility, however. A physician's decision to prescribe a generic drug may be related to the insurance coverage of the patient. We therefore evaluate split-sample estimates based on whether physicians receive a high or low share of Medicare and Medicaid patients, relative to the privately insured; separately, we repeat this split-sample analysis, distinguishing physicians based on whether a high or low share of their patients pay for prescriptions with cash. In both cases, we found negligible differences across groups.

VI. Physician Heterogeneity

As a final point, we consider whether the data support the idea that incomplete information may contribute to disparities in care across U.S. physicians. To evaluate this, we assign each physician to one of two groups based on her drug database registration status in December 2010. Within each group, we measure the extent of prescribing heterogeneity across physicians: specifically, we determine the unit vector of prescriptions written in December 2010 for each prescriber i. We then compute the distance between this physician-i unit vector and either a) the average unit vector among physicians using the database, if i is a database user, or if not, b) the average unit vector among physicians not using the database. If physician users more closely resemble eachother in prescribing than do nonusers, it would be consistent with information frictions contributing to prescribing heterogeneity; if not, it would suggest disparities are primarily the result of other factors.

Corresponding results appear in Table 12, Panel A. These indicate that the distribution of prescriptions among physicians using the database is more compressed than among non-users, in line with the idea that information reduces prescribing heterogeneity. Specifically, the average Euclidean distance between the physician-i unit vector and the group-specific average is higher for non-users (0.1762) than for users (0.1522), and the estimated difference in column 3 (-0.0236) is highly significant. Importantly, note that the results in Table 5 indicate physicians using the database prescribe a significantly more diverse set of products than non-users, so this does not imply a loss of variation in therapies generally. Rather, physicians resemble eachother more closely when connected to the same information source.

Replicating this exercise with prescription data from January 2000 and an indicator for database use in December 2010—applied prematurely, as though eventual users had already adopted in January 2000—in Panel B reveals two important observations. First, both groups

(users and non-users) experience considerable homogenization over time, with the average Euclidean distance declining by 0.040 points for database non-users and by 0.052 percentage points for users. Second, based on prescribing in January 2000, physicians that only later register as database users already exhibit greater homogeneity in prescribing than other physicians, suggesting database users form a selected sample in this respect, and indicating the importance of controlling for pre-existing physician characteristics. With this in mind, the estimates in Panel C indicate that despite these pre-existing differences, eventual database users undergo significantly more within-group homogenization than non-users, even when controlling for physician fixed effects.[22]

VII. Discussion and Concluding Remarks

This paper has examined whether access to product information encourages physician adoption of new cholesterol control drugs. Using data on prescriptions and electronic drug reference use for over 130,000 individual U.S. physicians and 18 products during January 2000—December 2010, our results indicate that physicians with access to pharmaceutical product information at the point of care prescribe a significantly more diverse set of products than other doctors, and first prescribe new products—particularly new generics—sooner than other doctors. These results hold after controlling for stable unobserved physician characteristics that may influence adoption and prescribing decisions. Such doctors are also, at any time, more likely to prescribe a new product that is within 24 months of its initial release, particularly if it is generic, than are other doctors.

Taken together, our results support the hypothesis that physicians' use of the electronic database caused a small but statistically significant shift in their prescribing patterns. That the effects were small should not be surprising: physicians who prescribe cholesterol drugs are generally well-informed about the therapeutic options, and adopting the database therefore did not present a significant shock their information sets. However, our results indicate that even such incremental changes in physicians' access to information impact their behavior. And, given the size of the market for cholesterol drugs, even small shifts in physicians' prescribing patterns can have substantial effects on aggregate drug costs.

These findings thus speak to policy debates regarding U.S. healthcare provision, particularly those concerning wide variation in the observed cost and quality of medical care across U.S. locations (Wennberg et al 1996). Specifically, our estimates suggest that uneven access to information may contribute to heterogeneity in U.S. healthcare provision and costs. In this, an implication of our findings is that connecting physicians to electronic information

[22]For clarity, Panel C reports coefficients from a least-squares regression of the Euclidean distance to the mean D_{it} for doctor i at $t = \{\text{January 2000, December 2010}\}$ on an indicator I_{2010} for December 2010, and its interaction $Z_{2010} \times I_{2010}$ with an indicator $Z_{i,2010}$ for physician-i database access in December 2010, and physician fixed effects.

resources has the potential to reduce disparities in care. Importantly, this reduction need not come at the expense of careful medical decision-making: indeed, our observation that physicians with access to electronic product information prescribe a more diverse set of products suggests a closer match between patient characteristics and available therapies.

REFERENCES

[1] **Agha, Leila.** 2012. "The Effects of Health Information Technology on the Costs And Quality of Medical Care." Boston University mimeo.

[2] **Agha, Leila and David Molitor.** 2015. "The Local Influence of Pioneer Investigators in Technology Adoption: Evidence from New Cancer Drugs." Boston University mimeo.

[3] **Arrow, Kenneth J.** 1969. "Classificatory Notes on the Production and Transmission of Technological Knowledge." *American Economic Review, Papers and Proceedings*, vol. 59, no. 2, pp. 29–35.

[4] **Athey, Susan, and Scott Stern** 2002. "The Impact Of Information Technology On Emergency Health Care Outcomes." *RAND Journal of Economics*, vol. 33, no. 3, pp. 399–432.

[5] **Attewell, Paul.** 1992. "Technology Diffusion and Organizational Learning: The Case of Business Computing." *Organizational Science*, vol. 3, no. 1, pp. 1–19.

[6] **Berndt, Ernst R., Robert S. Gibbons, Anton Kolotilin, and Anna Levine Taube.** 2015. "The Heterogeneity of Concentrated Prescribing Behavior: Theory and Evidence from Antipsychotics." *Journal of Health Economics*, vol. 40, pp. 26–39.

[7] **Black, Sandra E., and Lisa M. Lynch.** 2001. "How To Compete: The Impact of Workplace Practices and Information Technology on Productivity." *Review of Economics and Statistics*, vol. 83 no. 3, pp. 434–445.

[8] **Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen.** 2009. "The Distinct Effects of Information Technology and Communication Technology on Firm Organization." National Bureau of Economic Research (NBER) Working Paper 14975.

[9] **Bloom, Nicholas, Raffaella Sadun, and John Van Reenen.** 2012. "Americans Do IT Better: US Multinationals and the Productivity Miracle." *American Economic Review*, vol. 102, no. 1, pp. 167–201.

[10] **Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt.** 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics*, vol. 117, no. 1, pp. 339–76.

[11] **Bresnahan, Timothy, and Shane Greenstein.** 1996. "Technical Progress and Co-invention in Computing and in the Uses of Computers." Brookings Papers on Economic Activity Microeconomics: pp. 1–8.

[12] **Brooks, John M., Elizabeth A. Cook, Cole G. Chapman, et al.** 2014. "Geographic Variation in Statin Use for Complex Acute Myocardial Infarction Patients." *Medical Care*, vol. 52, no. 3 Suppl 2, pp. S37-S44.

[13] **Brynjolfsson, Erik, and Lorin M. Hitt.** 2003. "Computing Productivity: Firm-Level Evidence." *Review of Economics and Statistics*, vol. 85, no. 4, pp. 793–808.

[14] **Cannon, Christopher P., Eugene Braunwald, Carolyn H. McCabe, et al.** 2004. "Intensive Versus Moderate Lipid Lowering with Statins After Acute Coronary Syndromes." *New England Journal of Medicine*, vol. 350, no. 15, pp. 1495–1504.

