Knowledge Diffusion Through Networks*

Treb Allen
Dartmouth & NBER

L. Kamran Bilir
Wisconsin & NBER

Zhang Chen
Princeton

Christopher Tonetti
Stanford GSB & NBER

December 2019

PRELIMINARY

Abstract

How do geography and other barriers to the free flow of information shape the rate of knowledge diffusion? To address this question, we develop an empirical model of product discrete choice with Bayesian learning on a social network. Estimating this model using monthly data on the cholesterol-drug prescription decisions of over 50,000 U.S. physicians during January 2000 through December 2010, we find that the evolution of product choice efficiency is highly responsive to network structure changes, particularly targeted friction reductions that strengthen the strongest bilateral links.

*We thank Murray Aitken, Allen Campbell, seminar participants at AEA, CEMFI, Penn State, Princeton, MIT, Notre Dame, Science Po, SED, UBC, and UCSD-UCLA for very helpful comments. E-mail: treb@dartmouth.edu, bilir@wisc.edu, zhangc@princeton.edu, tonetti@stanford.edu.

1 Introduction

How do geography and other barriers to the free flow of information shape the rate of knowledge diffusion? In a range of relevant contexts, barriers impeding knowledge flows across individuals negatively impact aggregate efficiency, motivating policies that aim to reduce such frictions.[1] However, relatively little is known about how to efficiently implement these policies, particularly when the policymaker is not fully-informed about the underlying knowledge that is diffusing.[2] To achieve the largest productivity gains, is it more effective to target the distribution of information across agents in the network, or the structure of the network itself? In the latter case, is it more efficient to target the weakest links, or the strongest?

To shed light on these questions, we develop an empirical model of product discrete choice with learning dynamics on a social network. Under the assumptions of Bayesian learning and static multinomial choice, we show that the model parameters may be estimated using a simple non-linear estimator. We apply this to quantify the model using monthly data on the cholesterol-drug prescription decisions of over 50,000 U.S. physicians during January 2000 through December 2010, a period featuring patent expirations, drug entry, and major aggregate shifts in product choice. To understand how barriers to knowledge diffusion and learning impact the evolution of choice efficiency, we simulate the quantified model under a series of targeted policy interventions. These indicate that the evolution of efficiency is highly responsive to network structure changes, particularly friction reductions that target the initially strongest bilateral links.

The model features a set of risk averse professionals (agents) positioned in social network. In every period, multiple clients arrive for each agent, who makes a distinct, discrete product choice on behalf of each. Agents do not know the true quality of each available product, but instead hold idiosyncratic beliefs about these product qualities that evolve over time. In particular, agents update idiosyncratic beliefs as information arrives from two sources: from the observation of their own clients' outcomes, and in addition, from learning through the social network about the outcomes of others' clients. The model specifies that agents receive more signals, and hence learn more from, nearby professionals with high decision volumes. Thus, while each agent is connected to every other agent in the network, the degree of connectedness differs across links, with strong links implying a large flow of signals per period, and with weak links implying the opposite.

We show that this setup indicates a straightforward link between agents' unobserved beliefs about product qualities and their observable product choice shares. Moreover, when combined with our network structure and Bayesian learning assumptions, this further implies a nonlinear estimating equation characterizing the change in an agent's relative prescription choice shares as a function of three fundamentals: first, the precision of an agent's own initial beliefs; second, the distribution (mean and variance) of signals traveling over the network; and third, the complete set of bilateral network links connecting each pair of agents. Because separately estimating N^2

[1] For example, knowledge frictions across individual agents have been shown to partially explain the slow diffusion of a cost-reducing innovation (Griliches 1957), as well as deviations from the law of one price (Jensen 2007).

[2] Banerjee et al (2013, 2014) and Akbarpour, Malladi, and Saberi (2018) both provide important results about how to optimally target networked agents with an information seeding policy. These results are particularly relevant when the policymaker holds the information about which agents in the social network are learning.

unrestricted bilateral proximity parameters is infeasible given the large size of our dataset, we parameterize the connections between agents, as well as their initial beliefs precision levels, as functions of observable attributes. We show that, under this parameterization, the model can be estimated using a simple, restricted nonlinear least-squares estimator.

This estimator is applied to quantify the model using comprehensive data from IMS Health (now IQVIA) on the cholesterol-drug prescription decisions of U.S. physicians over the 11-year period spanning January 2000 through December 2010. For each doctor and month, we observe the number of prescriptions corresponding to each of the cholesterol therapies available in that month, as well as the location (five-digit zipcode) and a decomposition of prescriptions by patient insurance type for each doctor. We match these data with information on each doctor's primary medical specialty, medical school attended, medical school graduation year, total and cholesterol-drug specific advertising exposure, and location-specific household income and population density. Consistent with agents learning about unconditional product qualities, the prescription data reveal substantial shifts in aggregate choice shares over the sample period (Figure 1) that differ across U.S. locations (Figures 2 and 3).

Our estimates indicate that geographic and cohort proximity as well as shared medical school and specialty are all statistically important determinants of network connectedness; moreover, it is primarily distance variation that explains the shape of network connectedness across agents. Using our product quality estimates, we further demonstrate that—although our estimation procedure does not restrict this to be the case—agents' choice shares are, on average, evolving over time in a way that is consistent with agents' idiosyncratic beliefs converging to the true product qualities, a fundamental prediction of the model. We find that individual convergence rates are faster for agents that are relatively well-connected to the network and for agents with relatively imprecise initial beliefs, also in line with the predictions of the model. On the other hand, our estimates indicate convergence rates are not systematically related to doctors' exposure to direct pharmaceutical advertising.

To understand the implications of the model for how information barriers shape the rate of knowledge diffusion, we simulate our quantified model under a series of counterfactual scenarios. These suggest that policies impacting the structure of the network itself may be particularly influential, especially when efforts to strengthen links target the network ties that are already the strongest. By contrast, we find a relatively limited role for interventions targeting either weak network links or the initial distribution of information across individual agents. Specifically, providing certain individuals with additional signals in the initial period, which increases the precision of their prior beliefs, or even providing agents the true product qualities, has only a negligible impact on the aggregate rate of convergence. It is also important to note that this latter form of policy intervention is unavailable in the context of our model, as the true product qualities are fundamentally unobserved in real time. While one may estimate these *ex post* given historical data, it would actually not be feasible to learn these qualities parameters at the time such knowledge could have been valuable for information injection purposes.

This paper is related to a growing literature on knowledge diffusion. In particular, the theoretical framework in our paper resembles Eaton and Kortum (1996) in that we model individual's

knowledge evolution as determined, in part, by ideas arriving from others, with ideas arriving more intensely from relatively nearby and informed sources. Our panel data allow us to estimate the model using observed dynamics in agents' product choices rather than cross-section differences in productivity. Although we do not observe a direct measure of idea flows between physicians, we are nevertheless able to estimate the role of the network in governing knowledge flows by relying on the assumptions of Bayesian learning and static multinomial choice. This paper is also conceptually related to Buera and Oberfield (2019) in that our model implies reductions in network frictions, related to the idea of trade liberalization, increase the rate of knowledge diffusion, thus increasing the rate of efficiency gain as agents' beliefs thereby converge more rapidly to the truth.

More broadly, recent work considers firm networks as key for explaining trade flows, and in particular has emphasized that the structure of network connections is important for both firm-level and aggregate outcomes that may result from infrastructure improvements or reductions in variable trade costs (Chaney 2014, Bernard and Dhingra 2016, Bernard and Moxnes 2018, Bernard, Moxnes, and Saito 2019). Our model considers a related question in a distinct context. In particular, we consider a parallel set of counterfactuals, but focus on different outcomes: instead of considering the efficiency of bilateral, buyer-supplier matches that result after a reduction in bilateral frictions, we evaluate the quantitative impact of this reduction in frictions on efficiency gains as agents' rate of knowledge acquisition rises, improving decision quality. We further assess the value of targeting reductions in frictions to certain parts of the network, and show that aggregate efficiency is especially responsive to friction reductions aimed at the initially strongest network links.

Our model and results are related to recent work evaluating the dynamics of social learning in networks (Banerjee et al 2013, 2014, Akbarpour, Malladi, and Saberi 2018). We emphasize that the empirical model we develop may be used to characterize transition dynamics as knowledge diffusion through the social network leads to aggregate efficiency gains. For this, our empirical approach relies on techniques developed in models of static discrete choice (Train 2009) and of product choice under Bayesian learning (Crawford and Shum 2005).

Because our quantification of the model involves a medical context, our work is also related to two literatures in health economics. First, our framework is related to models of learning in pharmaceutical markets including Erdem and Keane (1996), Ackerberg (2003), Crawford and Shum (2005), and Arrow, Bilir and Sorensen (2018). We build on this work by modeling agents' learning about unobserved drug qualities as determined in part by the social network: in our model, agents learn not only by their own experience, but also from the experiences of their social contacts. This, combined with the population-level data we analyze, implies our analysis has direct implications both for individuals' learning and efficiency gains and for aggregate efficiency gains.[3]

Second, our model and results have implications related to the work documenting variations in U.S. medical care (e.g. Wennberg et al 1996, Munson et al 2013, and Cooper et al 2015). We contribute to this work by establishing a network-learning mechanism that can rationalize static

[3]Our results also add to the literature examining the determinants of new medical technology diffusion including Coleman, Katz, and Menzel (1957, 1996), Skinner and Staiger (2007), and Agha and Molitor (2015). Note that, because our data do not include individual patient characteristics, we are not able to estimating a model featuring learning about match quality within each patient-physician pair, as in Crawford and Shum (2005).

productivity disparities, and also explain how these disparities evolve over time. In addition, our quantified model can provide guidance for policies aimed at reducing knowledge-based disparities by intervening to change the strength of network connections and the distribution of initial information.

The rest of the paper is organized as follows. Section 2 presents a learning model of product choice under uncertainty. Section 3 outlines the baseline estimation strategy. Section 4 describes the data and provides descriptive evidence about aggregate evolution of prescription choice shares. Section 5 presents our main estimates, section 6 provides estimates for alternative specifications, and section 7 discusses quantitative implications of these estimates. Section 8 concludes. Derivations and additional results may be found in the Appendix.

2 Empirical Model

This section presents an empirical model of learning among individual professionals positioned in a social network. The framework is developed in general terms; in sections 3 and 4, we go on to provide estimation and data details specific to the medical setting we evaluate as a quantitative application of the model.

2.1 Setup

Consider a set $\mathcal{I}$ of agents (indexed by either i or j) arrayed on a network. Suppose that in each period $t = 1, 2, ..., T$, agent i faces the arrival of a measure R_{it} of clients ν_i. For each client ν_i, i chooses a single product (indexed by d) in period t from a set $\mathcal{D}_t$ of available products. Suppose client ν_i receives the following reward from product d at t

$$u_{dt}(\varepsilon_{dt}(\nu_i)) = \beta_d^T + \varepsilon_{dt}(\nu_i),$$

where β_d^T captures the true unconditional quality of product d. These true qualities $\{\beta_d^T\}_{d\in\mathcal{D}_t}$ are not fully known to agent i. In particular, we assume that for each d, i holds beliefs about the value of β_d^T at t that are summarized by a normal distribution with mean β_{idt} and variance σ_{idt}^2. The match values $\varepsilon_{dt}(\nu_i)$ are, on the other hand, observed perfectly by agent i at t (though not by the econometrician); these are assumed to follow a Gumbel distribution $F(\varepsilon)$ with shape parameter θ.[4]

Letting $U_{idt}(\varepsilon_{dt}(\nu_i), \beta_d)$ denote the agent-i utility of choosing product d for client ν_i at t given beliefs $\beta_d \sim N(\beta_{idt}, \sigma_{idt}^2)$ about the quality of d, we specify that

$$U_{idt}(\varepsilon_{dt}(\nu_i), \beta_d) = -\exp(-\alpha(\beta_d + \varepsilon_{dt}(\nu_i))),$$

where $\alpha > 0$ is the coefficient of absolute risk aversion. This implies (see Appendix A.1) that the

[4] That is, $F(x) = P\{\varepsilon_{dt}(\nu_i) \leq x\} = e^{-e^{-x/\theta}}$.

expected utility for i of choosing product d for client ν_i at t is, given her current beliefs,

$$\begin{aligned}\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i), \beta_{idt}, \sigma_{idt}) &\equiv \mathbb{E}_{\beta_d}[U_{idt}(\varepsilon_{dt}(\nu_i), \beta_d)|\beta_{idt}, \sigma_{idt}] \\ &= -\exp\left(-\alpha\beta_{idt} + \frac{\alpha^2\sigma_{idt}^2}{2} - \alpha\varepsilon_{dt}(\nu_i)\right). \end{aligned} \tag{1}$$

For each client ν_i, i thus compares (1) across products $d \in \mathcal{D}_t$ and chooses that which maximizes i's expected utility from the product choice for ν_i,

$$U_{it}(\boldsymbol{\varepsilon_t}(\nu_i), \boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t) \equiv \max_{d\in\mathcal{D}_t}\{\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i), \beta_{idt}, \sigma_{idt})\},$$

where $\boldsymbol{\beta}_{it}$, $\boldsymbol{\sigma}_{it}$, and $\boldsymbol{\varepsilon_t}(\nu_i)$ are vectors that contain, respectively, each product-specific value of β_{idt} and σ_{idt} corresponding to i at t, as well as each product-specific value of $\varepsilon_{dt}(\nu_i)$ corresponding client ν_i at t. Considering all of the clients served at t, i's period payoff $W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t)$ is therefore

$$W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t) = R_{it}E_\varepsilon[U_{it}(\boldsymbol{\varepsilon_t}(\nu_i), \boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t)]. \tag{2}$$

2.2 Conditional Product Choice

We first characterize the decision rule for each agent, and the corresponding choice shares, conditional on a set of current beliefs. We then characterize the expected utility of agents given their optimal choice shares and beliefs. Specifically, let

$$d_{it}^*(\boldsymbol{\varepsilon_t}(\nu_i), \boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = \operatorname*{argmax}_{d\in\mathcal{D}_t}\{\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i), \beta_{idt}, \sigma_{idt})\}$$

be the optimal product chosen for client ν_i at t by agent i. Letting $\pi_{idt}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it})$ denote the share of clients for whom i chooses product d during period t, the law of large numbers implies,

$$\pi_{idt}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = \Pr\{d = d_{it}^*(\boldsymbol{\varepsilon_t}(\nu_i), \boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it})\}.$$

Applying this, along with the distributional assumption for $\varepsilon_t(\nu_i)$, the following result summarizes the relationship between agent-i's conditional choice shares and current beliefs about the quality of each product $d \in \mathcal{D}_t$.

Result 1: *Given agent i's current beliefs summarized by $(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it})$ about product qualities $\{\beta_d^T\}_{d\in\mathcal{D}_t}$, the share of product d in agent i's overall decision outcomes in period t is*

$$\pi_{idt}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = \frac{\exp(\beta_{idt} - \alpha\sigma_{idt}^2/2)}{\sum_{d'\in\mathcal{D}_t}\exp(\beta_{id't} - \alpha\sigma_{id't}^2/2)}. \tag{3}$$

The proof for Result 1 appears in Appendix A.2. This result has straightforward implications for the expected period payoff of agent i, which we characterize next.

2.3 Expected Utility with Optimal Product Choice

Specifically, given that each agent i serves a continuum of clients per period, and in light of Result 1, it is possible under the assumptions above to characterize the expected payoff of i at t given i's current beliefs. The following result thus characterizes i's expected period utility conditional on her current beliefs.

Result 2: *Given agent i's current beliefs and the corresponding optimal product choices made on behalf of clients in period t, the expected utility* (2) *of agent i in period t is*

$$W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t) = -R_{it} \times \Gamma(1+\theta\alpha) \times \left(\sum_{d \in \mathcal{D}_t} \exp\left[\frac{1}{\theta} \left(\beta_{idt} - \frac{\alpha \sigma_{idt}^2}{2} \right) \right] \right)^{-\theta\alpha}. \tag{4}$$

The proof for Result 2 appears in Appendix A.3. Equation (4) indicates the agent-i period payoff is increasing in i's patient volume R_{it}, i's perceived mean quality of each drug $\boldsymbol{\beta}_{it}$, and, importantly, the precision of i's beliefs $\boldsymbol{\sigma}_{it}$ about the quality of each drug, with the latter effect amplified by the degree of risk aversion α. With this result in hand, we now specify the assumptions that govern the determination and evolution of agent-i beliefs. To highlight the network structure of the model, we focus on the case in which the evolution of beliefs is exogenously determined.