[15] **Cannon, Christopher P., Michael A. Blazing, Robert P. Giugliano, et al.** 2015. "Ezetimibe Added to Statin Therapy After Acute Coronary Syndromes." *New England Journal of Medicine*, vol. 372, no. 25, pp. 2387–2379.

[16] **Chernew, Michael E., Richard A. Hirth and David M. Cutler.** 2009. "Increased Spending On Health Care: Long-Term Implications For The Nation." *Health Affairs*, vol. 28, no. 5, pp. 1253–1255.

[17] **Coleman, James, Elihu Katz, and Herbert Menzel.** 1957. "The Diffusion of an Innovation Among Physicians." *Sociometry*, vol. 20, no. 4, pp. 253–270.

[18] **Coleman, James, Elihu Katz, and Herbert Menzel.** 1966. *Medical Innovation: A Diffusion Study.* New York: Bobbs-Merrill Co.

[19] **Comin, Diego and B. Hobijn.** 2004. "Cross-Country Technology Adoption: Making Theory Face the Facts." *Journal of Monetary Economics*, vol. 51, pp. 39–83.

[20] **Cooper, Zack, Stuart Craig, Martin Gaynor, and John Van Reenen.** 2015. "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured." Yale University mimeo.

[21] **Crawford, Gregory S. and Matthew Shum.** 2005. "Uncertainty and Learning in Pharmaceutical Demand." *Econometrica*, vol. 73, no. 4, pp. 1137–1173.

[22] **Datta, Anusua and Dhaval M. Dave.** 2013. "Effects of Physician-Directed Pharmaceutical Promotion on Prescription Behaviors: Longitudinal Evidence." National Bureau of Economic Research (NBER) Working Paper 19592.

[23] **David, Paul A.** 1966. "The Mechanization of Reaping in the Ante-Bellum Midwest," Chapter 1 of H. Rosovsky, ed., *Industrialization in Two Systems: Essays in Honor of Alexander Gerschenkron*, New York: Wiley and Sons.

[24] **Dickstein, Michael J.** 2015. "Efficient Provision of Experience Goods: Evidence from Antidepressant Choice." New York University mimeo.

[25] **Dranove, David, Chris Forman, Avi Goldfarb, and Shane Greenstein.** 2014. "The Trillion Dollar Conundrum: Complementarities and Health Information Technology." *American Economic Journal: Economic Policy*, vol. 6, no. 4, pp. 239–270.

[26] **Duggan, Mark and Fiona Scott Morton.** 2010. "The Effect of Medicare Part D on Pharmaceutical Prices and Utilization." *American Economic Review*, vol. 100, no. 1, pp. 590–607.

[27] **Escarce, Jose.** 1996. "Externalities in Hospitals and Physician Adoption of A New Surgical Technology: An Exploratory Analysis." *Journal of Health Economics*, vol. 15, pp. 715–734.

[28] **Forman, Chris, Avi Goldfarb, and Shane Greenstein.** 2005. "How Did Location Affect the Adoption of the Commercial Internet? Global Village vs. Urban Leadership." *Journal of Urban Economics*, vol. 58, no. 3, pp. 389–420.

[29] **Fuchs, Victor R.** 1974. "Who Shall Live?" Basic Books, New York.

[30] **Gawande, Atul**. 2009. "The Cost Conundrum," *The New Yorker*. June 1.

[31] **Grabowski, Henry G. and John M. Vernon.** 1979. "Substitution Laws and Innovation in the Pharmaceutical Industry." *Law and Contemporary Problems*, vol. 43, no. 1, pp. 43–66.

[32] **Griliches, Zvi.** 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica*, vol. 25, no. 4, pp. 501–522.

[33] **Hall, Bronwyn H., Adam B. Jaffe, and Manuel Tratjenberg.** (2001). "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools," NBER Working Paper 8498.

[34] **Hubbard, Thomas N.** 2003. "Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking." *American Economic Review*, vol. 93, no. 4, pp. 1328–53.

[35] **Kastelein, John J., Fatima Akdim, Erik S. Stroes, et al.** 2008. "Simvastatin With or Without Ezetimibe in Familial Hypercholesterolemia." *New England Journal of Medicine*, vol. 358, no. 14, pp. 1431–1443.

[36] **Klenow, Peter and Andres Rodriguez-Clare.** 1997. "Economic Growth: A Review Essay." *Journal of Monetary Economics*, vol. 40, pp. 597–618.

[37] **Law, M.R., N.J. Wald, and A.R. Rudnicka.** 2003. "Quantifying Effect of Statins on Low Density Lipoprotein Cholesterol, Ischaemic Heart Disease, and Stroke: Systematic Review and Meta-Analysis." *British Medical Journal*, vol. 326, pp. 1–7.

[38] **Munson, Jeffrey C., Nancy E. Morden, David C. Goodman, Luca F. Valle, and John E. Wennberg.** 2013. *The Dartmouth Atlas of Medicare Prescription Drug Use*, American Hospital Association Press, *at* http://www.dartmouthatlas.org.

[39] **Parente, Stephen L. and Edward C. Prescott.** 1994. "Barriers to Technology Adoption and Development." *Journal of Political Economy*, vol. 102, no. 2, pp. 298–321.

[40] **Skinner, Jonathan and Douglas Staiger.** 2007. "Technology adoption from hybrid corn to beta-blockers." In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches.* NBER Chapters, National Bureau of Economic Research, Inc, pp. 545–570.

[41] **Sohn, Emily.** 2016. "The Geography of Discovery." *Nature*, vol. 533, pp. S40–S43.

[42] **Solow, Robert M.** 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics*, vol. 70, no. 1, pp. 65–94.

[43] **Solow, Robert M.** 1987. "We'd Better Watch Out." *New York Times*, July 12, 36.

[44] **Wennberg, John E. and Megan M. Cooper, eds.** 1996. *The Dartmouth Atlas of Health Care*, American Hospital Association Press, *at* http://www.dartmouthatlas.org.

Table 1: Heterogeneity in Initial Prescription of a New Pharmaceutical Drug

Drug Name	Release Date	FDA Category	Mean	SD	Final Share
Lescol XL	October 2000	New dosage form	28.856	23.486	0.628
Advicor	December 2001	New combination	64.752	15.434	0.298
lovastatin	December 2001	New generic version	19.914	22.688	0.936
Altoprev	June 2002	New dosage form	42.504	24.057	0.154
Zetia	October 2002	New molecular entity	15.181	17.384	0.941
Pravigard PAC	June 2003	New combination	7.328	5.954	0.038
Crestor	August 2003	New molecular entity	22.668	21.823	0.936
Vytorin	July 2004	New combination	14.122	13.498	0.903
Lovaza	November 2004	New molecular entity	34.039	17.075	0.670
pravastatin	April 2006	New generic version	7.451	12.332	0.922
simvastatin	June 2006	New generic version	3.128	7.497	1.000
Simcor	February 2008	New dosage form	12.353	9.188	0.228

Notes: This table summarizes the variation across individual U.S. physicians in the initial use of twelve new pharmaceutical products, each aimed at controlling blood cholesterol or lipid levels. Each product was initially released for sale in the U.S. market on the date indicated. New drug applications are categorized by the FDA based on whether the proposed product is a new molecular entity, a new drug combination, a new dosage form, or a new generic equivalent. The distribution of initial prescription dates is described by the mean and standard deviation (SD) in months from initial U.S. market approval. The share of physicians that prescribe the product at least once by December 2010 (Final Share) ranges from 15 percent (Altoprev) to 100 percent (simvastatin). Prescription data are from IMS Health.