2.4 Evolution of Beliefs: Knowledge Diffusion in a Network

Suppose that in the initial period $t = 0$, agent i has a set of prior beliefs about the true quality β_d^T of each product d. In particular, assume that for each d, agent i's initial beliefs about β_d^T are summarized by a normal distribution with mean β_{id0} and variance σ_{id0}^2. Let S_{i0} denote the general precision level of agent i's prior beliefs, and assume it is common across products so that $\sigma_{id0}^2 = \sigma_d^2 / S_{i0}$, where σ_d^2 is the fundamental variance associated with d. Agents with relatively high levels of precision thus also have beliefs with relatively low levels of initial variance.[5]

In each period t, suppose agent i receives f_{ijt} product quality signals from every other agent $j \in \mathcal{I}$ in the network (including $j = i$). Assume further that each signal reflects new information realized at t as the outcome of agent j's own period-t client outcomes, and is a vector with one element per product $d \in \mathcal{D}_t$. While we assume new signals contribute to the precision of posterior beliefs, this contribution is subject to decay over time. Specifically, we assume that in period t, the precision of agent i's beliefs is determined by

$$S_{it+1} = \delta S_{it} + \sum_{j \in \mathcal{I}} f_{ijt},$$

[5] Assuming differences across agents i in the precision S_{i0} of initial beliefs captures that, in many empirical settings, including the one we consider below as a quantitative application, the initial observation in the data includes agents that are heterogenous in both professional tenure (age) and decision volume (R_{it}). Along with other characteristics, these may contribute to empirically relevant differences in the precision of initial beliefs by the first observation date. Additional data details appear in section 4.

where $\delta \in [0,1]$ governs the rate of signal decay. Similar to Crawford and Shum (2005), we assume that the value x_{ijdnt} of each signal $n = 1, ..., f_{ijt}$ received by i from j about product d at t is an independent draw from an unbiased normal distribution with mean β_d^T and variance σ_d^2. Notice that we may therefore express the mean value of signals received by i from j about d at t as follows, $\overline{x}_{ijdt} = \sum_{n=1}^{f_{ijt}} x_{ijdnt}/f_{ijt} = \beta_d^T + e_{ijdt}$, where $e_{ijdt} \sim N(0, \sigma_d^2/f_{ijt})$. We specify that the volume of signals f_{ijt} received by i from j at t is exogenous and proportional to both the network proximity τ_{ij} between i and j, and the measure R_{jt} of clients served by j in period t

$$f_{ijt} = \tau_{ij} R_{jt}.$$

Thus, i receives more signals from agents j that are relatively nearby in the network and that generate relatively more information through client decision outcomes of the current period.

Agent i updates her beliefs in each period on the basis of this information. Importantly, this setup implies agent i's posterior beliefs in period $t+1$ may be expressed recursively as a sequence of normal distributions (e.g. DeGroot 1970) with mean

$$\beta_{idt+1} = \frac{\beta_{idt}\delta S_{it} + \sum_{j \in \mathcal{I}} (\beta_d^T + e_{ijdt}) f_{ijt}}{S_{it+1}}, \tag{5}$$

where the discount factor δ, precision S_{it}, flow of new signals f_{ijt}, true product-d quality β_d^T, and noise e_{ijdt} are defined above. The variance of i's posterior beliefs is similarly updated as follows

$$\sigma_{idt+1}^2 = \sigma_d^2 / S_{it+1}. \tag{6}$$

3 Estimation

In this section, we derive estimating equations using the model described in section 2. We show that these equations may be combined with available data to recover the parameters governing the extent of knowledge diffusion across individual agents in the network and the corresponding rate of change in agents' product choices.

3.1 Estimation approach

To estimate the model, consider a spell during which the set of products remains unchanged. Rearranging terms in the choice share equation (3) implies the following relationship between choice shares π_{idt} and unobserved mean beliefs (β_{idt}, σ_{idt}^2) about the quality of product d,

$$\beta_{idt} = \ln \pi_{idt} + \frac{\alpha \sigma_{idt}^2}{2} + \eta_{it}, \tag{7}$$

where $\eta_{it} \equiv \ln \left(\sum_{d' \in \mathcal{D}_t} \exp(\beta_{id't} - \alpha \sigma_{idt}^2/2) \right)$ is an individual-month specific unobserved term. The evolution of mean beliefs β_{idt} is captured by the Bayesian updating rule (5), which combined with

the static discrete choice equation (7) above implies

$$\ln \pi_{idt+1} + \frac{\alpha\sigma^2_{idt+1}}{2} + \eta_{it+1} = \frac{\delta S_{it}}{S_{it+1}} \left(\ln \pi_{idt} + \frac{\alpha\sigma^2_{idt}}{2} + \eta_{it} \right) + \sum_{j \in \mathcal{I}} \frac{\tau_{ij} R_{jt}}{S_{it+1}} (\beta^T_d + e_{ijdt}).$$

Differencing with respect to a reference good d', and using the variance updating rule $\sigma^2_{idt+1} = \sigma^2_d / S_{it+1}$ in (6), we thus arrive at the following equation

$$\begin{aligned}[\ln \pi_{idt+1} - \ln \pi_{id't+1}] - [\ln \pi_{idt} - \ln \pi_{id't}] = \\ - \frac{\alpha(1-\delta)(\sigma^2_d - \sigma^2_{d'})}{2S_{it+1}} + \sum_{j \in \mathcal{I}} \frac{\tau_{ij} R_{jt}}{S_{it+1}} (\beta^T_d - \beta^T_{d'} - [\ln \pi_{idt} - \ln \pi_{id't}]) + u_{idt}, \quad (8)\end{aligned}$$

where the precision of agent-i beliefs at t follows $S_{it+1} = \delta S_{it} + \sum_{j \in \mathcal{I}} \tau_{ij} R_{jt}$, and the error term is $u_{idt} \equiv \sum_{j \in \mathcal{I}} \tau_{ij} R_{jt} (e_{ijdt} - e_{ijdt'}) / S_{it+1}$.

Notice that while (8) summarizes the restrictions implied by the model for the prescription shares of agent i, product d, and month t given all underlying parameters, the set of parameters in (8) includes N^2 proximity terms τ_{ij}, N initial precision levels S_{i0}, and $2(|\mathcal{D}_t| - 1)$ relative drug quality and variance terms. Even with a relatively small population of $N = 5,000$ individual agents, and even with a sufficiently large set of product-time observations per agent, this would amount to over 25 million parameters, the estimation of which, given the nonlinearities in (8), is infeasible. We therefore redefine both the proximity τ_{ij} and precision S_{i0} terms as functions of observable variables. In particular, we propose that the proximity between agents i and j may be summarized by a linear function of observable bilateral proximity variables including geographic proximity, the extent of professional experience, professionals school attended, and so on. We specify that $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$, where $\boldsymbol{b}_\tau$ is the vector of coefficients governing the contribution of bilateral proximity variables $\boldsymbol{Y}_{ij}$ to the strength of the social network connection between agents i and j. Similarly, we specify that the initial precision of agent-i beliefs may be summarized by a linear function of observable agent-i characteristics, including local demographics, professional specialty, experience, and per-period choice volume among others. Thus, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$, where $\boldsymbol{b}_S$ is the vector of coefficients governing the initial precision of agent-i beliefs as a function of observed period-0 characteristics $\boldsymbol{X}_{i0}$. Using these two parameterizations, we modify (8) above to arrive at

$$\begin{aligned}[\ln \pi_{idt+1} - \ln \pi_{id't+1}] - [\ln \pi_{idt} - \ln \pi_{id't}] = \\ - \frac{\alpha(1-\delta)(\sigma^2_d - \sigma^2_{d'})}{2S_{it+1}(\boldsymbol{b}_\tau, \boldsymbol{b}_S, \delta)} + \sum_{j \in \mathcal{I}} \frac{\tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij}) R_{jt}}{S_{it+1}(\boldsymbol{b}_\tau, \boldsymbol{b}_S, \delta)} (\beta^T_d - \beta^T_{d'} - [\ln \pi_{idt} - \ln \pi_{id't}]) + u_{idt}, \quad (9)\end{aligned}$$

where the following restriction on the precision of beliefs is satisfied in every period

$$S_{it+1}(\boldsymbol{b}_\tau, \boldsymbol{b}_S, \delta) = \delta^{t+1} S(\boldsymbol{b}_S, \boldsymbol{X}_{i0}) + \sum_{u=0}^{t} \delta^{t-u} \sum_{j \in \mathcal{I}} \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij}) R_{ju}. \quad (10)$$

To estimate (9), we set the quarterly discount factor to $\delta = 0.9873$, implying an annual discount factor of 0.95 and assume a value for the coefficient of absolute risk aversion α using an estimate from the recent literature.[6] In addition, to establish the scale of the estimation, we set one of the parameters of the function $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$ to take a value of 1; notice that without fixing the scale, there are multiple sets of parameters that solve (9). With these restrictions, the set of unknown parameters to estimate in this single-spell baseline model is thus $\Theta = \{\{\beta^T, \sigma^2\}_{d \in \mathcal{D}_t \setminus d'}, \boldsymbol{b}_\tau, \boldsymbol{b}_S\}$. Estimates of Θ may be recovered by evaluating (9) above using nonlinear least squares, restricting the beliefs precision levels according to (10).

To capture a time horizon featuring changes in the set of available products $\mathcal{D}_t$, we generalize the specification above. In particular, we divide the sample period into distinct spells $k = 1, 2, .., K$ demarcated by product entry events, and allow the β_d^T to differ across spells. We thus estimate a modified version of (9),

$$[\ln \pi_{idt+1} - \ln \pi_{id't+1}] - [\ln \pi_{idt} - \ln \pi_{id't}] = \\ - \frac{\alpha(1-\delta)(\sigma_d^2 - \sigma_{d'}^2)}{2S_{it+1}(\boldsymbol{b}_\tau, \boldsymbol{b}_S, \delta)} + \sum_{j \in \mathcal{I}} \frac{\tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij}) R_{jt}}{S_{it+1}(\boldsymbol{b}_\tau, \boldsymbol{b}_S, \delta)} (\beta_{dk}^T - \beta_{dk'}^T - [\ln \pi_{idt} - \ln \pi_{id't}]) + u_{idt}, \quad (11)$$

and the set of parameters is $\Theta = \{\{\beta_{dk}^T\}_{d \in \mathcal{D}_t \setminus d', k \in 1,..,K}, \{\sigma^2\}_{d \in \mathcal{D}_t \setminus d'}, \boldsymbol{b}_\tau, \boldsymbol{b}_S\}$.

With an estimate $\hat{\Theta}$ in hand, it is straightforward to estimate the standard error for each of its elements. In particular, the standard error for the k^{th} parameter of Θ is the square root of the k^{th} diagonal element of the estimated variance-covariance matrix

$$V = (\hat{J}'\hat{J})^{-1} \hat{J}' \hat{\Omega} \hat{J} (\hat{J}'\hat{J})^{-1},$$

where $\hat{J}$ is the Jacobian matrix evaluated at $\hat{\Theta}$, and $\hat{\Omega} = diag(\hat{u}_{idt}^2)$ for residuals u_{idt} defined by either (9) or (11) above when evaluated at $\hat{\Theta}$.

4 Data and Measurement

In this section, we describe the context and dataset we use for our quantitative application of the model, and present descriptive statistics on the population of agents, product choice set, and product choice dynamics at both aggregate and individual levels.

4.1 Prescriptions by U.S. Physicians

Estimating the model described in section 2 above requires data on all product choice decisions for agents positioned in the social network. Decisions by each agent must be observed for the full product choice set $\mathcal{D}_t$, and on a repeated basis over a period of time. Finally, it is necessary to observe the location and other characteristics of each individual agent that may be relevant for determining the network structure.

[6] The value $\alpha = 0.99$ is from Crawford and Shum (2005); see section 5.

Our quantitative application relies on a dataset from IMS Health (IQVIA) that satisfies each of these requirements. In particular, our analysis uses physician-level prescription choice data from the IMS Health Xponent database for the complete class of drugs targeting cholesterol control.[7] These data provide a quarterly prescription count during January 2000 through December 2010 for each U.S. prescriber associated with a minimum of ten cholesterol-drug prescriptions in 2010. Given the low level of this threshold, our dataset includes the full decision history for nearly all U.S. cholesterol drug prescribers during this period. Importantly, the comprehensive coverage of our data imply that our estimation procedure is able to avoid sampling bias that may arise in network settings with partial data (e.g. Chandrasekhar and Lewis 2016).

Each prescriber in the Xponent dataset is identified not only by a unique medical education number, but also by name (first name, last name, middle name), and location (a five-digit U.S. zipcode). Using these latter identifiers, we match each prescriber in our Xponent database with a) attributes data contained in the CMS Physician Compare database, and b) drug marketing and advertising data disclosed under the Physician Payments Sunshine Act, part of the 2010 Affordable Care Act, which were obtained from ProPublica. Observable attributes in a) include the medical school attended, medical school graduation year, primary medical specialty, and gender.[8] In addition, beginning in 2006, our Xponent prescription data are reported separately according to four patient insurance-coverage categories: Medicare, Medicaid, privately-insured, and cash payer. We use these insurance data to construct a time-invariant public insurance share for each doctor i.[9] Finally, we include location-specific demographic data (county-level household income and population density) from the U.S. Census 2000, matched to physicians based on a link between five-digit zipcodes and U.S. counties. Using the advertising data for pharmaceutical products b), we similarly construct time-invariant, doctor-specific measures of drug marketing exposure at two levels: including all drugs, and restricting attention to the class of cholesterol drugs as in our quantification of the model. This measurement approach is motivated by the delayed coverage of the advertising data, which span the dates August 2013 through December 2015.[10]

Importantly, because the attributes data described above help to ensure that our characterization of physicians, including the physician social network, is sufficiently rich, our quantitative analysis restricts attention to prescribers appearing in the Xponent data, the Physician Compare dataset, and the Propublica advertising dataset. Note that the CMS Physician Compare database includes essentially the full population of *physicians*; prescribers excluded from our analysis are primarily nurse practitioners, registered nurses, clinical pharmacists, and physician assistants who

[7] IQVIA, formerly IMS Health, also maintains data on additional drug classes. These additional data are not available for this study, however, as the customized data extracts involved are unusually large.

[8] The data also include a group practice identifier and an identifier for the hospital at which the doctor has admitting privileges, but in practice these attributes overlap substantially with the five-digit zipcode and thus are currently not included in our analysis.

[9] While it would have been ideal to consider the full detail of the patient insurance data, we take the approach of building a fixed coverage index as these data are available only during the final five years of the sample period.

[10] Information relevant to measuring dynamics in the physician network are also observed, including a) changes in an individual doctor's location (zipcode), and b) entry of new doctors. The latter is likely to be a relevant source of network variation, as approximately 40 percent of matched physicians enter the dataset after the completion of the first quarter observed.

may be viewed as inherently less likely to interact as agents within the physician social network.[11]

Because it is unlikely that different formulations of the same molecule have relevant differences in perceived quality, we aggregate drugs to the molecule level; for example, Mevacor and its generic equivalent, lovastatin, are treated as the same product in our empirical analysis. To account for (unobserved) differences in pricing that may coincide with the introduction of a new or generic version of a drug, however, we allow true molecule qualities (the common patient payoff β_d^T) to differ in our estimation across spells, demarcated respectively by the successive introduction dates of three generic versions and three new drugs containing novel molecules. In summary, for our application, an agent is thus a doctor, a client is a patient, a period is a quarter, and a product is a molecule, the chemical compound that defines an associated prescription drug.

As an important goal in our analysis is to determine the aggregate implications of our individual-level learning model (e.g. the evolution of aggregate choice shares), we restrict attention to a) physicians with total prescription volumes at the 25^{th} percentile or above, and b) molecules with at least three percent of the total prescription volume over the sample period. With these two restrictions in place, we retain over 90 percent of the prescriptions, while reducing the presence of zeros in the prescription share data by a factor of approximately three.[12] Our final estimation sample has 53,040 doctors, 44 quarters, and up to 7 molecules.

4.2 The Choice Set and Drug Entry

While the U.S. market for cholesterol drugs is already of immediate interest given its size, an important feature of this drug class for our analysis is the number of significant events over the sample period affecting this class of drugs.[13] Specifically, in the initial quarter in the data, four relevant products were available: Lipitor, Mevacor, Pravachol, and Zocor.[14] Patents expired for three of these products during the sample period, resulting in the entry of generic lovastatin (December 2001), pravastatin (April 2006), and simvastatin (June 2006). In addition, three drugs based on new molecules were approved for sale in the United States during the sample period: Crestor (August 2003), Vytorin (July 2004), and Zetia (November 2004). The complete list of molecules and respective introduction dates appears in Table 1.

These six entry events are important for our analysis because they are consistent with the idea that physicians' effective knowledge depreciates over time (captured by δ in our model). This, in turn, provides a motive for doctors' ongoing acquisition of knowledge. For example, suppose that some or even all physicians begin their medical career fully informed regarding the relative patient

[11] Specifically, the Physician Compare database includes data for each physician treating patients that participate in Medicare or Medicaid. Recent work shows that between 92 and 96 percent of U.S. physicians accept Medicare patients during the sample period (Bishop, Federman, and Kayhani 2011).

[12] Gandhi, Lu, and Shi (2017) describes the bias that may arise in the presence of zeros in choice-share data.

[13] Chronic hypercholesterolemia and dyslipidemia, conditions in which abnormal levels of cholesterol or lipids are present in the bloodstream, are common in the United States: according to the Centers for Disease Control and Prevention, approximately 71 million U.S. adults suffer these chronic conditions. As a result, cholesterol drug sales amounted to over $18 billion U.S. dollars in 2011 (Ledford 2013). In addition, two of the drugs considered in our estimation, Lipitor and Crestor, are among the top ten pharmaceutical drugs according to 2010 U.S. sales. For further discussion of this drug class, see Arrow, Bilir, and Sorensen (2019).

[14] Lescol was also available during this time, but due to its negligible share in total prescriptions and disproportionately high share of zero-shares in the data, it is omitted from our analysis as noted above.

payoffs of the different currently-available cholesterol drugs; if the product space is static, it is unclear why such doctors rationally pay attention to signals about drug quality that other doctors may send. Conversely, if knowledge depreciates due to changes in the set of available products and their pricing, agents will rationally pay attention to signals about product quality. This feature of our empirical setting is thus important for the assumptions of our model to be satisfied.

The seven molecules considered in our analysis are therapeutic substitutes: each aims at the clinical endpoint of cholesterol or trigliceride reduction, and patients are typically prescribed only one molecule rather than multiple products.[15] Importantly, however, these molecules are only imperfect substitutes. First, while most of the cholesterol therapies we study are pure statins, which act to reduce cholesterol synthesis in the liver by inhibiting a specific coenzyme [these include Lipitor (atorvastatin), Mevacor (lovastatin), Pravachol (pravastatin), Zocor (simvastatin), and Crestor (rosuvastatin)], other products rely on different mechanisms of action. In particular, Zetia (ezetimibe), and thus also Vytorin (ezetimibe and simvastatin), are distinct in that cholesterol reduction is achieved by reducing intestinal absorption of cholesterol. Beyond mechanisms, the molecules we consider differ in therapeutic intensity. High doses of Lipitor and Crestor are typically more effective at lowering low-density lipoprotein (LDL) cholesterol than alternatives (Law et al 2003), for example. Therapeutic intensity is relevant not only because it may imply different molecules are appropriate for different patients depending on disease severity, but also because evidence suggests it is correlated with the intensity of adverse side effects. For example, evidence suggests high doses of intense therapies such as Lipitor and Crestor may raise the incidence of adverse reactions, while also indicating that therapies such as Vytorin may for certain patients be more appropriate care for cases of severe cholesterol abnormality (Kastelein et al 2008).

As the above discussion suggests, abundant clinical evidence suggests that the benefits and risks associated with statins are heterogeneous across patients.[16] Our model captures this through the match quality terms $\varepsilon_{dt}(\nu_i)$ in the patient-level physician payoff function. Because our data do not include patient-level information, we abstract from the possibility that physicians learn about the value of $\varepsilon_{dt}(\nu_i)$.[17] The evidence presented below nevertheless indicates learning about molecules' average qualities is likely to be highly relevant given the substantial changes in aggregate market shares observed during the sample period.

4.3 Descriptive Statistics

Evolution in Aggregate Choice Shares

Figure 1 plots the evolution in aggregate product choice volumes during the sample period. Consistent with learning about β_d^T, substantial shifts in aggregate choice shares are evident. For example, the aggregate choice share of Lipitor is initially above 50 percent, but falls to approximately 20 percent by the end of the sample period. This opposite occurs for simvastatin which accounts for approximately 20 percent of the market initially, ultimately rising to around 45 per-

[15] See Arrow, Bilir, and Sorensen (2019) for a related discussion.

[16] See, for example, Brooks et al (2014).

[17] Crawford and Shum (2005) estimate a model of learning specifically about the patient-drug match for a distinct drug class.

cent. Figure 1 also reveals that the pattern of change in aggregate choice shares is gradual, in line with the learning mechanism of the model outlined in section 2.

Across U.S. locations, the evolution of choice shares follows distinct patterns relative to the aggregate shifts in Figure 1. For example, Figure 2 indicates New York, NY (three-digit zipcode 100), despite having similar initial choice shares, ultimately has larger shares for both Lipitor and Crestor, and smaller shares for each of the three generics. Figure 3 considers instead the relatively remote Hemet, CA (three-digit zipcode 925). Expansion in Hemet's generic choice share is large relative to Figures 1 and 2, while the Crestor share expands relatively less. These local differences in choice share evolution are consistent with persistent information differences across U.S. locations, but importantly, could also be explained by heterogeneity in patient payoffs β_d^T across locations.[18]

Table 2 thus considers the decision over whether to prescribe the branded versus generic version of a particular molecule. Specifically, for each of the three molecules that experienced generic entry during the sample period, Table 2 summarizes the distribution of generic prescription shares for that molecule at different time horizons. The advantage of focusing on these shares is that it is possible to compare prescribing of a branded product with its molecularly-equivalent generic, two distinct drugs that have no relevant clinical differences. And, by examining these shares at different time horizons, it is possible to determine whether stable heterogeneity in β_d^T across locations is likely to be the only explanation for differences in choice shares across locations. Consider lovastatin, for example, in column 1: after six months, the average generic prescription share for lovastatin—that is, the share of lovastatin plus Mevacor prescriptions that are accounted for by generic lovastatin—across sample physicians was 83.2 percent, with a doctor at the fifth percentile prescribing only Mevacor, the relatively expensive branded version. After 12 months, the average rises to 90 percent, but the fifth percentile physician still prescribes only Mevacor. By contrast, at the end of the sample period in December 2010, the average generic share is essentially 100 percent, and even the fifth percentile doctor prescribes exclusively generic lovastatin. This pattern of delayed substitution between two molecularly equivalent products is evident for each of the three generics in our dataset, and strongly suggests factors other than time-invariant patient heterogeneity contribute to prescribing differences across locations, and also indivdual physicians. Importantly, this pattern of delayed substitution is consistent with the influence of information frictions.[19]

5 Main Results

We estimate the model described in (10) and (11), specifying bilateral network proximity τ_{ij} as a function of individual characteristics observed in the data described above (section 4). In particular, the network proximity function is parameterized as

$$\begin{aligned}\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij}) = b_{\tau,g}\,\text{Geographic Proximity}_{ij} + b_{\tau,s}\,\text{Same Medical School}_{ij} \\ + b_{\tau,c}\,\text{Cohort Proximity}_{ij} + b_{\tau,sp}\,\text{Same Medical Specialty}_{ij},\end{aligned}$$

[18] In section 6 below, we extend our estimation approach to account for β_d^T heterogeneity across locations.
[19] See also Arrow, Bilir, and Sorensen (2019).

where Geographic Proximity$_{ij} \equiv (1+distance_{ij})^{-1}$ is the inverse of one plus the geographic distance separating doctors i and j. Cohort Proximity$_{ij}$ is the analogous measure replacing distance$_{ij}$ with the absolute difference between the respective medical school graduation years of doctors i and j. Same Medical School$_{ij}$ and Same Medical Specialty$_{ij}$ are dummies indicating whether doctors i and j either attended the same medical school, or practice within the same primary medical specialty. Similarly, the initial precision of doctor-i beliefs S_{i0} is parameterized as

$$\begin{aligned} S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0}) = b_{S,0} + b_{S,income}\,&\text{Mean Household Income}_{i0} + b_{S,density}\,\text{Population Density}_{i0} \\ &+ b_{S,r}\text{Decision Volume}_{i0} + b_{S,e}\,\text{Experience}_{i0} + b_{S,f}\,\text{Female}_i \\ &+ b_{S,c}\,\text{Cardiology}_i + b_{S,im}\,\text{InternalMedicine}_i + b_{S,fp}\,\text{FamilyPractice}_i. \end{aligned}$$

The first two variables, Mean Household Income$_{i0}$ and Population Density$_{i0}$, are demographic data from the U.S. Census 2000, and correspond to the county in which doctor i is located. Decision Volume$_{i0}$ (R_{i0}) is the number of doctor-i prescriptions for the period $t = 0$, Experience$_{i0}$ is the difference in years between $t = 0$ and the medical school graduation date of doctor i. Female$_i$ indicates whether doctor i is female, and Cardiology$_i$, InternalMedicine$_i$, and FamilyPractice$_i$ are primary medical specialty dummies indicating, respectively, whether i is a cardiologist, an internist, or in family practice. The latter three specialties account for approximately 90 percent of physicians in the data. As noted above, we set the annual discount factor to 0.95, implying a quarterly discount factor $\delta = 0.9873$. In addition, we use a value for the risk aversion parameter from the literature. Specifically, Crawford and Shum (2005) estimates that $\alpha = 0.990$ for a related empirical setting, and we use this value for our estimation.[20]

5.1 Baseline Estimates by Professional Specialty

We begin by considering separately each of the three major medical specialties in the data: cardiology, internal medicine, and family practice. In particular, we consider doctors i within each specialty to be positioned in a closed network consisting exclusively of other doctors within the same medical specialty as i. Corresponding estimates of (11) appear in Tables 3 through 5.

Estimates for cardiologists appear in Table 3, columns 1 and 3, with standard errors in columns 2 and 4 the right of each respective point estimate. Consider the estimates for bilateral proximity $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$ in column 1, which evaluates the complete network specification. These indicate that both a shared medical school and proximity in graduation dates are highly significant determinants of network connectedness, relative to geography. In particular, the coefficient on the indicator for shared medical school implies that, for geography to have the same contribution to τ_{ij} as a common medical school, cardiologists i and j would need to be located just four miles apart. The coefficient on graduation year proximity implies that if doctors i and j graduate from medical school within three years of each other, the contribution to τ_{ij} would be equivalent to the impact of the two doctors being separated by a distance of 61 miles. Nevertheless, variation in the underlying data

[20]Notice that α and δ are not separately identified from the signal variance terms σ_d^2 in (10). Thus, to the extent that setting $\alpha = 0.990$ or $\delta = 0.9873$ is incorrect, it would affect the magnitude of these variance estimates.

also determines the contribution of each component of bilateral connectedness to overall network proximity τ_{ij}. Interestingly, the data reveal that the relative importance of components is not uniform across the distribution of connectedness across doctors i. To understand the respective contributions of location, school, and graduation cohort to network proximity, Figure 4 plots the distribution of $\overline{\tau}_i \equiv \frac{1}{N}\sum_j \tau_{ij}$ across cardiologists i (solid line), which can also be interpreted as an index of network centrality, as well as the mean value of each subcomponent. It is clear that although the shape of the curve is determined by variation in graduation cohort among less-connected doctors, this is not true among relatively connected doctors, for whom variation in location is a more important determinant of the $\overline{\tau}_i$ distribution.

The estimates of doctors' initial beliefs precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$ indicate that household income, population density, decision volume, and years of experience are all statistically important determinants of an agent's beliefs precision, where recall that in the model, higher values of S_{i0} imply a slower rate of beliefs updating, all else equal. Of these four components, only experience enters negatively, suggesting those with more recent medical degrees have the most precise initial beliefs, and in turn, the slowest rates of beliefs updating conditional on network position. This result is consistent with the idea that knowledge depreciates over time, so that agents with less recent medical training—which is relevant because both the set of drugs and available clinical evidence about relative qualities change over time—have a relatively small effective stock of accumulated signals, all else equal, when compared with recent graduates. Figure 5 shows the distribution of implied S_{i0} values across cardiologists i (solid line) as well as each of its separate components. Similar to the case of network connectedness $\overline{\tau}_i$, the figure suggests the factors determining the shape of the S_{i0} differ across the distribution; in particular, S_{i0} is strongly correlated with experience among doctors with low S_{i0} values, but depends primarily on variation population density and decision volume for those with relatively high S_{i0} values. Table 3 also shows estimates of the true, unconditional drug qualities relative to the reference drug Lipitor, $\beta^T_{dk} - \beta^T_{d;k}$, for the first and last spells. These estimates indicate not only a clear ranking of products based on quality, but also relevant changes in this ranking between the spells. In particular, the estimated relative quality of simvastatin is higher than that of pravastatin in the first spell, which is, in turn, higher than that of lovastatin. While this relative ranking is maintained in spell 6, each of the relative quality estimates increases to a highly significant extent, in part reflecting the price changes that occurred as patents for each of these three molecules expired. Indeed, by the final spell, the positive estimate for simvastatin indicates its unconditional quality surpasses that of Lipitor. The final-spell estimates for new drugs indicate that the qualities of Zetia and Vytorin are both low relative to Lipitor, while the quality of Crestor is significantly higher than Lipitor. Estimated relative variance terms appear just below the quality parameters and indicate, unsurprisingly, that the signal distribution is relatively imprecise (high variance) for newer varieties, with the exception of Crestor which is in this respect indistinguishable from Lipitor.

Estimates for physicians specialized in internal medicine appear in Table 4, and estimates for family practice doctors are in Table 5. While proximity in graduation dates and a shared medical school are statistically important determinants of τ_{ij} in all three professional specialties, it is clear that the relative importance of a shared school is highest among cardiologists. Conversely,

the relative importance of distance is substantially higher for family practice physicians than for those in either of the other two specialties. These differences are also evident when comparing the components of network connectedness $\overline{\tau}_i$ among doctors in internal medicine (Figure 6) and family practice (Figure 8) with the analogous figure for cardiologists (Figure 4) discussed above. Specifically, Figure 8 indicates that distance variation determines the shape of $\overline{\tau}_i$ across essentially the complete population of family practice physicians, while internists resemble cardiologists in that variation in cohort proximity is predominant in the lower range of the distribution. Regarding the determinants of initial beliefs precision, comparing Figures 5, 7, and 9 reveal that the shapes of the respective S_{i0} curves are similarly determined by underlying components across medical specialties.

The estimates of unconditional drug qualities reveal an important distinction between cardiologists and physicians specialized in family practice. In particular, relative to Lipitor, cardiologists view Crestor favorably while other doctors do not; at the same time, family practice physicians are substantially more favorable about all generic molecules (lovastatin, pravastatin, and simvastatin) when compared with doctors in other specialties. Intuitively, this likely reflects sorting across specialties based on patients' disease severity. As noted above, Lipitor and Crestor are high-intensity products that are, accordingly, more appropriate care for those with severe cholesterol abnormalities. If patients with such conditions are disproportionately served by cardiologists, it is to be expected that cardiologists will choose Crestor and Lipitor high a higher propensity than the more general physicians specialized in internal medicine and family practice. Alternatively, another possibility is that patients seeing the latter physician types are more price-sensitive than those seeing cardiologists, and hence prefer to avoid relatively high-cost options like Crestor and Lipitor. This result is initially harder to support, as Zetia and Vytorin are also relatively expensive, active-patent molecules, yet both are viewed (relative to Lipitor) more favorably by family practice doctors than by cardiologists. To better distinguish between these mechanisms, we extend the model in section 6 below to include a time-varying measure of drug prices.

5.2 Network Specification with All Specialties

Table 6 provides estimates of the model using a sample of all doctor specialties and locations. In particular, given the size of the complete bilateral network ($53,040^2 = 2.8$ billion), we estimate this comprehensive model with a ten percent random sample of the 53,040 physicians.[21] Column 1 presents estimates of (11), while for comparison, column 3 imposes no time discounting, $\delta = 1$. This latter parameter restriction implies that the first term in (11) is equal to zero. Consider first the estimates in column 1. These indicate that each network proximity covariate—proximity in graduation dates, an indicator for shared medical school, and an indicator for common medical specialty—has a positive and highly significant contribution to the network connectedness of doctors in the data. Among these, a shared medical school is a particularly strong force for network connectedness. For two doctors i and j, attending the same medical school is associated with an increase in τ_{ij} that is, on average, equivalent to the contribution of geographic proximity if i and j were located only 9 miles apart. The distance equivalent for sharing a medical specialty is an

[21] Note that the region- and specialty-specific estimates are based on complete populations rather than samples.

order of magnitude larger (92 miles), while that for sharing the same medical school graduation year is four times larger (40 miles). The latter effect declines rapidly, moreover, and by ten years of separation in graduation years, the impact of proximity falls to a distance equivalent of 409 miles.

To understand the full contribution of each determinant of network connectedness, we combine the estimates with the data to compute each element of the implied τ_{ij} matrix. Taking the mean value for each doctor i, across all other doctors j, we build an index of mean network connectedness for each doctor that appears as the solid line in Figure 4. This line reveals a distribution centered around a median value of 0.0116, with a standard deviation of 0.004 and a skewness of 2.81 resulting from a long upper tail of highly connected physicians. Figure 5 decomposes this line into its four subcomponents, which indicate that, given the patterns in the data, geography is highly influential in determining the shape of the distribution. Highly connected doctors in the upper tail have unusually high geographic proximity to other doctors, but a similar range to doctors around the median for school, cohort, and specialty. On the other hand, the very least-connected physicians at the low end of the distribution have unusually low values among all four components of the τ_{ij} function. That variation in geography is a particularly important determinant of network connectedness is reinforced by the distance-only model considered in Table 3.[22] The distribution of the mean network connectedness index for this simpler model also appears in Figure 4 (dotted line) and follows almost exactly the shape of the distribution for the full network specification.

The estimates of the initial beliefs precision function S_{i0} in column 1 further indicate doctors in locations with high levels of household income and population density have significantly more precise prior beliefs about product quality. Doctors with high quarterly decision volumes also have significantly more precise initial beliefs. While experience and gender are not important determinants of initial beliefs' precision in the overall physician sample, medical specialty does impact S_i0. In particular, internists have the most precise initial beliefs, followed by family practitioners, and then by cardiologists, who have less-precise initial beliefs. These impacts of medical specialty are relatively large compared with the median implied value of S_{i0}, ranging from 15 to 25 percent of the median, and also raise the possibility that other relevant differences exist across medical specialties. Below, we thus estimate specialty-specific models that allow all parameters to differ depending on whether a physician is a cardiologist, internist, or family practitioner.