Table 2: Regression Summary Statistics

Variable	Mean	SD	Min	Max
Physician-Drug-Month Level:				
Number of Prescriptions	4.411	12.560	0	2383
Indicator for Positive Prescriptions	0.358	0.479	0	1
Log Prescriptions, Conditional on Positive Prescriptions	1.725	1.256	0	7.776
Physician-Month Level:				
Drug Database Indicator	0.242	0.429	0	1
Drug Database and Use Indicator	0.131	0.338	0	1
Number of Unique Drugs Prescribed	5.210	2.798	1	16
Prescription Herfindahl Index	0.447	0.232	0.097	1
Physician-Drug Level:				
Months to First Prescription	19.140	21.950	0	122
Drug-Month Level:				
Indicator for New Drug, 24 months	0.159	0.365	0	1
General:				
Number of Unique Prescribers	131323			
Number of Drugs, January 2000	6			
Number of Drugs, January 2000 - December 2010	18			

Notes: This table summarizes the physician-level prescription and information access data used in the analysis. Statistics correspond to U.S. physicians that prescribe a minimum of ten statin or lipid-lowering products both during January-December 2000 and January-December 2010. The Drug Database indicator is a binary variable reflecting physician-month level information access, Drug Database and Use indicates physicians that use the information database for cholesterol drugs, and prescription diversity is summarized by Number of Unique Drug Prescribed and the physician-month level Herfindahl index. Drugs are considered New if within 24 months of market approval by U.S. Food and Drug Administration. Prescription variables are from IMS Health and database registration data are from a leading U.S. point-of-care medical applications firm. .

Table 3: Descriptive Statistics

	Physician Level, Within-Zipcode Average					Physician Level	
Product:	lovastatin	pravastatin	simvastatin	Lipitor	Generic	Lipitor	Generic
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Share in total prescriptions							
Final month, December 2010							
Mean	0.0648	0.0986	0.4304	0.1805	0.5938	0.1920	0.5695
Standard deviation	0.0606	0.0764	0.1323	0.0795	0.1248	0.1103	0.1692
Min	0.0279	0.0462	0.3024	0.1111	0.4651	0	0
Max	0.1457	0.1931	0.5586	0.2696	0.7107	1	1
Max-min gap	0.1178	0.1469	0.2562	0.1584	0.2456	1	1
Observations	11950	11950	11950	11950	11950	81412	81412
B. Share in molecule-specific prescriptions							
Six months after generic release							
Mean	0.8357	0.8725	0.8267				
Standard deviation	0.2306	0.1144	0.2159				
Min	0.6370	0.7330	0.5746				
Max	0.9286	0.9475	0.9402				
Max-min gap	0.2916	0.2145	0.3656				
C. Share in molecule-specific prescriptions							
Final month, December 2010							
Mean	0.9997	0.9949	0.9980				
Standard deviation	0.0011	0.0138	0.0051				
Min	0.9973	0.9669	0.9874				
Max	1.0000	0.9990	0.9995				
Max-min gap	0.0027	0.0322	0.0122				

Notes: This table describes prescription heterogeneity across U.S. physicians. Statistics correspond to U.S. physicians that prescribe a minimum of 30 statin or lipid-control products in the observation month indicated. Panel A describes prescribing in December 2010 across all physicians (columns 6, 7), and the average zipcode-specific value across U.S. zipcodes (columns 1-5). Panels B and C describe physicians' within-molecule substitution toward generics for lovastatin (column 1), pravastatin (column 2), and simvastatin (column 3); all report average zipcode-specific values across U.S. zipcodes. Panel B describes this substitution six months after generic release, while Panel C describes the final sample period in December 2010. The upper-left number in Panel A (mean, lovastatin, 0.0648) is the average across U.S. zipcodes in the mean share (across physicians in that zipcode) of all prescriptions accounted for by lovastatin in December 2010; the upper-left number in Panel B (mean, lovastatin, 0.8357) is the average across U.S. zipcodes in the mean share of lovastatin and Mevacor prescriptions accounted for by lovastatin, six months after lovastatin's release; the upper-left number in Panel C (mean, lovastatin, 0.9997) is the average across U.S. zipcodes in the mean share of lovastatin and Mevacor prescriptions accounted for by lovastatin in December 2010. Generic approval dates are from the U.S. Food and Drug Administration; all other variables are from IMS Health.

Table 4: Time to First Prescription of New Drug, U.S. Physicians, 2000—2010

Dependent Variable:	Log(time to first prescription of drug *j* by physician *i*)					
	All physicians				Eventual users	
	(1)	(2)	(3)	(4)	(5)	(6)
Physician-Drug Information	-0.0691***	-0.0257***	-0.0414***	-0.0010	0.0138**	0.0147*
	0.0036	0.0069	0.0040	0.0073	0.0056	0.0083
Physician-Drug Information x Generic			-0.0751***	-0.0578***	-0.0470***	-0.0469***
			0.0079	0.0069	0.0120	0.0106
Drug FE	Y	Y	Y	Y	Y	Y
Doctor FE	N	Y	N	Y	N	Y
N	935745	935745	935745	935745	204386	204386
R-Squared	0.3928	0.6202	0.3929	0.6202	0.4210	0.6340

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table provides least-squares estimates of equation (6) for U.S. physicians' prescription of cholesterol drugs first introduced to the U.S. market during January 2000 through December 2010. The dependent variable captures the time lapse between FDA approval of drug j and physician i's initial prescription of it for the full sample of physicians (columns 1-4) and for the subset of physicians that eventually use the electronic reference for cholesterol drugs (columns 5-6). Information is an indicator variable that is equal to 1 for physicians with information access through the electronic drug reference at the time of drug j's FDA approval and is otherwise zero. Generic is an indicator for the products pravastatin, lovastatin, and simvastatin. Regressions include drug fixed effects (columns 1-6) and physician fixed effects (columns 2, 4, 6). Qualitatively identical results obtain with Poisson and Negative Binomial estimation. Robust standard errors appear below each point estimate.

Table 5: Prescription Diversity, U.S. Physicians, 2000—2010

Dependent Variable:	All physicians			Eventual users		
	Number of Unique Drugs		Drug HHI	Number of Unique Drugs		Drug HHI
	(1)	(2)	(3)	(4)	(5)	(6)
Physician-Month Information	0.2517***	0.0237***	-0.0115***	0.0721***	0.0074***	-0.0019***
	0.0022	0.0007	0.0002	0.0027	0.0019	0.0003
Doctor FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N	16378056	16378056	16378056	3515201	3515201	3515201
R-Squared	0.7496	-	0.5756	0.7524	-	0.5710

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table provides estimates of equation (7) for cholesterol drug prescriptions by U.S. physicians during January 2000 through December 2010, including all sample physicians (columns 1-3) and the subset of physicians that eventually use the electronic reference for cholesterol drugs (columns 4-6). The dependent variable in columns 1, 2, 4, and 5 captures the prescription diversity of physician i as captured by the number of unique drugs j she prescribes during month t. The dependent variable in columns 3 and 6 is the prescription Herfindahl-Hirschman index for physician i in month t. Each column includes all 132 months in the sample period. Columns 1, 3, 4, and 6 provide least-squares estimates; columns 2 and 5 provide Poisson estimates. Information is an indicator variable that is equal to 1 for physicians with information access through the electronic drug reference database in month t and is otherwise zero. All regressions include physician and month fixed effects. Results are robust to including zipcode-month fixed effects. Robust standard errors appear below each point estimate.

Table 6: Prescription Propensity, U.S. Physicians, 2000-2010

Dependent Variable:	1{Prescriptions of drug *j* > 0}	
	All physicians	Eventual users
	(1)	(2)
Information x New x Generic	0.0161***	0.0151***
	0.0005	0.0005
Information x New x Branded	-0.0025***	-0.0017***
	0.0002	0.0002
Information x Old x Generic	0.0228***	0.0176***
	0.0004	0.0004
Information x Old x Branded	0.0041***	0.0020***
	0.0001	0.0001
Drug x Month FE	Y	Y
Physician x Product FE	Y	Y
N	238351245	106290030
R-Squared	0.6645	0.6660

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table provides least-squares estimates of equation (8) for prescriptions of cholesterol drugs by U.S. physicians during January 2000 through December 2010, including all sample physicians (column 1) and the subset of physicians that eventually use the electronic reference for cholesterol drugs (column 2). The dependent variable is an indicator for whether the doctor *i* prescribes drug *j* during month *t.* Information is an indicator variable that is equal to 1 for physicians with information access through the electronic drug reference database and is otherwise zero. New is an indicator that is equal to 1 for drug products that are within 24 months of initial approval by the U.S. FDA, and is otherwise zero. Generic indicates the products pravastatin, lovastatin, and simvastatin. All regressions include drug-month fixed effects and physician-drug fixed effects. Results are robust to logit estimation and zipcode-month fixed effects. Standard errors appear below each point estimate.

Table 7: Placebo Test, Time to First Prescription of New Drug, U.S. Physicians, 2000—2010

Dependent Variable:	Log(time to first prescription of drug *j* by physician *i*)			
	All physicians			
	(1)	(2)	(3)	(4)
Physician-Drug Information	-0.0688***	-0.0266***	-0.0406***	-0.0014
	0.0036	0.0069	0.0041	0.0073
Physician-Drug Information x Generic			-0.0768***	-0.0594***
			0.0080	0.0070
Physician-Drug Placebo Information	0.0029	-0.0128*	0.0079	-0.0091
	0.0042	0.0071	0.0048	0.0076
Physician-Drug Placebo Information x Generic			-0.0145	-0.0113
			0.0091	0.0079
Drug FE	Y	Y	Y	Y
Doctor FE	N	Y	N	Y
N	935745	935745	935745	935745
R-Squared	0.3928	0.6202	0.3929	0.6202

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table provides least-squares estimates of equation (6) for U.S. physicians' prescription of cholesterol drugs first introduced to the U.S. market during January 2000 through December 2010. The dependent variable captures the time lapse between FDA approval of drug j and physician i's initial prescription of it for the full sample of physicians. Information is an indicator variable that is equal to 1 for physicians with information access through the electronic drug reference at the time of drug j's FDA approval and is otherwise zero; Placebo Information indicates physicians that have database access but during the sample period never use it to look up a cholesterol drug. Generic is an indicator for the products pravastatin, lovastatin, and simvastatin. Regressions include drug fixed effects (columns 1-4) and physician fixed effects (columns 2, 4). Qualitatively identical results obtain with Poisson and Negative Binomial estimation. Robust standard errors appear below each point estimate.

Table 8: Placebo Test, Prescription Diversity, U.S. Physicians, 2000—2010

Dependent Variable:	All physicians			Eventual adopters		
	Number of Drugs		Drug HHI	Number of Drugs		Drug HHI
	(1)	(2)	(3)	(1)	(2)	(3)
Physician-Month Information	0.2614***	0.0248***	-0.0121***	0.1335***	0.0106***	-0.0048***
	0.0021	0.0007	0.0002	0.0024	0.0018	0.0003
Physician-Month Placebo Information	0.0825***	0.0094***	-0.0056***	-0.0420***	-0.0038**	0.0015***
	0.0021	0.0006	0.0002	0.0023	0.0018	0.0002
Doctor FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N	16378056	16378056	16378056	7338162	7338162	7338162
R-Squared	0.7497	-	0.5756	0.7497	-	0.5723

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table provides estimates of equation (7) for cholesterol drug prescriptions by U.S. physicians during January 2000 through December 2010, including all sample physicians (columns 1-3) and the subset of physicians that eventually use the electronic reference for cholesterol drugs (columns 4-6). The dependent variable in columns 1, 2, 4, and 5 captures the prescription diversity of physician i as captured by the number of unique drugs j she prescribes during month t. The dependent variable in columns 3 and 6 is the prescription Herfindahl-Hirschman index for physician i in month t. Columns 1, 3, 4, and 6 provide least-squares estimates; columns 2 and 5 provide Poisson estimates. Information is an indicator variable that is equal to 1 for physicians with information access through the electronic drug reference database in month t and is otherwise zero; Placebo Information indicates physicians that have database access but durign the sample period never use it to look up a cholesterol drug. All regressions include physician and month fixed effects. Results are robust to including zipcode-month fixed effects. Robust standard errors appear below each point estimate.

Table 9: Placebo Test, Prescription Propensity, U.S. Physicians, 2000-2010

Dependent Variable:	1{Prescriptions of drug *j* > 0}	
	All physicians	Eventual adopters
	(1)	(2)
Information x New x Generic	0.0164***	0.0166***
	0.0005	0.0005
Information x New x Branded	-0.0021***	0.0005**
	0.0002	0.0003
Information x Old x Generic	0.0238***	0.0176***
	0.0004	0.0005
Information x Old x Branded	0.0042***	0.0013***
	0.0001	0.0001
Placebo Information x New x Generic	0.0134**	0.0136**
	0.0005	0.0005
Placebo Information x New x Branded	0.0060***	0.0089***
	0.0002	0.0003
Placebo Information x Old x Generic	0.0129***	0.0073***
	0.0003	0.0004
Placebo Information x Old x Branded	0.0024***	-0.0001
	0.0001	0.0001
Drug x Month FE	Y	Y
Physician x Product FE	Y	Y
N	238351245	106290030
R-Squared	0.6645	0.6660

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table provides least-squares estimates of equation (8) for prescriptions of cholesterol drugs by U.S. physicians during January 2000 through December 2010, including all sample physicians (column 1) and the subset of physicians that eventually use the electronic reference for cholesterol drugs (column 2). The dependent variable is an indicator for whether the doctor i prescribes drug j during month t. Information is an indicator variable that is equal to 1 for physicians with information access through the electronic drug reference database in month t and is otherwise zero; Placebo Information indicates physicians that have database access but durign the sample period never use it to look up a cholesterol drug. New is an indicator that is equal to 1 for drug products that are within 24 months of initial approval by the U.S. FDA, and is otherwise zero. Generic indicates the products pravastatin, lovastatin, and simvastatin. All regressions include drug-month fixed effects and physician-drug fixed effects. Results are robust to logit estimation and zipcode-month fixed effects. Standard errors appear below each point estimate.

Table 10: Mandatory Substitution Laws by Zipcode, Time to First Prescription, U.S. Physicians, 2000—2010

Dependent Variable:	With Mandatory Substitution Law				Without Mandatory Substitution Law			
	Log Time to First Rx							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Physician-Drug Information	-0.0580***	-0.0297**	-0.0277***	0.0005	-0.0738***	-0.0238***	-0.0468***	-0.0012
	0.0065	0.0127	0.0082	0.0137	0.0044	0.0085	0.0055	0.0092
Physician-Drug Information x Generic			-0.0809***	-0.0688***			-0.0736***	-0.0534***
			0.0134	0.0118			0.0091	0.0081
Drug FE	Y	Y	Y	Y	Y	Y	Y	Y
Doctor FE	N	Y	N	Y	N	Y	N	Y
N	304153	303778	304153	303778	631529	630916	631529	630916
R-Squared	0.3949	0.6262	0.3950	0.6262	0.3928	0.6179	0.3926	0.6180

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table provides least-squares estimates of equation (6) for U.S. physicians' prescription of cholesterol drugs first introduced to the U.S. market during January 2000 through December 2010. The estimates are presented for two subsamples: physicians located in states with active mandatory substitution laws (columns 1-4) and those without such laws (columns 5-8). The dependent variable captures the time lapse between FDA approval of drug j and physician i's initial prescription of it. Information is an indicator variable that is equal to 1 for physicians with information access through the electronic drug reference at the time of drug j's FDA approval and is otherwise zero. Generic is an indicator for the products pravastatin, lovastatin, and simvastatin. Regressions include drug fixed effects (columns 1-8) and physician fixed effects (columns 2, 4, 6, 8). Qualitatively identical results obtain with Poisson and Negative Binomial estimation. Robust standard errors appear below each point estimate.