The specification further provides estimates for the parameters (mean and variance) of the signal distribution for each product. Considering first the true drug-d quality values β_{d1}^T for the initial spell, the estimates indicate that the signal distributions for Lovastatin, Pravastatin, and Simvastatin all have lower mean values than the distribution for Lipitor, in line with the initial dominance of Lipitor in the aggregate choice shares plotted in Figure 1. Also consistent with the market shares in Figure 1, the signal distribution for Simvastatin is higher in the first spell than that for Pravastatin, which is, in turn, higher than that for Lovastatin. Considering now the final spell, the signal distribution for Simvastatin rises to a level a significantly higher mean than that for Lipitor, and values for both Lovastatin and Pravastatin, while still negative, are also significantly higher than in the first spell. The substantial increases in β_{dk}^T for these three products

[22]The specifications in Table 3 consider network proximity τ_{ij} simply as a function of distance, restricting $b_{\tau,s}$, $b_{\tau,c}$, and $b_{\tau,sp}$ to take values of zero.

likely reflects the patent expiration and resulting price declines experienced by each molecule over the sample period, and moreover indicate the importance of allowing time variation in the signal distribution. Interestingly, Zetia and Vytorin have the lowest β_{dk}^T values in spell 6 and also the highest relative variance estimates, suggesting the signal distribution for these newer products is highly dispersed around a relatively low mean. On the other hand, the mean and variance values for Crestor are not significant, suggesting that the signal distribution parameters for Crestor are statistically indistinguishable from those of Lipitor.

Comparing the estimates with time discounting $\delta < 1$ in column 1 with those that restrict $\delta = 1$ in column 3 reveals that this restriction is primarily reflected by the β_{dk}^T estimates. The spell-1 signal value is higher for Lovastatin, and lower for both Pravastatin and Simvastatin. Interestingly, the estimated variance in the unrestricted model is positive for Lovastatin and negative for Pravastatin and Simvastatin, suggesting the restricted model inability to control for differences in signal dispersion may translate into biased quality estimates. Spell-6 mean signal values are higher for all drugs in the restricted model. These qualitative differences are also evident in Table 7, which replicates Table 6 with the added restriction that the network specification depends only on geographic proximity.

6 Robustness and Alternative Specifications

6.1 Advertising and Price Effects

The baseline specification estimated in section 5 above considers the product qualities β_{dk}^T about which agents learn as capturing not only the true, unconditional *efficacy* of product d, but also any other common determinants of product choice including product prices. While existing evidence supports the idea that physicians learn and are thus uncertain about drug prices, the model above considers quality as fixed within each spell k and is thus unable to address the possibility of within-spell price changes.[23] Indeed, if agents are sensitive to product price changes within a spell, some of the variation in relative choice shares may be explained by prices rather than by learning about β_{dk}^T. This may be relevant in the case of patent expirations, where initial generic entry is restricted, implying a larger long-run than short-run price impact for affected molecules.

Relatedly, the decision by physician i to choose drug d rather than an alternative for a given patient may depend not only on the unconditional quality of d, but also on i's incentives to choose d, including targeted advertising efforts by the firm promoting product d. This consideration is particularly important if drug advertising targets doctors differentially depending on the characteristics that determine learning rates in the model, including network connectedness τ_{ij} and initial beliefs precision S_{i0}. If this were to be the case, this underlying correlation could impact the interpretation of our estimates in a specification that omits advertising.

To address this concern, as well as the potential for within-spell price variation, we generalize the static choice equation to include two additional covariates, a) total drug advertising spending received by doctor i, and b) average U.S. pharmacy prices by molecule and date. Because our

[23] See, for example, Arrow, Bilir, and Sorensen (2019).

measure of advertising is time-invariant, and thus already incorporated in the static choice equation through η_{it}, we allow advertising exposure to differentially impact molecules that are still protected by active patents by interacting this measure with D_{dt}, which indicates the patent status of d at t and takes a value of 1 if the patent has expired and generic competition exists for d at t.[24,25]

With these changes, the static choice equation for agent i reflects price and advertising effects as follows

$$U_{idt}(\varepsilon_{dt}(\nu_i), \beta_d) = -\exp(-\alpha(\beta_d + \varepsilon_{dt}(\nu_i) + \xi P_{dt} + \gamma Ads_i \times D_{dt}))$$

where P_{dt} is the product-d price at t, Ads_i is pharmaceutical advertising spending received by agent i, and D_{dt} indicates generic competition for molecule d at t. Notice that normal price sensitivity would suggest $\xi < 0$, while the differential influence of marketing on choice for on-patent molecules suggests $\gamma < 0$.

With the preferences above, the expected, agent-i utility for choosing product d for client ν_i at t is, given her current beliefs,

$$\begin{aligned}\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i),&\beta_{idt}, \sigma_{idt}) \equiv \mathbb{E}_{\beta_d}[U_{idt}(\varepsilon_{dt}(\nu_i), \beta_d)|\beta_{idt}, \sigma_{idt}] \\ &= -\exp\left(-\alpha\beta_{idt} + \frac{\alpha^2\sigma_{idt}^2}{2} - \alpha\xi P_{dt} - \alpha\gamma Ads_i \times D_{dt} - \alpha\varepsilon_{dt}(\nu_i)\right). \qquad (12)\end{aligned}$$

Agent i thus selects the product $d \in \mathcal{D}_t$ that maximizes i's expected utility from the product choice for ν_i. As before, defining this optimized expected utility,

$$U_{it}(\boldsymbol{\varepsilon_t}(\nu_i), \boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t) \equiv \max_{d \in \mathcal{D}_t}\{\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i), \beta_{idt}, \sigma_{idt})\}$$

agent i's period payoff $W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t)$ is

$$W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t) = R_{it}E_{\varepsilon}[U_{it}(\boldsymbol{\varepsilon_t}(\nu_i), \boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}, \mathcal{D}_t)]. \qquad (13)$$

By the law of large numbers, it remains true that

$$\pi_{idt}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = \Pr\{d = d_{it}^*(\boldsymbol{\varepsilon_t}(\nu_i), \boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it})\}.$$

Applying this, along with the distributional assumption for $\varepsilon_t(\nu_i)$, physician i's conditional choice shares and current beliefs about the quality of each product $d \in \mathcal{D}_t$ are related according to the

[24]The advertising data are from ProPublica (see section 4 for details) and the price information is from the Marketscan Redbook database. The time coverage of the advertising data (2013-2015) unfortunately rule out including a more detailed marketing measure varying by doctor, drug, and quarter.

[25]Beyond both considerations, a doctor may take patient-level pricing into account when selecting a product. Although we do not have direct data on patient-specific pricing, beginning in 2006 our data indicate whether the purchasing patient paid for the product using a private insurance plan, Medicaid, Medicare, or cash. Using this information, we have constructed a time-invariant measure of public and private insurance shares that we interact with product prices as an additional covariate in the physician choice specification, but our estimates indicate that this is not a statistically important determinant of product choice.

expression

$$\pi_{idt}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = \frac{\exp(\beta_{idt} - \alpha\sigma_{idt}^2/2 + \xi P_{dt} + \gamma Ads_i \times D_{dt})}{\sum_{d' \in \mathcal{D}_t} \exp(\beta_{id't} - \alpha\sigma_{id't}^2/2 + \xi P_{dt} + \gamma Ads_i \times D_{dt})}. \tag{14}$$

To estimate the model with advertising and price effects, we note that (3) implies

$$\beta_{idt} = \ln \pi_{idt} + \frac{\alpha\sigma_{idt}^2}{2} - \xi P_{dt} - \gamma Ads_i \times D_{dt} + \eta_{it}, \tag{15}$$

where $\eta_{it} \equiv \ln\left(\sum_{d' \in \mathcal{D}_t} \exp(\beta_{id't} - \alpha\sigma_{idt}^2/2 + \xi P_{dt} + \gamma Ads_i \times D_{dt})\right)$ is an individual-month specific unobserved term. The evolution of mean beliefs β_{idt} is captured by the Bayesian updating rule (5), which combined with the static discrete choice equation (15) above implies

$$\begin{aligned}\ln \pi_{idt+1} + \frac{\alpha\sigma_{idt+1}^2}{2} - \xi P_{dt+t} - \gamma Ads_i \times D_{dt+1} + \eta_{it+1} = \\ \frac{\delta S_{it}}{S_{it+1}}\left(\ln \pi_{idt} + \frac{\alpha\sigma_{idt}^2}{2} - \xi P_{dt} - \gamma Ads_i \times D_{dt} + \eta_{it}\right) + \sum_{j \in \mathcal{I}} \frac{\tau_{ij} R_{jt}}{S_{it+1}}(\beta_d^T + e_{ijdt}).\end{aligned}$$

Differencing with respect to a reference good d', for which $D_{d't} = 0$ holds for all t, and using the variance updating rule $\sigma_{idt+1}^2 = \sigma_d^2/S_{it+1}$ in (6), we thus arrive at the revised equation

$$\begin{aligned}[\ln \pi_{idt+1} - \ln \pi_{id't+1}] - [\ln \pi_{idt} - \ln \pi_{id't}] = \gamma Ads_i \left(D_{dt+1} - \frac{\delta S_{it}}{S_{it+1}} D_{dt}\right) \\ + \xi\left(P_{dt+1} - P_{d't+1} - \frac{\delta S_{it}}{S_{it+1}}[P_{dt} - P_{d't}]\right) - \frac{\alpha(1-\delta)(\sigma_d^2 - \sigma_{d'}^2)}{2S_{it+1}} \\ + \sum_{j \in \mathcal{I}} \frac{\tau_{ij} R_{jt}}{S_{it+1}}(\beta_d^T - \beta_{d'}^T - [\ln \pi_{idt} - \ln \pi_{id't}]) + u_{idt},\end{aligned} \tag{16}$$

where the precision of agent-i beliefs at t follows $S_{it+1} = \delta S_{it} + \sum_{j \in \mathcal{I}} \tau_{ij} R_{jt}$, and the error term is $u_{idt} \equiv \sum_{j \in \mathcal{I}} \tau_{ij} R_{jt}(e_{ijdt} - e_{ijdt'})/S_{it+1}$. As in (11), specification (16) above can also be extended to allow for differences in product qualities β_d^T across spells k.

Estimates of (16) appear in Tables 8 through 11. Tables 8 and 9, which allow spell-specific β_d^T terms as in the baseline model, indicate that for each medical specialty (cardiology, internal medicine, and family practice) as well as for the sample that includes all doctor specialties, agents are highly responsive to advertising exposure. As expected, doctors with higher levels of direct pharmaceutical marketing payments prescribe significantly lower shares of off-patent molecules facing generic competition; these doctors thus prescribe significantly higher shares of relatively expensive therapies protected by active patents. Interestingly, however, price sensitivity differs across doctor groups: the estimates indicate statistically significant price sensitivity only among family practice physicians.

On the other hand, estimated price sensitivity in a specification that allows spell variation in product qualities may be low precisely because these quality terms already capture most of the

relevant price variation in the data. A key motivation for allowing quality variation across spells was precisely to capture the possibility of unobserved price changes resulting from product entry and patent expiration events. Allowing for these changes, the price parameter ξ in Tables 8 and 9 is identified using only within-spell variation in prices, which is evidently relatively limited. On the other hand, (16) may be estimated directly, without allowing spell variation in product qualities. Accordingly, Tables 10 and 11 provide estimates that hold fixed β_d^T across the sample period. These estimates indicate a highly significant degree of price sensitivity across all three specialties.

Importantly, across Tables 8 through 11, the estimates governing network proximity τ_{ij} and initial beliefs precision S_{i0} are essentially unchanged. Differences in estimated β_d^T and β_{dk}^T parameters may impact the implications, however, and these alternative models are thus considered when investigating the quantitative implications of the model in section 7 below.

6.2 Estimates by U.S. Region

An alternative assumption to the one implicit in our baseline estimation, in which physicians participate in a nationwide social network, is that interactions are substantially more local, with only negligible interactions occurring between agents on opposite coasts, for example. Given the available data, we could estimate the model with nearly any geographic restriction (e.g. by U.S. state, county, or MSA). In this section, we consider a broad division of the United States into four distinct geographic regions and provide separate estimates for each region. These are defined by one-digit zipcodes: New England (0 and 1), the East (2, 3, and 4), Central States (5, 6, and 7), and the West (8 and 9). While the resulting estimates assume no social interactions across these regions, they are otherwise more general as all parameters may differ across regions, including those governing the strength of network connections, the initial precision of beliefs, and product qualities.

Region-specific estimates appear in Table 12 (New England, East) and Table 13 (Central, West). Broadly, the estimates are similar to the baseline results in that they revealing the statistical importance of each network covariate relative to geography (shared medical school, shared specialty, and cohort proximity). It is notable that in the West, the estimated network coefficients are all highest, consistent with a higher relative importance of proximity in school and specialty, compared with geographic proximity. Indeed, the coefficient on graduation year proximity is statistically indistinguishable from that on geographic proximity. The rank-ordering of the network determinants is nevertheless similar across regions: the coefficient on graduation year proximity exceeds the coefficient on shared medical school school, which in turn exceeds that on shared medical specialty. In contrast to the baseline estimates, it is also notable that the initial precision of doctors' initial beliefs S_{i0} is only responsive to local incomes in New England.

Estimated product qualities also differ substantially across the four regions. In New England and the Eastern states, the relative quality estimates for Crestor are high and for Zetia and Vytorin are low, when compared with Central and Western states. Thus, conditional on prescribing an expensive, active-patent drug, high intensity products Lipitor and Crestor are systematically preferred in the East and Northeast. As described in the paragraph above, these effects seem more likely to be explained by differences in disease severity across regions than purely by price; this is

because the same regions that place a high preference on Lipitor and Crestor view Vytorin and Zetia as relatively lower-quality options. Conversely, there is broad agreement regarding the relative quality of simvastatin, which in all regions has the highest estimated quality parameter. By contrast, the quality of generic molecules pravastatin and simvastatin are highest in Eastern states, and both exceed that of Lipitor, consistent with the idea that patients may have higher levels of price sensitivity in these states.

7 Quantitative Implications

In this section, we investigate the quantitative implications of the model in section 2 and corresponding estimates described in section 5. First, we evaluate the extent to which the data are consistent with two main predictions of the model regarding the convergence of agents' beliefs over time. We then simulate the parameterized model and consider the quantitative implications of policy interventions that impact either the network structure or the initial distribution of information across agents. Throughout the section, we focus on the estimates in Table 6, column 1 that include physicians in all specialties and regions.

7.1 Convergence in Agents' Idiosyncratic Beliefs

A straightforward implication of the model is that agents' beliefs regarding the true, unconditional product qualities converge over time to their true values. That is, within each spell k, the mean β_{idt} of individual i's idiosyncratic beliefs about the true, unconditional quality of product d, β^T_{dk}, converges over time to this true value, β^T_{dk}. This occurs in the model because agents are Bayesian and the distribution of signals is unbiased. Thus, as time progresses, agents' beliefs increasingly reflect the mean value of signals received, which itself converges to the true value, and to a decreasing extent reflect prior beliefs about product quality.

To assess whether the data are consistent with this prediction of the model, we proceed in three steps. First, we map the estimated β^T_{dk} values to the prescription shares that they would imply for an agent holding such beliefs. For this, we apply the static choice equation (7) as follows

$$\underbrace{\ln \pi^T_{idtk} - \ln \pi^T_{id'tk}}_{\text{Convergence Target}} = \beta^T_{dk} - \beta^T_{d'k} - \alpha(\sigma^2_d - \sigma^2_{d'})/(2S_{it}). \tag{17}$$

Notice that the expression (17) above implies that two agents with identical (and correct) beliefs about product qualities would nevertheless make different prescription choices in period t due to differences in the precision of their beliefs, S_{it}, in period t. Second, we calculate the Euclidean distance between a) the prescription share vector for a agent i that would be implied by 'correct' beliefs, with elements $\ln \pi^T_{idtk} - \ln \pi^T_{id'tk}$ from (17), and b) agent i's actual prescription vector $\ln \pi_{idt} - \ln \pi_{id't}$ which is observed in the data. Third, we examine the evolution of this Euclidean 'distance to the truth' measure over time by considering how its mean value across agents changes within a spell. An important feature of this setup is that the estimation procedure itself does not place restrictions on the true, relative drug qualities $\beta^T_{dk} - \beta^T_{d'k}$ that would imply convergence; these

quality parameters serve a role similar to that of simple product fixed effects reflecting average, relative prescription shares during an entire spell.

Figure 12 plots the mean value of this Euclidean distance to the truth, for the final spell considered in our analysis. This is a useful starting point because for 18 consecutive quarters, there are no changes in the set of products available, nor their patent status. Consistent with the within-spell convergence prediction of the model, the line in Figure 12 is monotonically decreasing, with a highly significant downward slope of -0.291 (std. error 0.012). Because the prediction should hold in all spells, not only the final spell, Figure 13 replicates the graph for the complete time horizon of our analysis. This latter figure reveals convergence within each spell, with an average estimated slope controlling for spell fixed effects of -0.331 (std. error 0.024).

Notice that Figure 13 also shows large increases in the Euclidean distance measure between the first period of spells three and five relative to the final period of the preceding spells (spells two and four). This likely reflects, in part, an mechanical increase in the distance measure, which itself increases in the length of the vector considered: for example, as a new product is introduced and the set of choices increases from three to four, the Euclidean distance measure increases even if the mean prescription distance for each vector were held constant. To adjust for this, Figures 14 and 15 replicate Figures 12 and 13 but restrict attention to the set of four initially-available products, Lovastatin, Simvastatin, Pravastatin, and the reference drug Lipitor. These restricted figures confirm the pattern of within-spell convergence, with respective downward, within-spell slopes of -0.0601 (std. error 0.0042) and -0.0851 (std. error 0.016), both highly significant. Importantly, Figure 15 is further consistent with the idea that there is aggregate convergence in agents' beliefs not only within but also *across* spells.

7.2 Agents' Convergence Rates, Network Position, and Initial Beliefs Precision

Beyond aggregate convergence, the model predicts heterogeneity in agents' respective rates of convergence to the truth in beliefs about product qualities. Specifically, the Bayesian updating expression (5) implies that when agent i forms posterior beliefs at t, the weight placed on new signals received is strictly increasing in the ratio $\sum_j \tau_{ij} R_{jt}/S_{it+1}$. It is thus simple to observe that, all else equal, agents with higher degrees of network connectedness, as well as agents with less-precise initial beliefs, converge to the truth at a greater rate. To evaluate whether the data are consistent with this prediction regarding heterogeneous rates of convergence, we again consider the measure of agent i's Euclidean 'distance to the truth' in period t. Using this distance measure, as a first step we obtain agent-specific estimates of the average linear rate of convergence as follows,

$$\text{Euclidean Distance to Truth}_{it} = \lambda_i \times t + \eta_i + \eta_t + \mu_{it}. \tag{18}$$

Second, we project the estimated rates of convergence λ_i on the two agent-i characteristics that determine convergence rates in the model: network connectedness $\overline{\tau}_i \equiv \frac{1}{N}\sum_j \tau_{ij}$, and the precision of initial beliefs S_{i0}. The estimates appear in Table 14. Consistent with the model, the estimates reveal a negative and highly significant correlation between the average linear decline in the Euclidean distance to the truth across agents i and their index of network connectedness $\overline{\tau}_i$ which is

constructed using the data and estimates in Table 6. Moreover, the correlation remains significant when including zipcode fixed effects in the regression. This implies that, for two agents located in the same five-digit zipcode, for whom the only variation in network connectedness is due to differences in specialty, school, or cohort, it remains true that relatively well-connected physicians converge to the truth at correspondingly faster rates on average. Column 3 adds the initial beliefs precision, again constructed using the data and estimates in Table 6; the positive and highly significant coefficient is again consistent with the model, suggesting that the beliefs of agents with high levels of initial beliefs precision are also slower to converge to the truth.

Taken together, the results in Figures 12 through 15 and Table 14 are consistent with the model predictions regarding aggregate beliefs convergence to the truth and also the heterogeneity across agents in idiosyncratic convergence rates given variation in agents' network position and the precision of initial beliefs. These support the validity of the model and its assumptions for the data setting we consider.

7.3 Simulating the Model

To further understand the model and its quantitative implications given the estimates described in section 5, we simulate the full model and consider two sets of policy interventions that impact individual and thus also aggregate rates of convergence. The first set of interventions we evaluate are targeted changes to the network structure that change the τ_{ij} values. Such interventions are conceptually related to trade policy (e.g. Eaton Kortum 2002, Donaldson 2018). The second set of interventions changes the distribution of initial beliefs' precision S_{i0} across agents i and is thus loosely related to the notion of information injection (e.g. Banerjee et al 2013, Akbarpour, Malladi, and Saberi 2018). A key question in this literature that our analysis will consider is whether it is more efficient to target well-connected versus peripheral agents and links.

To simulate the model, we begin by specifying agents' prior beliefs β_{id0} to be consistent with their own observed product choice shares in the initial period of the data, given their estimated initial stock of signals S_{i0}. That is, we define

$$\beta_{id0} - \beta_{id'0} = \underbrace{\ln \pi_{id0} - \ln \pi_{id'0}}_{\text{data}} + \underbrace{\alpha(\sigma_d^2 - \sigma_{d'}^2)/(2S_{i0})}_{\text{estimates and data}}.$$

Using the values of $\{\tau_{ij}\}$ and $\{S_{i0}\}$ implied by our data and the parameter estimates of the model, we calculate the complete (deterministic) path of signal stocks S_{it} for each agent i and period $t > 0$. Then, using the Bayesian updating rule and the estimated signal distribution parameters $\{\beta_d^T\}$ and $\{\sigma_d^2\}$, we simulate mean beliefs β_{idt} for all agents i, products d, and time periods $t > 1$. Importantly, although the signal distribution is centered around β_d^T, the fact that agents draw only a finite number of signals per period implies beliefs evolve only gradually over time. Using the simulated data, we recompute the measure of agents' Euclidean distance to the truth in prescription shares that is implied by their beliefs in each period, and replicate the aggregate convergence graph, Figure 12, described above. As in Figure 12, this simulation considers the final spell only, in which the set of available products is fixed.

The simulated version of Figure 12 appears in Figure 16, and shows a similarly decreasing curve with a negative and highly significant slope of -0.502 (std. error 0.030). This matches the aggregate figure qualitatively, and as a moment that is not targeted in our estimation, is reasonably close in magnitude as well. That said, it is worth noting that the model predicts somewhat faster convergence than exists in the data.

7.4 The Influence of Network Structure on Convergence

Having established that the simulated model features aggregate convergence to the truth with a magnitude similar to that in the data, we now proceed to evaluate the quantitative implications of policy interventions. We begin by considering two symmetric adjustments to the structure of the network. The first intervention is to strengthen the weakest network links, and the second intervention is its opposite, to weaken then strongest network links. To implement this, we simulate the model after replacing all first-quartile τ_{ij} values with the 25^{th}-percentile τ_{ij} value, thereby flattening the bottom quartile of the τ_{ij} distribution. Then, we resimulate the model after replacing all fourth-quartile τ_{ij} values with the 75^{th}-percentile τ_{ij} value, flattening the top quartile of the τ_{ij} distribution. Intuitively, the first intervention should speed aggregate convergence by increasing the flow of signals spread over the network in each period, and for the same reason, the second should slow aggregate convergence. To understand the quantitative impact of each intervention, we replicate the aggregate convergence graph for the simulated baseline model in Figure 12 using both counterfactual networks.

The resulting graphs appear in Figures 17 and 18, which show the aggregate convergence curve for the simulation under the counterfactual network (green line) and for comparison, under the true network (black line). It is evident that, despite the qualitative symmetry in the two network adjustments, the latter intervention targeting the strongest links with the highest τ_{ij} values has a substantially larger impact on convergence. Quantitatively, this intervention implies a slope of -0.446 (std. error 0.021), which is statistically different from that of the simulated baseline model -0.502 (std. error 0.030). On the other hand, the intervention targeting the weakest network links has a negligible impact on convergence, as the resulting slope -0.507 (std. error 0.031) is not statistically different from the baseline model. Thus, this simple policy intervention reveals that adjusting the network structure can significantly impact the aggregate rate of beliefs convergence, and that the largest impacts are achieved by targeting the strongest network links.

7.5 Initial Beliefs and the Rate of Convergence

A substantial literature has considered the importance of social networks for channeling knowledge flows across individuals, and thereby determining aggregate rates of information 'infection.' In our model, because agents always share unbiased signals about the true, unconditional drug qualities about which they are learning, interventions that 'seed' the network with information by giving certain agents the true drug qualities would not have a meaningful impact on the spread of knowledge or convergence outcomes. Moreover, such policies would not be feasible as true drug qualities are fundamentally unknown, including to policymakers. Nevertheless, it is possible to intervene in

a way that relates to agents' information by making changes to the distribution of initial signals stocks, S_{i0}. Again, the question we are interested in answering is whether targeting those with high versus low values is more efficient in achieving faster convergence to the truth.

Similar to the network interventions described above, we proceed by considering two symmetric adjustments. The first increases the initial beliefs precision S_{i0} for those with the lowest estimated values, by replacing the first-quarter values with the 25^{th}-percentile S_{i0} value. The second decreases S_{i0} for agents with the highest values, replacing top-quartile values with the 75^{th}-percentile S_{i0} value. We resimulate the model under both changes, where intuitively, the first intervention should slow convergence and the second should increase its pace. To understand the quantitative impact of each intervention, we again replicate the aggregate convergence graph for the simulated baseline model in Figure 12 using both counterfactual distributions of S_{i0}.

Aggregate convergence graphs for each of the two simulations appear in Figures 19 and 20. As described above, the green line is the simulated convergence curve under the counterfactual model, and the black line shows the baseline simulated model, for comparison. While the interventions have the predicted effects on aggregate convergence, it is striking to observe that neither intervention has a quantitatively important impact on convergence. Increasing precision for low-S_{i0} agents decreases the convergence-curve slope from -0.502 (std. error 0.030) to -0.491 (std. error 0.028), a difference that is not statistically different from zero. Similarly, decreasing S_{i0} for agents with the highest values speeds convergence, but the resulting slope is only -0.504 (std. error 0.031), which is again statistically identical to the baseline model.

Without knowing the relative costs of intervening to change S_{i0} or τ_{ij} values, it is of course not possible to undertake a full welfare analysis of which type of intervention is more efficient for stimulating aggregate convergence. Moreover, the analysis above is best viewed as capturing short-run implications; this is because our partial-equilibrium model does not permit agents to determine their network position nor the intensity of their exposure to new signals through active search. Despite these caveats, the simulation results strongly suggest that targeted interventions impacting the structure of the network are substantially more potent than interventions affecting the distribution of information (initial signal stocks) across agents.

8 Conclusion

This paper examines the diffusion of knowledge within a social network. We develop an empirical model capturing general features of learning among professionals that make repeated decisions on behalf of clients. Importantly, these professionals face uncertainty regarding the true, unconditional qualities of the available choices, and it is about these unobserved qualities that agents in our model learn. We show that under the dual assumptions that agents a) make static multinomial choices in each period, and b) update idiosyncratic prior beliefs regarding product qualities in a Bayesian fashion, we obtain a framework that offers a full characterization of not only the evolution of choice efficiency at the agent level, but also at the aggregate level. This latter fact allows us to use the model to consider the implications of micro-level policy interventions for aggregate 'productivity gains' in the form of increased average efficiency of individuals' decisions. Moreover, we show that

estimating the parameters of the model along with standard errors is feasible using a restricted nonlinear least-squares estimator.

To understand the qualitative and quantitative implications of the model, including possible policy interventions that affect the fundamental distribution of knowledge or the network structure itself, we use estimate the model for a context in which we are able to observe the choices of a complete professional network over an 11-year period. In particular, we observe the cholesterol-drug prescription decisions of over 50,000 U.S. physicians at a quarterly frequency during January 2000 through December 2010, a period during which the set of available products and their perceived relative qualities experienced substantial change.

Among the parameters that we estimate are the true, unconditional product qualities that agents learn about in the model. We use these estimates and the individual-level choice share data to assess two of the model's predictions regarding the convergence of agents' beliefs over time to their true values. In addition to predicting that agents' beliefs about product qualities converge to the truth, the model reveals two fundamental determinants of heterogeneity in agents' idiosyncratic rates of convergence: the beliefs of agents with relatively high levels of network connectedness and relatively less-precise initial beliefs should converge to the truth at systematically higher rates. In validation of our model and its assumptions, we find that the data are strongly consistent with aggregate convergence to the truth, and also indicate the statistical importance of both sources of heterogeneity in individual agents' convergence rates.

Finally, our simulation of the model and policy interventions reveal novel insights relevant for policies that aim to facilitate increases in the 'productivity' of agents decisions. These simulations suggest that policies affecting the network structure may be particularly influential, especially when targeting those network links that are already the strongest. By contrast, we find a relatively limited role for interventions targeting weak links or the distribution of information across individual agents.

REFERENCES

[1] **Ackerberg, Daniel A..** 2003. "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination." *International Economic Review*, vol. 44, no. 3, pp. 1007–1040.

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

[3] **Akbarpour, Mohammad, Suraj Malladi, and Amin Saberi.** 2018. "Just a Few Seeds More: Value of Network Information for Diffusion." Working Paper.

[4] **Arrow, Kenneth J., L. Kamran Bilir, and Alan Sorensen.** 2019. "The Impact of Information Technology on the Diffusion of New Pharmaceuticals." *American Economic Journal:*

Applied Economics, forthcoming.

[5] **Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson.** 2013. "The Diffusion of Microfinance." *Science*, vol. 341, no. 6144.

[6] **Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson.** 2014. "Gossip: Identifying Central Individuals in a Social Network." Technical report, National Bureau of Economic Research.

[7] **Bernard, Andrew B and Swati Dhingra.** 2019. "Importers, Exporters and the Division of the Gains from Trade." Working Paper.

[8] **Bernard, Andrew B. and Andreas Moxnes.** (2018). "Networks and Trade." *Annual Review of Economics*, vol. 10, pp. 65–85.

[9] **Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito.** 2019. "Production Networks, Geography, and Firm Performance." *Journal of Political Economy*, vol. 127, no. 2, pp. 639–688.

[10] **Bishop, Tara F., Alex D. Federman, and Salomeh Keyhani.** 2011. "Declines in Physician Acceptance of Medicare and Private Coverage." *Archives of Internal Medicine*, vol. 171, no. 12, pp. 1117–8.

[11] **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.

[12] **Buera, Francisco J. and Ezra Oberfield.** 2019. "The Global Diffusion of Ideas." *Econometrica, forthcoming.*

[13] **Chandreasekhar, Arun and Randall Lewis.** 2016. "Econometrics of Sampled Networks." Stanford University mimeo.

[14] **Chaney, Thomas.** 2014. "The Network Structure of International Trade." *American Economic Review*, vol. 104, no. 11, pp. 3600–3634.

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

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

[17] **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." *The Quarterly Journal of Economics*, vol. 134, no. 1, pp. 51–107.

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

[19] **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.

[20] **Donaldson, Dave.** 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*, vol. 108, no. 4-5, pp. 899–934.

[21] **Eaton, Jonathan, and Samuel Kortum.** 1996. "Trade in Ideas: Patenting and Productivity in the OECD." *Journal of International Economics*, vol. 40, pp. 251–278.

[22] **Eaton, Jonathan, and Samuel Kortum.** 2002. "Technology, Geography, and Trade." *Econometrica*, vol. 70, no. 5, pp. 1741–1779.

[23] **Erdem, Tülin J. and Michael P. Keane.** 1996. "Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." *Marketing Science*, vol. 15, no. 1, pp. 1–20.

[24] **Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi.** 2017. "Estimating Demand for Differentiated Products with Zeroes in Market Share Data." Working Paper.

[25] **Jensen, Robert.** 2007. "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector." *The Quarterly Journal of Economics*, vol. 122, no. 3, pp. 879–924.

[26] **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.

[27] **Ledford, Heidi.** 2013. "Cholesterol Limits Lose Their Lustre." *Nature*, vol. 494, pp. 410–11.

[28] **Degroot, Morris H.** (1970). *Optimal Statistical Decisions.* New York: McGraw-Hill.

[29] **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.

[30] **Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi.** 2017. "Estimating Demand for Differentiated Products with Zeroes in Market Share Data." University of Wisconsin-Madison mimeo.

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

[32] **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.

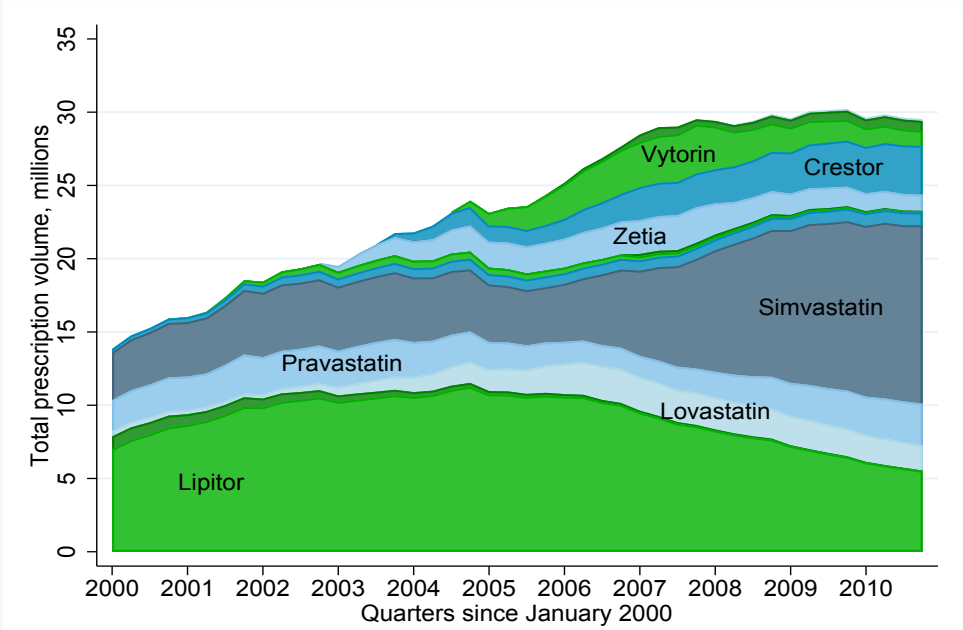
[33] **Manuelli, Rodolfo, and Ananth Seshadri.** 2014. "Frictionless Technology Diffusion: The Case of Tractors." *American Economic Review*, vol. 104, no. 4, pp. 1368–91.

[34] **Skinner, Jonathan.** 2012. "Causes and Consequences of Regional Variations in Health Care." *Handbook of Health Economics*, vol. 2, pp. 45–93.