Table 11: Medical Innovation by Zipcode, Time to First Prescription, U.S. Physicians, 2000—2010

Dependent Variable:	Bottom 5 Percent by Pharmaceutical Innovation				Top 5 Percent by Pharmaceutical Innovation			
	Log Time to First Rx							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Physician-Drug Information	-0.0647***	-0.0552***	-0.0336**	-0.0191	-0.0589***	-0.0628***	-0.0394**	-0.0427*
	0.0127	0.0228	0.0161	0.0249	0.0130	0.0236	0.0165	0.0258
Physician-Drug Information x Generic			-0.0823***	-0.0831***			-0.0517	-0.0468*
			0.0262	0.0233			0.0269*	0.0239
Drug FE	Y	Y	Y	Y	Y	Y	Y	Y
Doctor FE	N	Y	N	Y	N	Y	N	Y
N	49479	49479	49479	49479	45100	45100	45100	45100
R-Squared	0.3784	0.5519	0.3784	0.5520	0.3883	0.5596	0.3884	0.5596

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table provides least-squares estimates of equation (6) for U.S. physicians' prescription of cholesterol drugs first introduced to the U.S. market during January 2000 through December 2010. The estimates are presented for two subsamples: physicians located in U.S. 4-digit zipcodes in the bottom five percent based on the number of pharmaceutical patents filed with the USPTO since 1976 (columns 1-4) and those located in the top five percent (columns 5-8). The dependent variable captures the time lapse between FDA approval of drug j and physician i's initial prescription of it. Information is an indicator variable that is equal to 1 for physicians with information access through the electronic drug reference at the time of drug j's FDA approval and is otherwise zero. Generic is an indicator for the products pravastatin, lovastatin, and simvastatin. Regressions include drug fixed effects (columns 1-8) and physician fixed effects (columns 2, 4, 6, 8). Qualitatively identical results obtain with Poisson and Negative Binomial estimation. Robust standard errors appear below each point estimate.

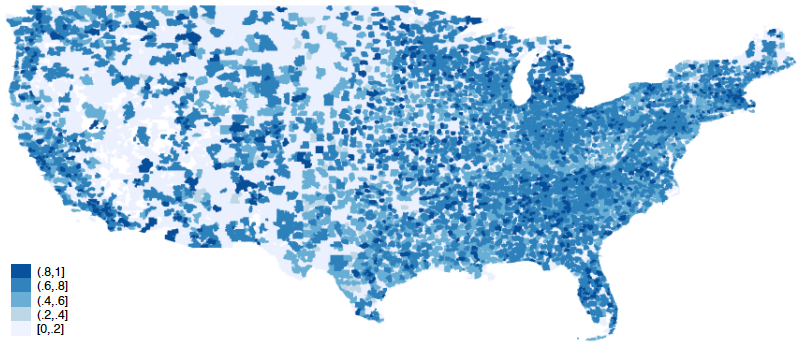
Table 12: Information and Prescribing Heterogeneity Among U.S. Physicians

	Rx Shares - Euclidean Distance to the Mean		
	Z = 0	Z = 1	
	(1)	(2)	(3)
A. December 2010			
Mean	0.1762	0.1522	
Estimated difference in means			-0.0236***
Standard error			0.0014
B. January 2000			
Mean	0.2162	0.2037	
Estimated difference in means			-0.0093***
Standard error			0.0015
C. January 2010 vs. 2000			
Difference in means	-0.0400	-0.0515	
Estimated difference in differences			-0.0107***
Standard error			0.0017
Estimated average change			-0.0447***
Standard error			0.0011

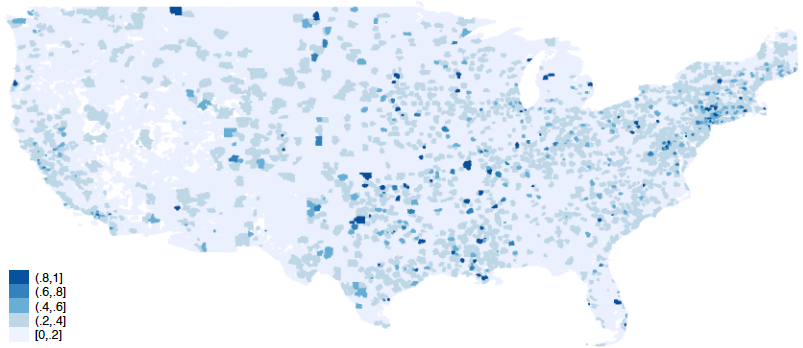
Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. This table summarizes prescription heterogeneity across U.S. physicians and over time. Columns 1 and 2 indicate the average Euclidean distance (norm) between the vector of physician-i prescription shares across drugs j and the vector of average prescription shares in December 2010 (Panel A) and in January 2000 (Panel B) for physicians without access to the electronic database in December 2010 (column 1) and for physicians with access in December 2010 (column 2). Column 3 presents estimates from two cross-section regressions in which the mean Euclidean distance between physician i and his group average is the dependent variable, regressed on an indicator for database access in December 2010 and zipcode fixed effects. Panel C provides difference-in-differences estimates with two time periods (January 2000 and December 2010); the dependent variable is as in Panels A and B, and is regressed on an indicator for December 2010, its interaction with the indicator for database access, and physician fixed effects. Standard errors appear below each point estimate.

Figure 1: Heterogeneity in Prescribing, by U.S. Zipcode, December 2010

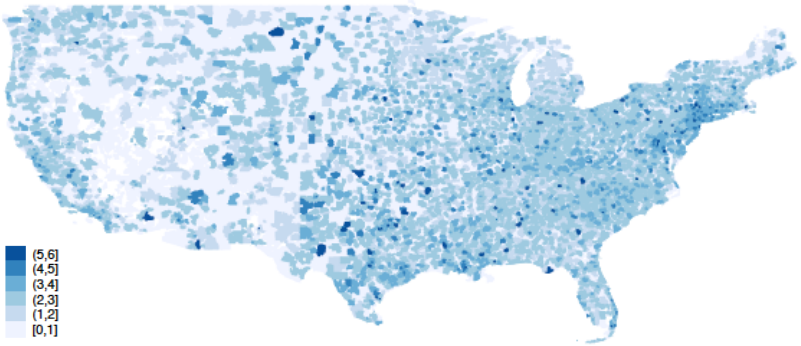
Panel A — Generic Share in Cholesterol Drug Prescriptions