[35] **Train, Kenneth.** (2009). *Discrete Choice Methods with Simulation.* Cambridge University Press.

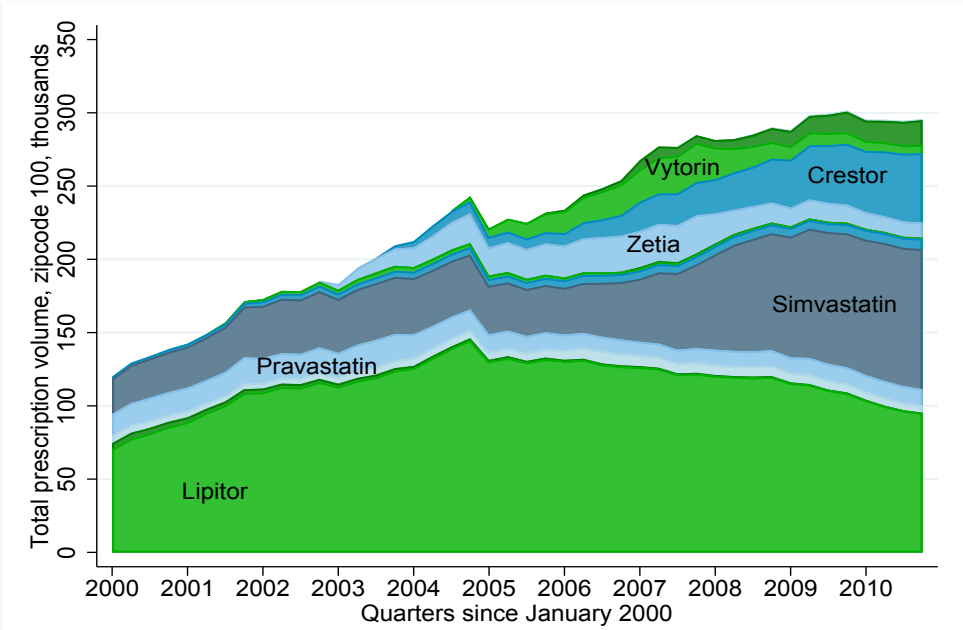
[36] **Wennberg, John E. and Megan M. Cooper, eds.** 1996. *The Dartmouth Atlas of Health Care*, American Hospital Association Press, *at* http://www.dartmouthatlas.org.

Figure 1: Aggregate Evolution in Product Choices, January 2000 to December 2010



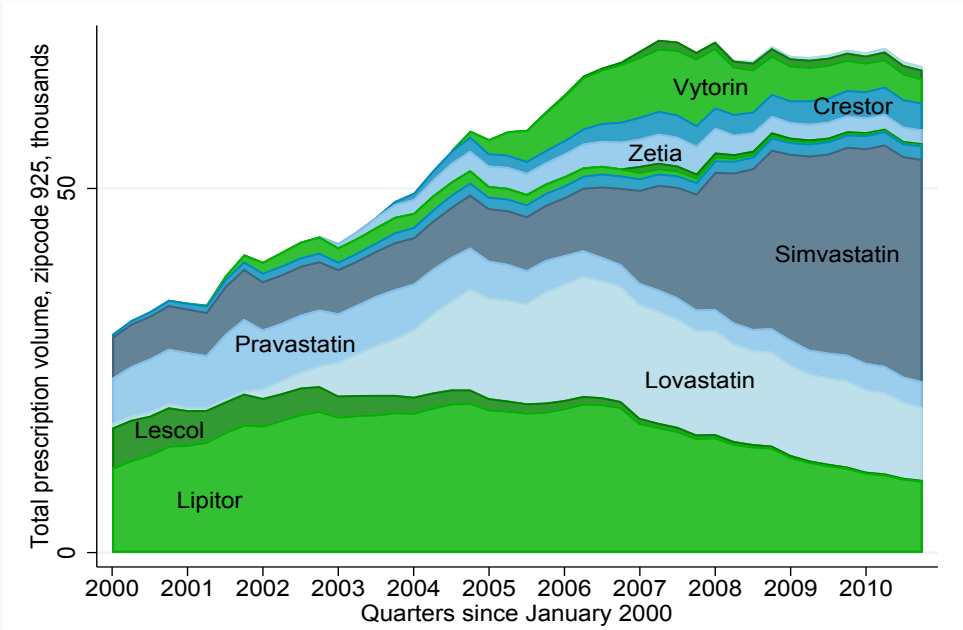
Notes: For each date indicated on the horizontal axis, the height of each shaded region indicates the aggregate prescription volume (number of prescriptions) filled during that period. Labels for specific drugs appear on the region corresponding to that drug.

Figure 2: Evolution in Product Choices in New York, NY, January 2000 to December 2010



Notes: For each date indicated on the horizontal axis, the height of each shaded region indicates the total prescription volume (number of prescriptions) filled in New York, NY (three digit zipcode 100) during that period. Labels for specific drugs appear on the region corresponding to that drug.

Figure 3: Aggregate Evolution in Product Choices in Hemet, CA, January 2000 to December 2010



Notes: For each date indicated on the horizontal axis, the height of each shaded region indicates the total prescription volume (number of prescriptions) filled in Hemet, CA (three digit zipcode 925) during that period. Labels for specific drugs appear on the region corresponding to that drug.

Table 1: Major U.S. Cholesterol Drug Approvals, January 2000–December 2008

Drug Name	Release Date	FDA Category
Lovastatin	December 2001	Generic version
Zetia	October 2002	Molecular entity
Crestor	August 2003	Molecular entity
Vytorin	July 2004	Combination
Lovaza	November 2004	Molecular entity
Pravastatin	April 2006	Generic version
Simvastatin	June 2006	Generic version

Notes: This table lists all major U.S. Cholesterol Drug Introductions during the period January 2000–December 2008. Each product was approved for sale in the United States on the date indicated. As indicated, new drug approvals are categorized by the FDA based on whether the product is a new molecular entity, a new drug combination, a new dosage form, or a new generic equivalent.

Table 2: Evolution in the Within-Molecule Substitution to Generic

	Generic Share in Prescriptions		
	Lovastatin	Pravastatin	Simvastatin
After six months			
Mean	0.8320	0.8240	0.8642
St Dev	0.3364	0.2808	0.2129
5^{th} Percentile	0	0	0.4286
95^{th} Percentile	1	1	1
After 12 months			
Mean	0.9057	0.8482	0.9779
St Dev	0.2606	0.2837	0.0866
5^{th} Percentile	0	0	0.8750
95^{th} Percentile	1	1	1
December 2010			
Mean	0.9996	0.9942	0.9974
St Dev	0.0170	0.0534	0.0271
5^{th} Percentile	1	1	1
95^{th} Percentile	1	1	1

Notes: This table describes within-molecule, generic prescription shares for lovastatin, pravastatin, and simvastatin across U.S. physicians. The top panel describes the distribution of doctor-specific prescription shares for each molecule six months after its patent expiration. The center panel provides the analogous statistics for 12 months after patent expiration, and the bottom panel does the same for the final month observed in our dataset, December 2010. The upper-left number in Panel A (mean, Lovastatin, 0.8320) is the mean, across physicians, fraction of cholesterol drug prescriptions for Lovastatin in May 2002 that are accounted for by generic lovastatin. Below the mean is the standard deviation of this fraction across physicians, followed by 5^{th} and 95^{th} percentile values. Generic approval dates are from the U.S. Food and Drug Administration; prescription data are from IMS Health.

Table 3: Baseline Estimates by Specialty – Cardiology

	Complete Network		Geography Only	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.2083^a	0.0432		
Graduation Year Proximity	0.0644^a	0.0060		
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.2192^a	0.0510	0.0832^a	0.0116
County Population Density	0.7089^a	0.1414	0.9705^a	0.0849
Initial Prescription Volume	49.560^a	6.6972^a	10.763	1.1623
Experience	-879.75^a	104.86	-152.67^a	17.777
Female	-34.032	24.867	-10.348^b	4.7554
Constant	42.6329^a	5.5213	4.6549^a	0.9094
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.7981^a	0.0821	-3.4937^a	0.0430
Pravastatin	-1.4330^a	0.0515	-1.3798^a	0.0274
Simvastatin	-0.6553^a	0.0436	-0.5992^a	0.0234
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-1.9895^a	0.0490	-1.8914^a	0.0388
Pravastatin	-0.7660^a	0.0366	-0.7456^a	0.0314
Simvastatin	1.4576^a	0.0260	1.4793^a	0.0204
Zetia	-1.9566^a	0.0576	-1.7430^a	0.0380
Crestor	0.4068^a	0.0792	0.3286^a	0.0484
Vytorin	-3.1202^a	0.0819	-2.3009^a	0.0507
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	57.870^a	15.4559	3.8734^a	1.4292
Pravastatin	-12.220	9.1205	-2.9075^a	0.8995
Simvastatin	-7.0912	7.8192	-2.0648^a	0.7724
Zetia	210.89^a	26.240	23.720^a	2.3047
Crestor	47.454	31.5914	16.063^a	3.2950
Vytorin	598.15^a	52.853	51.074^a	3.4685
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	7,069			
Number of Observations	1,519,835			

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (11) under the restriction (10) for all cardiologists in the data. Estimates of this baseline model appear in column 1. Column 3 provides estimates of a related model that further restricts the network proximity terms τ_{ij} to depend only on geography. Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 4: Baseline Estimates by Specialty – Internal Medicine

	Complete Network		Geography Only	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.1407^a	0.0187		
Graduation Year Proximity	0.0607^a	0.0030		
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.4754^a	0.0666	0.2362^a	0.0151
County Population Density	0.9073^a	0.1412	2.1766^a	0.1062
Initial Prescription Volume	237.26^a	14.065	44.258^a	2.0618
Experience	-1729.8^a	129.37	-315.66^a	23.710
Female	-36.741^c	20.660	-10.033^b	4.1700
Constant	90.8077^a	7.0067	7.8094^a	1.2672
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.4379^a	0.0465	-3.3739^a	0.0262
Pravastatin	-1.3600^a	0.0311	-1.4139^a	0.0176
Simvastatin	-0.8653^a	0.0279	-0.8883^a	0.0159
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-1.2068^a	0.0204	-1.2046^a	0.0162
Pravastatin	-0.4695^a	0.0172	-0.4844^a	0.0152
Simvastatin	1.7248^a	0.0125	1.7281^a	0.0101
Zetia	-2.2788^a	0.0312	-1.9769^a	0.0206
Crestor	-0.0601^a	0.0397	-0.1174^a	0.0244
Vytorin	-3.3133^a	0.0430	-2.3094^a	0.0269
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	43.060^a	17.735	8.8551^a	1.7601
Pravastatin	-86.869^a	12.286	-10.246^a	1.2281
Simvastatin	-58.332^a	10.941	-6.3407^a	1.1036
Zetia	683.28^a	41.797	61.738^a	3.3594
Crestor	24.473	45.371	20.508^a	4.5619
Vytorin	1963.7^a	87.216	145.37^a	5.0373
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	21,653			
Number of Observations	4,655,395			

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (11) under the restriction (10) for all internists in the data. Estimates of this baseline model appear in column 1. Column 3 provides estimates of a related model that further restricts the network proximity terms τ_{ij} to depend only on geography. Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 5: Baseline Estimates by Specialty – Family Practice

	Complete Network		Geography Only	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.0474^a	0.0049		
Graduation Year Proximity	0.0184^a	0.0008		
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.0854^a	0.0202	0.0243^a	0.0082
County Population Density	0.7110^a	0.1615	1.5493^a	0.1297
Initial Prescription Volume	83.761^a	4.2689	36.408^a	1.4121
Experience	-395.12^a	34.686	-103.02^a	12.0102
Female	-0.7985	5.4909	1.6414^a	2.1411
Constant	21.8084^a	1.8313	5.9075^a	0.6250
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.5033^a	0.0595	-3.3663^a	0.0364
Pravastatin	-1.0522^a	0.0376	-1.1997^a	0.0232
Simvastatin	-0.3812^a	0.0334	-0.5239^a	0.0205
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-0.9009^a	0.0214	-0.9009^a	0.0174
Pravastatin	-0.0876^b	0.0186	-0.0876^a	0.0169
Simvastatin	1.9999^a	0.0131	1.9999^a	0.0111
Zetia	-1.6286^a	0.0356	-1.6286^a	0.0244
Crestor	-0.0509	0.0430	-0.0509^c	0.0278
Vytorin	-1.9824^a	0.0471	-1.9824^a	0.0318
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	25.685^a	6.1975	5.8098^a	1.3277
Pravastatin	-59.1872^a	4.3634	-15.9954^a	0.8898
Simvastatin	-69.5182^a	4.1762	-19.8590^a	0.8040
Zetia	155.2922^a	13.3383	21.5680^a	2.7473
Crestor	-19.5912	16.2344	-5.0965	3.5122
Vytorin	720.4631^a	26.4837	92.8266^a	4.1204
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	18,302			
Number of Observations	3,934,930			

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (11) under the restriction (10) for all family practice physicians in the data. Estimates of this baseline model appear in column 1. Column 3 provides estimates of a related model that further restricts the network proximity terms τ_{ij} to depend only on geography. Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 6: Baseline Estimates, All Specialties

	$\delta < 1$		$\delta = 1$	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.0978^a	0.0426	0.1835^a	0.2899
Same Medical Specialty Indicator	0.0107^a	0.0021	0.0273^a	0.1054
Graduation Year Proximity	0.0244^a	0.0072	0.0171^a	0.0521
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.3559^a	0.0360	0.3788^a	0.0543
County Population Density	0.2654^a	0.0740	0.2903^a	0.0871
Initial Prescription Volume	50.547^a	5.7961	51.887^a	9.4723
Experience	28.166	45.005	20.709	51.943
Female	-2.4207	7.3330	-2.1054	8.8955
Internal Medicine Indicator	38.047^a	10.810	112.28^a	32.999
Cardiology Indicator	-41.457^a	9.1189	-36.435^a	11.152
Family Practice Indicator	24.012^b	9.7928	65.638^a	22.090
Constant	-12.486^a	2.0309	-13.9828^a	2.3325
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.4459^a	0.0962	-3.2655^a	-3.2655
Pravastatin	-1.3426^a	0.0621	-1.5550^a	-1.5550
Simvastatin	-0.7239^a	0.0576	-0.9762^a	-0.9762
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-1.1530^a	0.0416	-0.9963^a	0.0354
Pravastatin	-0.3226^a	0.0355	-0.1812^a	0.0355
Simvastatin	1.7918^a	0.0258	2.0130^a	0.0228
Zetia	-2.0715^a	0.0644	-1.5405^a	0.0338
Crestor	-0.0458	0.0801	0.2968^a	0.0343
Vytorin	-3.0325^a	0.0927	-1.4831^a	0.0339
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	13.193^a	6.3092		
Pravastatin	-13.701^a	4.2851		
Simvastatin	-16.567^a	4.1220		
Zetia	107.60^a	15.593		
Crestor	23.619	16.336		
Vytorin	321.19^a	34.886		
Discount Factor, δ	0.9873	Fixed	1	Fixed
Risk Aversion Parameter, α	0.9900	Fixed		
Number of Doctors	5,304			
Number of Observations	1,140,360			

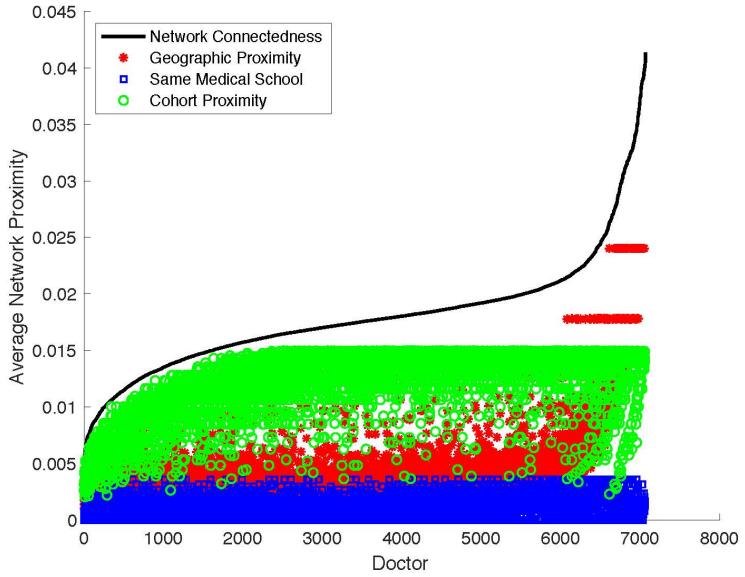
Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (11) under the restriction (10) for a ten percent random sample of all physicians in the data. Estimates of this baseline model appear in column 1. Column 3 provides estimates of a related model that further restricts the discount factor $\delta = 1$ implying no time discounting. Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 7: Model Estimates with Restricted Network Specification – Distance Only