Panel B — Lipitor Share in Cholesterol Drug Prescriptions



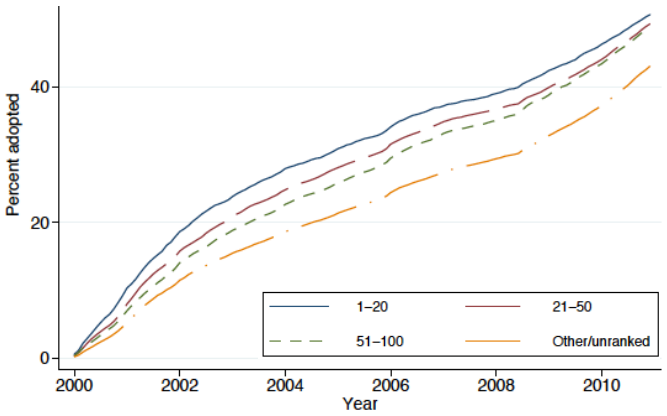
Panel C — Cost per Tablet Dispensed



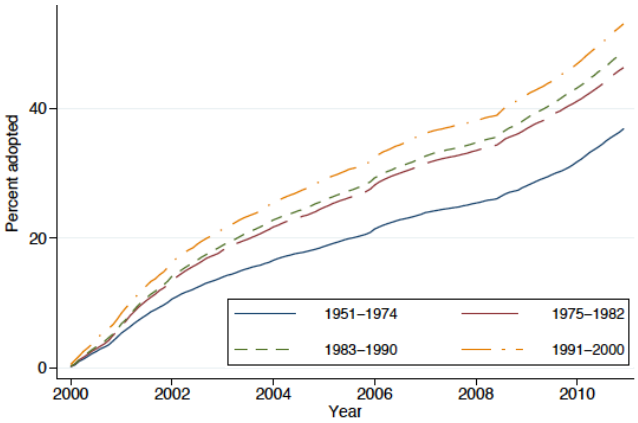
Notes: This figure illustrates the prescription of pharmaceutical drug products across U.S. zipcodes, specifically for products aimed at adjusting blood cholesterol and lipid levels. Panel A illustrates the share of dispensed prescriptions corresponding to generic products (lovastatin, pravastatin, and simvastatin) by zipcode; Panel B illustrates the share corresponding to Lipitor, the most expensive drug at the time; Panel C illustrates an approximation of the average cost per pill disensed using price data from the Centers for Medicare and Medicaid Services. In each panel, darker shades indicate higher values of the variable indicated. All figures correspond to prescriptions filled in December 2010 based on data from IMS Health.

Figure 2: Heterogeneity in Adoption of the Drug Reference Database, January 2000 - December 2010

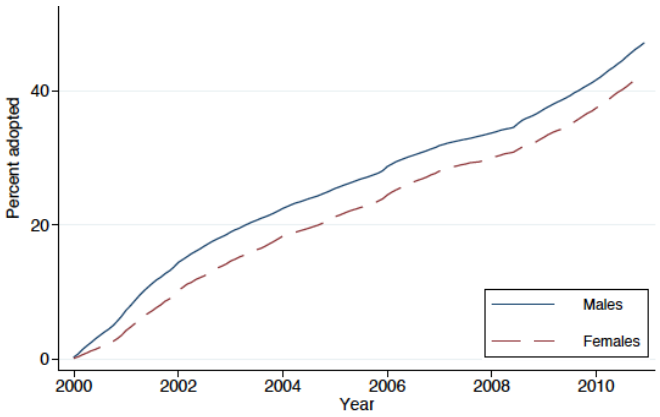
Panel A — By school rank



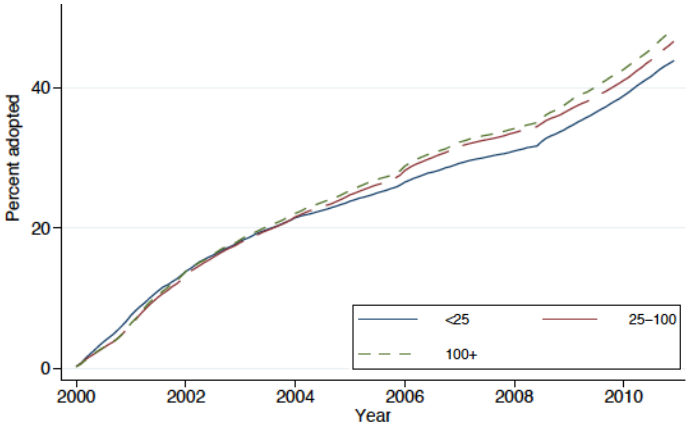
Panel B — By graduation year



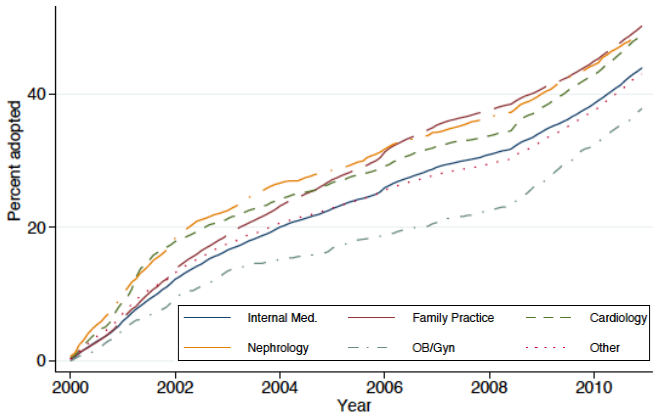
Panel C — By sex



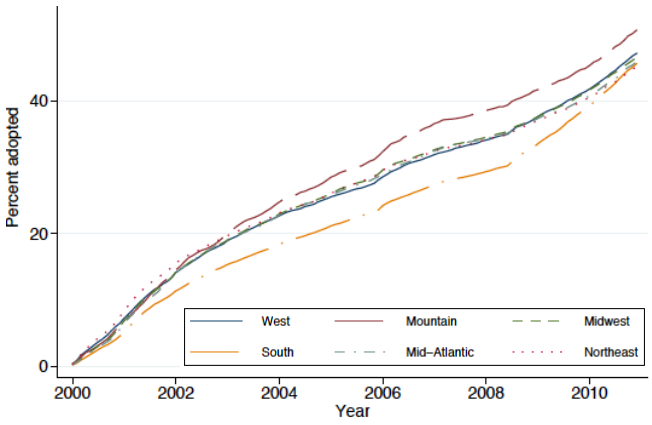
Panel D — By monthly prescription volume



Panel E — By medical specialty



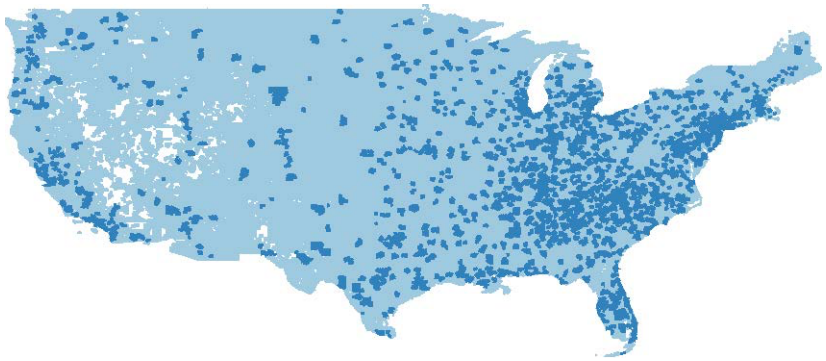
Panel F — By U.S. region



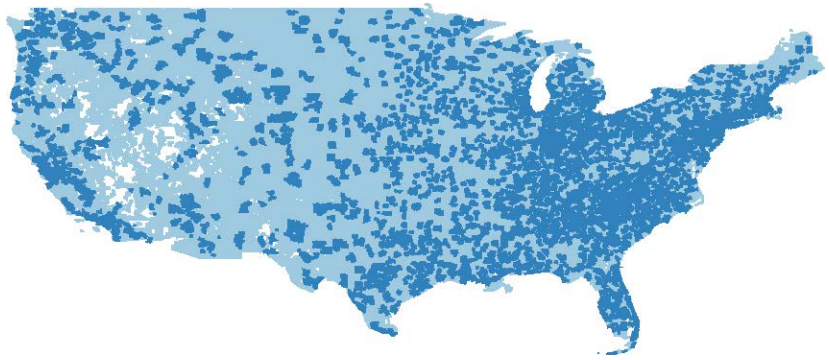
Notes: This figure plots the fraction of the 131,232 sample U.S. physicians that are registered users of the electronic drug reference database by the date indicated, and shows the extent to which adoption rates differ across physicians according their observable characteristics. Database registration data are from a leading U.S. point-of-care medical applications firm. Medical school rank is determined based on data from the U.S. News and World Report service, and all other variables are from the CMS Physician Compare database.