	$\delta < 1$		$\delta = 1$	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.0759^a	0.0062	0.0699^a	0.0055
County Population Density	0.4035^a	0.0481	0.4128^a	0.0469
Initial Prescription Volume	13.281^a	0.9393	12.139^a	0.8552
Experience	-12.483	10.289	-11.212	9.5379
Female	0.1436	1.8131	0.3724	1.6792
Internal Medicine Indicator	2.8213	2.4431	2.5150	2.2538
Cardiology Indicator	-9.3386^a	2.4251	-9.1982^a	2.1704
Family Practice Indicator	-0.1344	2.3825	-0.0643	2.2138
Constant	-1.8305^a	0.4721	-1.6941^a	0.4212
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.4235^a	0.0558	-3.2017^a	0.0323
Pravastatin	-1.3458^a	0.0357	-1.4920^a	0.0207
Simvastatin	-0.7122^a	0.0338	-0.9136^a	0.0193
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-1.1688^a	0.0348	-0.9936^a	0.0357
Pravastatin	-0.3233^a	0.0323	-0.1686^a	0.0358
Simvastatin	1.8026^a	0.0219	2.0369^a	0.0228
Zetia	-1.7899^a	0.0439	-1.5334^a	0.0341
Crestor	0.0055	0.0514	0.3194^a	0.0345
Vytorin	-2.1117^a	0.0575	-1.4847^a	0.0342
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	3.9703^a	0.8279		
Pravastatin	-2.8304^a	0.5595		
Simvastatin	-3.6635^a	0.5284		
Zetia	11.610^a	1.6477		
Crestor	3.4806	2.1602		
Vytorin	30.340^a	2.4694		
Discount Factor, δ	0.9873	Fixed	1	Fixed
Risk Aversion Parameter, α	0.9900	Fixed		
Number of Doctors	5,304			
Number of Observations	1,140,360			

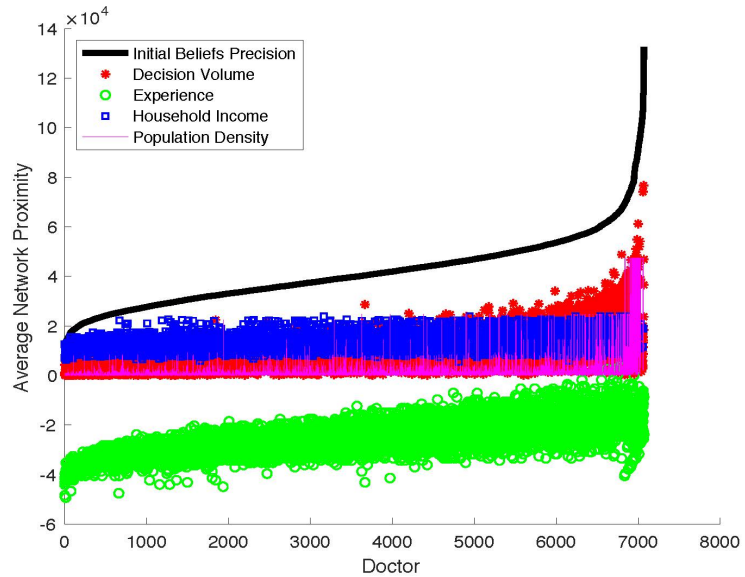
Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (11) under the restriction (10) for a ten percent random sample of all physicians in the data. Estimates of this baseline model, under the additional restriction that network proximity τ_{ij} depends only on geography, appear in column 1. Column 3 provides estimates of a related model that further restricts the discount factor $\delta = 1$ implying no time discounting. Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Figure 4: Decomposition of Network Connectedness by Component, Cardiology



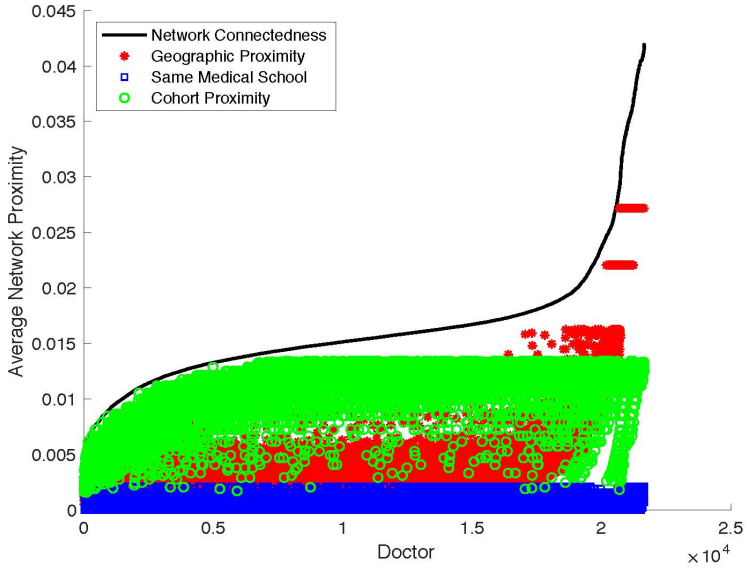
Notes: This figure provides the predicted value of mean network connectedness $\overline{\tau}_{ij}$ for each cardiologist i given the estimates in Table 3 and the data (solid black line), as well as a decomposition of this curve into its three underlying components: geographic proximity (red solid circles), shared medical school (blue outlined squares), and cohort proximity (green outlined circles).

Figure 5: Decomposition of Initial Beliefs Precision, Cardiology



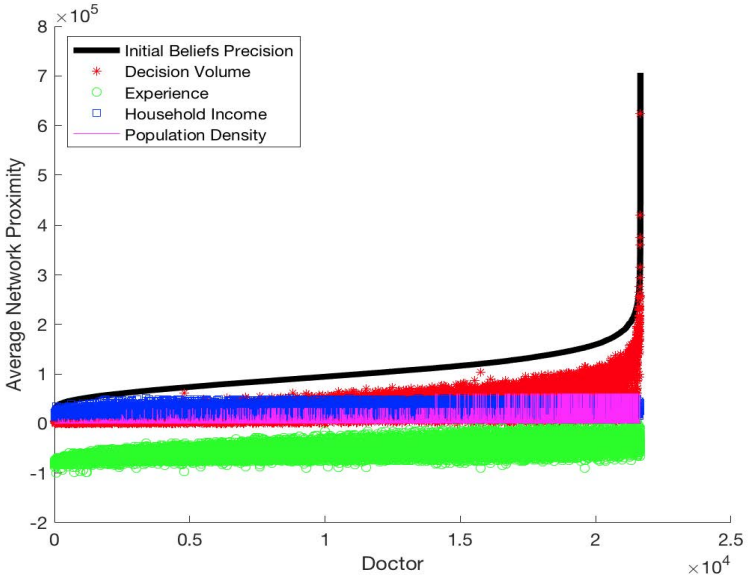
Notes: This figure provides the predicted value of initial beliefs precision S_{i0} for each cardiologist i given the estimates in Table 3 and the data (solid black line), as well as a decomposition of this curve into its four continuous components: decision volume (red solid circles), years of experience (green outlined circles), local household income (blue outlined squares), and local population density (dotted magenta line).

Figure 6: Decomposition of Network Connectedness by Component, Internal Medicine



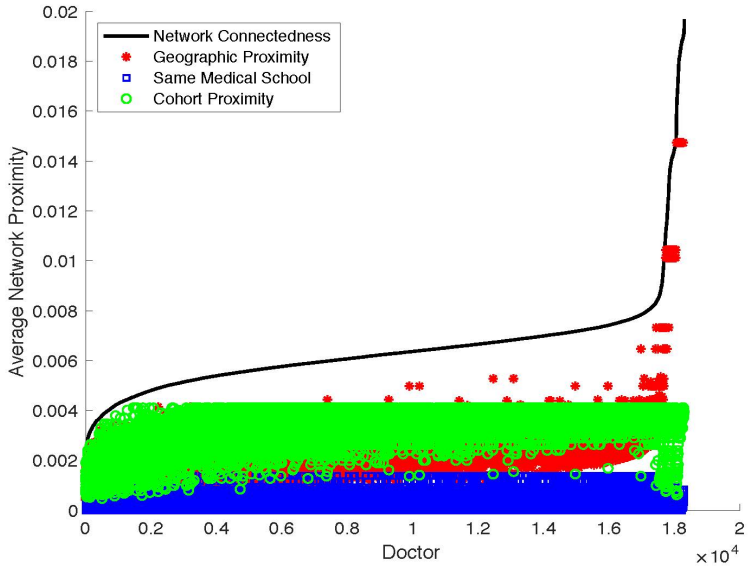
Notes: This figure provides the predicted value of mean network connectedness $\overline{\tau}_{ij}$ for each internist i given the estimates in Table 4 and the data (solid black line), as well as a decomposition of this curve into its three underlying components: geographic proximity (red solid circles), shared medical school (blue outlined squares), and cohort proximity (green outlined circles).

Figure 7: Decomposition of Initial Beliefs Precision, Internal Medicine



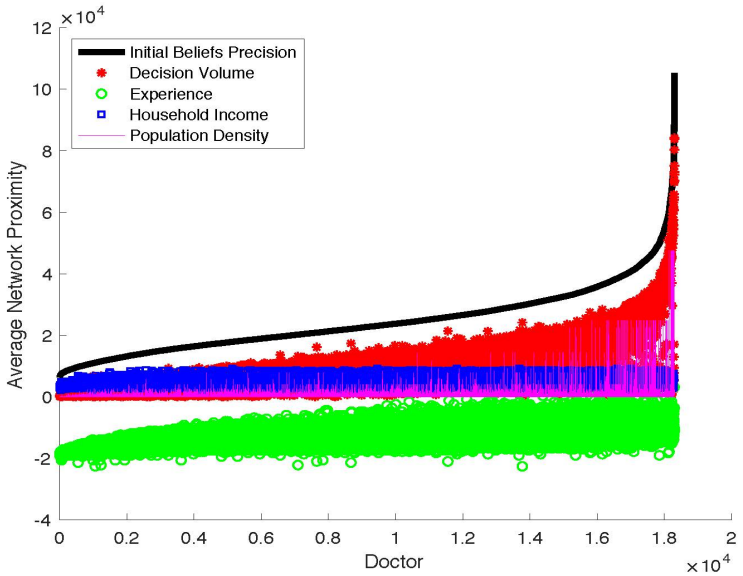
Notes: This figure provides the predicted value of initial beliefs precision S_{i0} for each internist i given the estimates in Table 4 and the data (solid black line), as well as a decomposition of this curve into its four continuous components: decision volume (red solid circles), years of experience (green outlined circles), local household income (blue outlined squares), and local population density (dotted magenta line).

Figure 8: Decomposition of Network Connectedness by Component, Family Practice



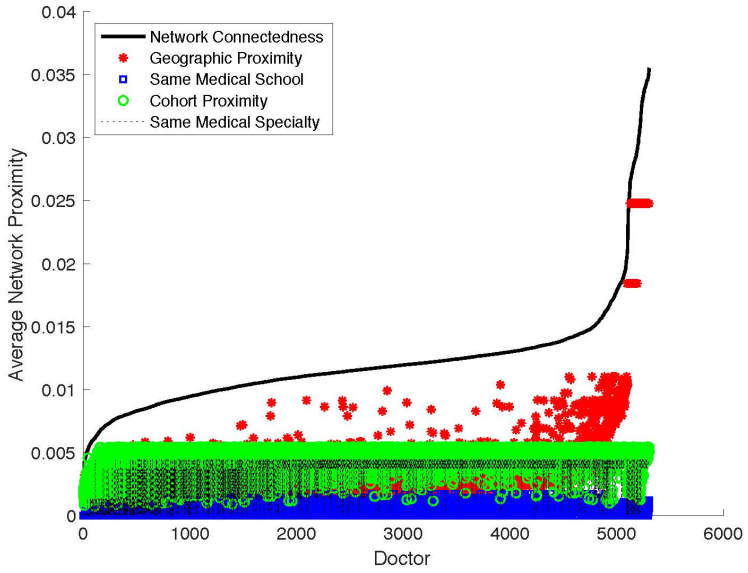
Notes: This figure provides the predicted value of mean network connectedness $\overline{\tau}_{ij}$ for each family practice physician i given the estimates in Table 5 and the data (solid black line), as well as a decomposition of this curve into its three underlying components: geographic proximity (red solid circles), shared medical school (blue outlined squares), and cohort proximity (green outlined circles).

Figure 9: Decomposition of Initial Beliefs Precision, Family Practice



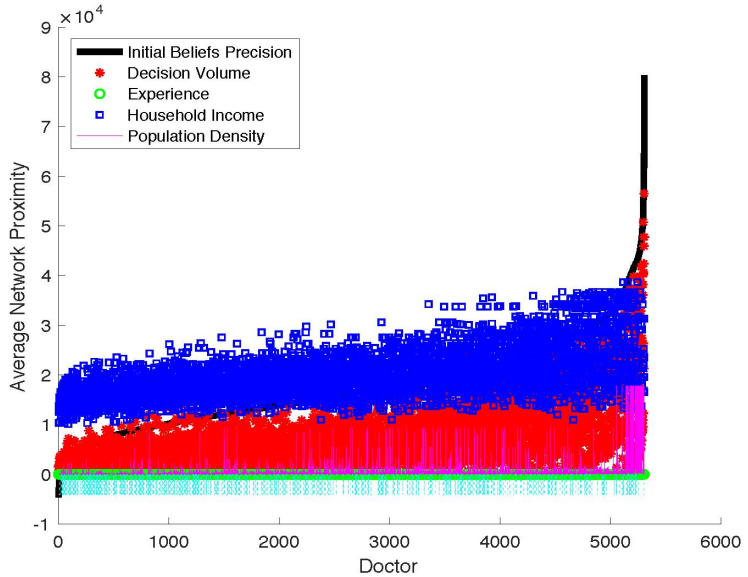
Notes: This figure provides the predicted value of initial beliefs precision S_{i0} for each family practice physician i given the estimates in Table 5 and the data (solid black line), as well as a decomposition of this curve into its four continuous components: decision volume (red solid circles), years of experience (green outlined circles), local household income (blue outlined squares), and local population density (dotted magenta line).

Figure 10: Decomposition of Network Connectedness by Component, All Specialties



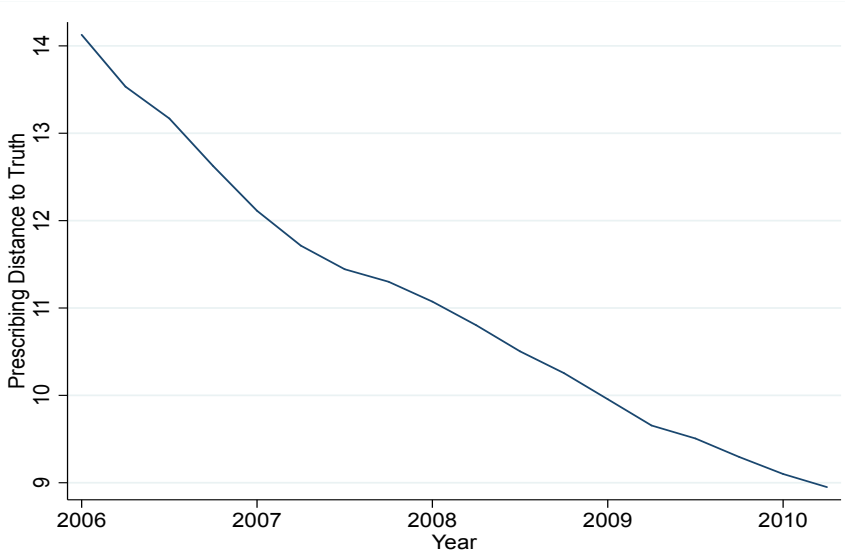
Notes: This figure provides the predicted value of mean network connectedness $\overline{\tau}_{ij}$ for each physician i in a ten percent random sample of all physicians, based on the estimates in Table 6 and the data (solid black line). The four separate components of this curve are also shown, including geographic proximity (red solid circles), shared medical school (blue outlined squares), cohort proximity (green outlined circles), and shared specialty (dotted black line).

Figure 11: Decomposition of Initial Beliefs Precision



Notes: This figure provides the predicted value of initial beliefs precision S_{i0} for each physician i in a ten percent random sample of all physicians, based on the estimates in Table 6 and the data (solid black line). The four separate components of this curve are also shown, including geographic proximity (red solid circles), shared medical school (blue outlined squares), cohort proximity (green outlined circles), and shared specialty (dotted black line).

Figure 12: Mean Euclidean Distance to the Truth, Final Spell

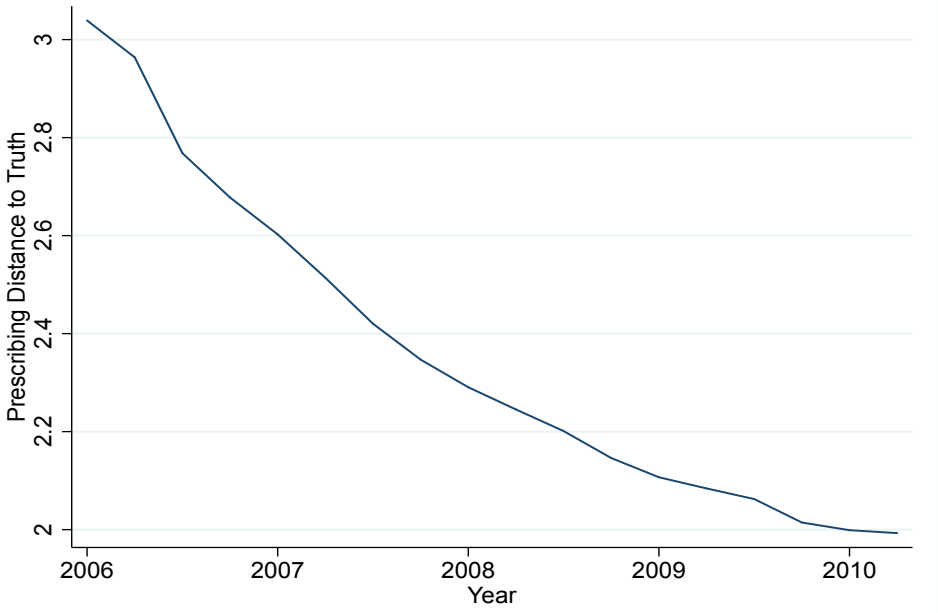


Notes: This figure plots the average (across doctors i) Euclidean distance between a) the vector of relative prescription shares $\ln \pi^T_{idtk} - \ln \pi^T_{id'tk}$ for doctor i that, as defined in (12), would be implied by doctor i choosing products d according to the true, unconditional product qualities β^T_{dk}, and b) the actual relative prescription shares $\ln \pi_{idt} - \ln \pi_{id't}$ of doctor i for each quarter t in the final spell.