Figure 3: Heterogeneity in the Initial Use of a New Medical Technology, by U.S. Zipcode

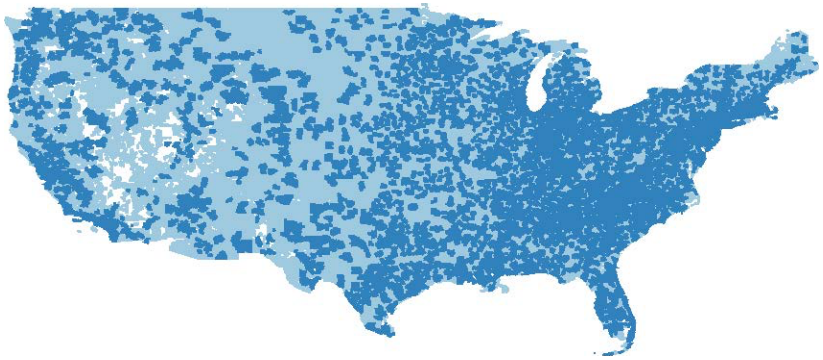
Panel A — One Month After Release



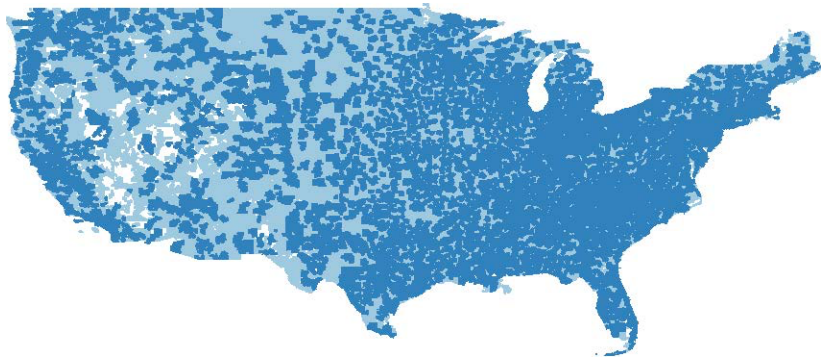
Panel B — Three Months After Release



Panel C — Six Months After Release



Panel D — Thirty-Six Months After Release



Notes: This figure illustrates the gradual diffusion of a new pharmaceutical drug, the statin Crestor, across zipcodes within the continental United States. Dark shades indicate zipcodes in which at least one prescription of Crestor has been written and filled, light shades indicate zipcodes in which Crestor has not yet been prescribed; areas shaded white contain no data. The four panels correspond to four points in time following the initial market introduction of Crestor in August 2003. These four points are September 2003 (Panel A), November 2003 (Panel B), February 2004 (Panel C), and August 2006 (Panel D). Prescription data are from IMS Health.

Appendix

A.1 Medical Innovation

Innovation in *hypercholesterolemia* and *dyslipidemia* therapy: Information about the evolving set of pharmaceutical therapies available for prescription was obtained from the U.S. Food and Drug Administration (FDA) for the period January 2000 through December 2010. Twelve new statin or lipid-lowering drugs, including new formulations, combinations, and versions, introduced during this period and are described below. These include three new molecular entities Crestor, Lovaza, and Zetia; three generic versions lovastatin (Mevacor), pravastatin (Pravachol), and simvastatin (Zocor); two new formulations Altoprev (extended-release Mevacor) and Lescol XL (extended-release Lescol); and four new drug combinations Advicor (extended-release niacin and Mevacor), Pravigard PAC (aspirin and Pravachol), Vytorin (Zetia and Zocor), Simcor (extended-release niacin and Zocor). A description of each drug innovation appears below based on publicly available data including approval letters and administrative, medical, and pharmacological review. Baycol was withdrawn early in the sample period in August 2001 and is thus omitted.

Existing therapies available in January 2000:
1. Lescol (fluvastatin) is a statin marketed by Novartis since its FDA approval as a new molecular entity on December 31, 1993; its patent protection expired in 2012. Like other statins, its mechanism of action is to limit a specific enzyme in the liver, preventing cholesterol synthesis.
2. Lipitor (atorvastatin) is a statin marketed by Pfizer. Its mechanism of action is similar to that of fluvastatin, but unlike other statins, atorvastatin is a synthetic compound. The therapy was approved by the FDA as a new molecular entity on December 17, 1996. Between 1996 and 2012, Lipitor was the best-selling drug globally; its patent expired in November 2011.
3. Mevacor (lovastatin) is the first statin to receive FDA approval. The drug was approved as a new molecular entity on August 31, 1987 for sale in the United States by Merck. The therapy was protected by patents through June 2001.
4. Niaspan (extended-release niacin) is vitamin B_3, or nicotinic acid, and is marketed by Abbott Laboratories. Extended-release niacin was approved for sale in the United States on July 28, 1997.
5. Pravachol (pravastatin) is a statin marketed by Bristol Myers Squibb since its FDA approval on October 31, 1991. In addition to inhibiting cholesterol synthesis, Pravachol also inhibits low-density lipoprotein synthesis. Two clinical trials, each completed in November 2003, suggest Pravachol is outperformed by both Zocor and Lipitor. Patent protection expired in June 2006.
6. Zocor (simvastatin) is a statin marketed by Merck since its FDA approval as a new molecular entity on December 23, 1991. Zocor outperformed Pravachol in its prevention of cholesterol synthesis in a clinical trial completed in November 2003. Patent protection expired in April 2006.

New chemical entities, January 2000–December 2010:
1. Crestor (rosuvastatin calcium) is a new molecular entity approved by the FDA for sale in the United States by Astra Zeneca Pharmaceuticals on August 12, 2003. The molecule acts by reducing intestinal absorption of cholesterol and related phytosterols, and is thereby distinct relative to other statin therapies. The drug was approved for use in treating primary hypercholesterolemia and mixed dyslipidemia (by reducing total-C, LDL-C, and Apo B), and as an adjunct to other lipid-lowering treatments. It was thus approved for use alone or with other statins. A 2008 clinical trial revealed additional evidence supporting the superior performance of Crestor compared with a placebo treatment. Patent protection expires in January 2016.
2. Lovaza (omega-3-acid ethyl esters) is a new molecular entity introduced by Abbott labs and approved by the FDA on November 10, 2004. It was initially introduced under the trade name Omacor. Unlike statins, Lovaza is aimed at reducing tricylerides rather than low-density lipopro-

teins and may thus be combined with a statin as an adjunct therapy. Patent protection expired in September 2012.
3. Zetia (ezetimibe) is a new molecular entity introduced by Schering and approved by the FDA on October 25, 2002 for sale in the United States. The molecule acts by reducing intestinal absorption of cholesterol and related phytosterols, and is thus distinct from statins. The drug was initially approved for use in treating hypercholesterolemia for use alone or with other statins. In January 2008, a clinical trial found Zetia performed poorly compared with other therapies, and it was at that time recommended that Zetia not be prescribed except in cases for which all other cholesterol drugs had previously failed. Patent protection expires in April 2017.

New generic versions, January 2000–December 2010:
1. Lovastatin is the generic equivalent of Mevacor, and was initially approved by the FDA for sale in the United States by Geneva Pharmaceuticals applied on December 17, 2001.
2. Pravastatin is the generic equivalent of Pravachol, and was initially approved by the FDA for sale in the United States by Teva Pharmaceuticals on April 24, 2006.
3. Simvastatin is the generic equivalent of Zocor, and was initially approved by the FDA for sale in the United States by Teva Pharmaceuticals on June 23, 2006.

New dosage forms, January 2000–December 2010:
1. Altoprev (extended-release lovastatin) is a new dosage form and was approved by the FDA on June 26, 2002 for sale in the United States, following a new drug application by Aura Pharmaceuticals, Inc. of March 30, 2001. The approval is for use of Altoprev for lowering cholesterol and LDL-C to target levels along with diet and exercise, to slow the progression of atherosclerosis in patients with coronary heart disease, and to reduce total-C, LDL-C, Apo B, and triclycerides and to increase HDL-C in patients with dyslipoproteinemia. The drug was found to outperform Mevacor (lovastatin). Altoprev is protected by patents though at least December 2017.
2. Lescol XL (extended-release Lescol) is a new dosage form and was approved by the FDA for sale in the United States by Novartis on October 6, 2000. Patent protection expired in 2012.

New drug combinations, January 2000–December 2010:
1. Advicor (Mevacor and extended-release Niacin) is a new drug combination approved by the FDA on December 17, 2001 for sale in the United States by Kos Pharmaceuticals. Advicor was approved for use in treating primary hypercholesterolemia and mixed dyslipidemia in two types of patients: a) those treated with lovastatin who require further triglyceride lowering or HDL raising who may benefit from adding niacin to their regimen, and b) patients previously treated with niacin who require further LDL lowering and may benefit from having lovastatin added to their regimen. Thus, Advicor was not approved as an initial therapy for lowering LDL levels. Moreover, in clinical trials, Advicor was found to perform no better than Mevacor as a first-line agent.
2. Pravigard PAC (Pravachol and aspirin) is a new drug combination approved by the FDA on June 24, 2003 for sale in the United States by Bristol Myers Squibb.
3. Vytorin (Zetia and Zocor) is a new drug combination approved by the FDA for use, along with diet or with other lipid-lowering treatments to reduce total C, LDL-C and raise HDL-C, on July 23, 2004 by MSP Singapore company, LLC. The drug combination was more effective at lowering lipids, but was also associated with more adverse events (both serious and leading to discontinuation) than either monotherapy. In January 2008, a completed clinical trial revealed Zetia, a component of Vytorin, performed poorly relative to other therapies.
4. Simcor (simvastatin and extended-release niacin) is a new drug combination approved by the FDA on February 15, 2008 for sale in the United States by Abbott Laboratories. Like Advicor, Simcor is approved only as a second-line treatment for cases in which the monotherapy is considered to be inadequate.

A.2 Data

U.S. Prescriptions for *Hypercholesterolemia* and *Dyslipidemia* Therapies: Prescription data for U.S. medical practitioners and each of the products described above were obtained from the IMS Health Xponent database. IMS Health draws its prescription data from a large but non-random sample of approximately 70 percent of U.S. pharmacies. As of 2011, Xponent includes direct information from over 38,000 retail stores, including approximately 119 mail-service pharmacies and 820 long-term care facilities; this compares with a universe of approximately 57,000 retail pharmacies, 327 mail-service outlets, and 3,000 long-term care facilities. To correct for sampling error and to ensure the data are representative, IMS Health has applied a proprietary re-weighting procedure to arrive at the prescription data provided to us for this study. Weights are constructed by combining pharmacy data with additional pharmaceutical sales data derived from multiple payor-based sources; these latter data provide detailed information regarding the prescription sales of nearly all non-sample pharmacies so that it may accurately represent the universe of prescription sales. If the response rate for a particular pharmacy is low relative to other pharmacies in a region, for example, IMS Health is able to re-weight observed prescription data for that pharmacy so that the total quantity purchased by a physician in that region is nevertheless representative. Importantly, this weighting procedure applies only to strictly positive prescription levels, but does not apply to zeros, enabling us to accurately track the initial adoption of new products over time for each physician.

The data IMS Health provided include prescriptions by 280,622 unique U.S. physicians for each product in each month during January 2000 through December 2010. To avoid studying physicians specialized outside cardiovascular care, we restrict analysis to physicians that prescribe at least some cholesterol products. Specifically, for a physician to be included in the dataset, he or she needs to have written at least ten filled prescriptions for cholesterol therapies during the calendar year 2010. The data provide precise identifying information for each prescribing physician, including the unique, 11-digit American Medical Association Medical Education Number, the first name, last name, and middle name, and the five-digit zipcode corresponding to the medical practice of the physician. From January 2006 through December 2010, the data provide additional detail regarding prescriptions: for each drug, a separate prescription count is observed for each of four payment methods, including Medicare Part D, Fee-for-Service Medicaid, cash, and commercial insurance. In the data, approximately half of dispensed prescriptions for cholesterol drugs correspond to individuals with commercial insurance; 34 percent obtain products through Medicare Part D, ten percent purchase medications with cash, and the remaining six percent are covered by Medicaid.

To prepare the data for analysis, we reshaped the files provided so that each row corresponds to a doctor-drug-month triplet. With guidance from IMS Health, zeros were explicitly introduced in this step for missing observations corresponding to existing products not associated with positive prescriptions in the IMS data. Starting in 2006, we aggregated prescriptions across methods of payment to arrive at a single number of prescriptions written by physician, drug, and month. We combined prescriptions for "Pravastatin" and "Pravastatin SOD", which are the same product, and did likewise for "Lovaza" and "Omacor", which are the same product. We dropped Baycol from the dataset. For some years, due to the projection calculation described above, the prescription variable was not a whole number; with guidance from IMS Health, we rounded the number of prescriptions to the nearest whole number. To abstract from physician entry during the sample period, we impose a sample restriction in addition to that described above: specifically, each physician included must prescribe at least ten cholesterol drugs during the calendar year 2000. Finally, we used information from the U.S. FDA to determine the approval date for each therapy. The first month after this date was determined to be the first month of a drug's market life in the United States. We created indicator variables for drugs that are new corresponding to the first six months of the drug's market

life in the United States, and separately, to the first 24 months of the drug's market life in the United States. We created indicator variables for generic products lovastatin, pravastatin, and simvastatin.

Electronic Database Use for *Hypercholesterolemia* and *Dyslipidemia* Therapies, by U.S. Physicians: We obtained data on individual physicians' information access from the leading U.S. point-of-care medical applications firm. For each physician, we observe the corresponding initial database registration date; this is used to construct the indicator variable Z_{it} that takes on a value of one for registered users, and that is otherwise zero. For each physician-product-month triplet, we also observe a proxy for the number of lookups completed. During January 2000 through December 2010, the share of sample physicians registered as database users rose from 0.003 to 0.446. Our analysis is thus based on a sample combining a) physicians that first registered during or before the sample period (44.3 percent), and b) physicians that registered before the sample period (0.3 percent) and c) physicians that never registered (55.4 percent). Each physician is identified in the data by a unique, 11-digit American Medical Association Medical Education Number, first name, last name, middle name, and five-digit zipcode. These characteristics form the basis for a merge with the prescription information described above.