Figure 13: Mean Euclidean Distance to the Truth, All Spells

Notes: This figure plots the average (across doctors i) Euclidean distance between a) the vector of relative prescription shares $\ln \pi^T_{idtk} - \ln \pi^T_{id'tk}$ for doctor i that, as defined in (12), would be implied by doctor i choosing products d according to the true, unconditional product qualities β^T_{dk}, and b) the actual relative prescription shares $\ln \pi_{idt} - \ln \pi_{id't}$ of doctor i for each quarter t during the sample period (all spells). Vertical lines demarcate the six spells.

Figure 14: Mean Euclidean Distance to the Truth, Initially-Available Products Only, Final Spell



Notes: This figure plots the average (across doctors i) Euclidean distance between a) the vector of relative prescription shares $\ln \pi^T_{idtk} - \ln \pi^T_{id'tk}$ for doctor i that, as defined in (12), would be implied by doctor i choosing products d according to the true, unconditional product qualities β^T_{dk}, and b) the actual relative prescription shares $\ln \pi_{idt} - \ln \pi_{id't}$ of doctor i for each quarter t in the final spell. To distinguish the convergence mechanism from mechanical changes in the scale of the Euclidean distance measure resulting from product entry, only the four initially-available products are included.

Table 8: Advertising and Price Effects

	Cardiology		Internal Medicine	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.1929^a	0.0417	0.1520^a	0.0205
Graduation Year Proximity	0.0625^a	0.0059	0.0636^a	0.0033
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.2065^a	0.0492	0.4252^a	0.0651
County Population Density	0.7178^a	0.1436	0.9205^a	0.1445
Initial Prescription Volume	46.241^a	6.3777	223.97^a	14.032
Experience	-820.08^a	101.66	-1651.1^a	129.63
Female	-33.279	24.567	-49.008^b	20.503
Constant	40.414^a	5.3486	87.184^a	7.0150
Product Choice Equation				
Advertising, γ	-0.0001^b	0.0001	-0.0002^a	0.0001
Drug Price, ξ	-0.0119	0.0168	0.0002	0.0098
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.7880^a	0.0835	-3.4649^a	0.0485
Pravastatin	-1.4140^a	0.0523	-1.3133^a	0.0321
Simvastatin	-0.6364^a	0.0443	-0.8448^a	0.0289
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-1.9703^a	0.0498	-1.2262^a	0.0213
Pravastatin	-0.7639^a	0.0373	-0.4735^a	0.0180
Simvastatin	1.4579^a	0.0266	1.7018^a	0.0132
Zetia	-1.9744^a	0.0584	-2.2696^a	0.0326
Crestor	0.4114^a	0.0801	-0.0340	0.0412
Vytorin	-3.1198^a	0.0832	-3.2648^a	0.0445
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	-50.604^a	14.774	-48.748^a	17.697
Pravastatin	13.323^a	8.8403	87.578^a	12.241
Simvastatin	8.2084	7.5706	55.853^a	10.853
Zetia	-209.96^a	25.675	-646.85^a	42.083
Crestor	-43.669	30.444	-12.177	45.030
Vytorin	-573.26^a	51.121	-1851.2^a	88.036
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	7,069		21,653	
Number of Observations	1,519,835		4,655,395	

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (16) under the restriction (10) for all cardiologists (column 1) and internists (column 3). Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 9: Advertising and Price Effects

	Family Practice		All Specialties	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.0519^a	0.0056	0.1305^a	0.0306
Graduation Year Proximity	0.0195^a	0.0009	0.0310^a	0.0036
Same Medical Specialty Indicator			0.0088^a	0.0011
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.0764^a	0.0202	0.0816^a	0.0234
County Population Density	0.6002^a	0.1536	0.0969^c	0.0516
Initial Prescription Volume	80.662^a	4.3291	38.100^a	4.5493
Experience	-382.37^a	35.485	-187.01^a	38.418
Female	-0.4409	5.6620	-4.2342	7.3666
Constant	21.381^a	1.8725	8.9733^a	2.0912
Product Choice Equation				
Advertising, γ	-0.0002^b	0.0001	-0.0002^a	0.0001
Drug Price, ξ	-0.0377	0.0116	0.0005^a	0.0208
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.4996^a	0.0623	-3.7049^a	0.1114
Pravastatin	-1.0091^a	0.0392	-1.2908^a	0.0725
Simvastatin	-0.3666^a	0.0347	-0.7011^a	0.0666
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-0.9439^a	0.0225	-1.2932^a	0.0465
Pravastatin	-0.0554^a	0.0195	-0.3616^a	0.0386
Simvastatin	1.9943^a	0.0139	1.7541^a	0.0288
Zetia	-1.8776^a	0.0369	-2.1133^a	0.0695
Crestor	0.0065	0.0445	-0.0550	0.0895
Vytorin	-3.1359^a	0.0484	-3.2065^a	0.0972
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	-22.044^a	6.2660	-24.635^a	7.4587
Pravastatin	58.936^a	4.4545	18.294^a	4.8564
Simvastatin	66.377^a	4.2185	15.971^a	4.5281
Zetia	-149.83^a	13.381	-107.62^a	15.189
Crestor	25.269	16.271	-13.823	17.851
Vytorin	-682.81^a	26.631	-347.92^a	32.226
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	18,302		5,304	
Number of Observations	3,934,930		1,140,360	

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (16) under the restriction (10) for all family practice physicians (column 1) and a ten percent random sample of all physicians (column 3). Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 10: Advertising and Price Effects with Fixed Product Qualities

	Cardiology		Internal Medicine	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.2109^a	0.0144	0.2983^a	0.0112
Graduation Year Proximity	0.0614^a	0.0020	0.0876^a	0.0019
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.1504^a	0.0495	0.4196^a	0.0949
County Population Density	0.4894^a	0.1305	0.3465^b	0.1667
Initial Prescription Volume	53.367^a	5.9840	340.12^a	16.778
Experience	-650.64^a	81.109	-2592.7^a	157.10
Female	-50.878^b	24.442	-61.782^b	29.676
Constant	37.633^a	4.3477	148.74^a	8.8604
Product Choice Equation				
Advertising, γ	0.0002^a	0.0001	0.0001^a	0.0000
Drug Price, ξ	-0.1596^a	0.0165	-0.1690^a	0.0097
Relative Mean by Drug, $\beta_d^T - \beta_{d'}^T$				
Lovastatin	-1.5501^a	0.0478	-0.4948^a	0.0207
Pravastatin	-1.4244^a	0.0342	-1.1252^a	0.0174
Simvastatin	0.4932^a	0.0274	1.0213^a	0.0141
Zetia	-3.2765^a	0.0584	-3.3073^a	0.0310
Crestor	-0.1601^b	0.0760	-0.6983^a	0.0379
Vytorin	-5.0508^a	0.0841	-4.9381^a	0.0435
Relative Variance by Drug, $\sigma_d^2 - \sigma_{d'}^2$				
Lovastatin	232.73^a	11.8830	1083.4^a	24.361
Pravastatin	26.745^a	6.7570	287.19^a	12.326
Simvastatin	184.54^a	7.6471	963.53^a	19.766
Zetia	-866.32^a	28.4163	-2575.0^a	55.887
Crestor	-259.56^a	26.8515	-1024.2^a	51.585
Vytorin	-1776.3^a	53.8500	-6031.4^a	118.602
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	7,069		21,653	
Number of Observations	1,519,835		4,655,395	

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (16) under the restriction (10) for all cardiologists (column 1) and internists (column 3). Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 11: Advertising and Price Effects with Fixed Product Qualities

	Family Practice		All Specialties	
	Estimate (1)	Standard Error (2)	Estimate (3)	Standard Error (4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.0588^a	0.0026	0.1916^a	0.0156
Graduation Year Proximity	0.0247^a	0.0005	0.0503^a	0.0023
Same Medical Specialty Indicator			0.0116^a	0.0006
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.1788^a	0.0284	0.1460^a	0.0369
County Population Density	0.4480^b	0.1726	0.0328	0.0647
Initial Prescription Volume	89.126	4.8958	55.315	5.8358
Experience	-420.96^a	40.587	-264.98^a	54.090
Female	-2.7455	7.7817	-8.1169	11.367
Constant	24.646^a	2.2725	29.042^a	3.7456
Product Choice Equation				
Advertising, γ	0.0002^b	0.0001	0.0000	0.0001
Drug Price, ξ	-0.2416^a	0.0115	-0.1670^a	0.0206
Relative Mean by Drug, $\beta_d^T - \beta_{d'}^T$				
Lovastatin	-0.0646^a	0.0206	-0.3441^a	0.0438
Pravastatin	-0.7493^a	0.0185	-0.9883^a	0.0374
Simvastatin	1.3596^a	0.0145	1.2217^a	0.0297
Zetia	-2.7842^a	0.0318	-3.3442^a	0.0647
Crestor	-0.9123^a	0.0376	-0.9130^a	0.0798
Vytorin	-4.9029^a	0.0440	-5.2804^a	0.0924
Relative Variance by Drug, $\sigma_d^2 - \sigma_{d'}^2$				
Lovastatin	312.19^a	5.8829	222.14^a	9.7413
Pravastatin	124.71^a	3.5107	68.684^a	5.0493
Simvastatin	289.16^a	4.7651	199.03^a	8.0205
Zetia	-549.16^a	12.067	-503.96^a	21.8208
Crestor	-371.79^a	14.377	-256.66^a	21.1777
Vytorin	-1766.1^a	28.474	-1250.6^a	48.2705
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	18,302		5,304	
Number of Observations	3,934,930		1,140,360	

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (16) under the restriction (10) for all family practice physicians (column 1) and a ten percent random sample of all physicians (column 3). Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 12: Baseline Estimates by U.S. Region

	Region 1 - New England		Region 2 - East	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.0670^a	0.0237	0.0136^a	0.0028
Same Medical Specialty Indicator	0.0215^a	0.0019	0.0054^a	0.0002
Graduation Year Proximity	0.0803^a	0.0065	0.0179^a	0.0008
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.3792^a	0.0802	0.0000	0.0438
County Population Density	0.8071^a	0.1170	4.6238^a	0.5135
Initial Prescription Volume	259.36^a	20.572	114.22^a	6.2029
Experience	-1422.8^a	188.39	-788.47^a	67.498
Female	-69.370^b	32.987	-23.628^c	12.507
Internal Medicine Indicator	290.69^a	45.769^a	97.299	17.281
Cardiology Indicator	57.083	48.760	-0.0596	20.1431
Family Practice Indicator	157.80^a	43.148	32.435^b	16.3352
Constant	70.064^a	9.7256	52.135^a	3.6631
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.4932^a	0.0667	-3.6234^a	0.0708
Pravastatin	-1.5828^a	0.0447	-1.2632^a	0.0458
Simvastatin	-0.7060^a	0.0386	-0.1296^a	0.0382
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-1.7735^a	0.0331	-1.1697^a	0.0274
Pravastatin	-0.8954^a	0.0263	0.0386^a	0.0212
Simvastatin	1.7515^a	0.0185	1.9473^a	0.0156
Zetia	-2.4872^a	0.0453	-2.4907^a	0.0450
Crestor	0.1936^a	0.0595	0.2850^a	0.0555
Vytorin	-3.3799^a	0.0688	-4.4467^a	0.0611
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	39.642^a	26.431	83.740^a	16.8502
Pravastatin	-60.388^a	17.717	-70.588^c	11.0803
Simvastatin	-157.50^a	17.672	-151.74^a	9.7644
Zetia	785.90^a	61.5746	771.23^a	34.6409
Crestor	-132.93^a	65.673	-89.390^a	41.0030
Vytorin	1662.7^a	113.74	2205.3^a	59.528
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	10,595		19,615	
Number of Observations	2,277,925		4,217,225	

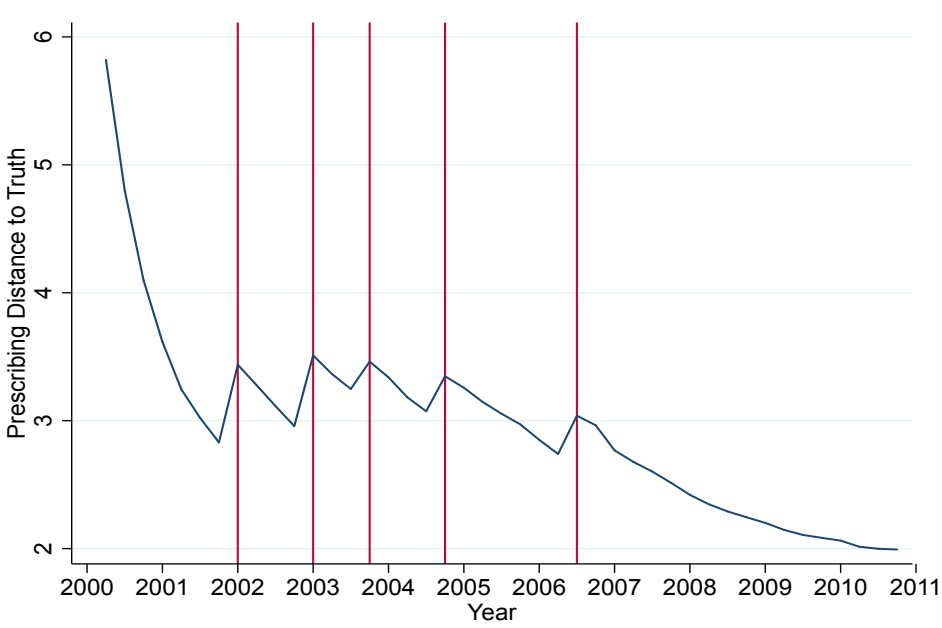
Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (11) under the restriction (10) for all physicians located in the U.S. Northeast (columns 1 and 2, zipcodes begin with 0 or 1) or the U.S. East (columns 3 and 4, zipcodes begin with 3 or 4). Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Table 13: Baseline Estimates by U.S. Region

	Region 3 - Central		Region 4 - West	
	Estimate	Standard Error	Estimate	Standard Error
	(1)	(2)	(3)	(4)
Bilateral Proximity, $\tau_{ij} = \tau(\boldsymbol{b}_\tau, \boldsymbol{Y}_{ij})$				
Geographic Proximity	1	Fixed	1	Fixed
Same Medical School Indicator	0.2806^a	0.0484	0.8460^a	0.2660
Same Medical Specialty Indicator	0.0791^a	0.0079	0.2822^a	0.0321
Graduation Year Proximity	0.3041^a	0.0289	0.9734^a	0.1060
Initial Precision, $S_{i0} = S(\boldsymbol{b}_S, \boldsymbol{X}_{i0})$				
County Mean Household Income	0.0406	0.3249	-0.3211	0.6149
County Population Density	22.558^a	2.8202	16.903^a	4.7964
Initial Prescription Volume	689.35^a	71.837	1899.6^a	216.96
Experience	-4076.8^a	529.11	-6000.4^a	1002.9
Female	-96.447	89.162	-286.20	178.79
Internal Medicine Indicator	777.14^a	132.24	1759.2^a	290.10
Cardiology Indicator	-80.964	122.70	210.56	257.07
Family Practice Indicator	586.44^a	115.22	1047.3^a	239.48
Constant	244.51^a	31.0297	433.68^a	65.4862
Distribution of Signal Values				
Relative Mean by Drug, First Spell, $\beta^T_{d1} - \beta^T_{d'1}$				
Lovastatin	-3.3633^a	0.0667	-3.2939^a	0.0670
Pravastatin	-1.5124^a	0.0454	-1.0155^a	0.0443
Simvastatin	-0.7300^a	0.0372	-1.1991^a	0.0428
Relative Mean by Drug, Final Spell, $\beta^T_{d6} - \beta^T_{d'6}$				
Lovastatin	-1.2066^a	0.0278	-0.6905^a	0.0244
Pravastatin	-0.3028^a	0.0244	-0.7918^a	0.0238
Simvastatin	1.7172^a	0.0167	1.7903^a	0.0179
Zetia	-1.8615^a	0.0396	-2.0765^a	0.0410
Crestor	-0.0058	0.0471	-0.4066^a	0.0521
Vytorin	-2.7809^a	0.0552	-2.6047^a	0.0574
Relative Variance by Drug, $\sigma^2_d - \sigma^2_{d'}$				
Lovastatin	142.29^c	83.721	-218.97	154.08
Pravastatin	-196.49^a	58.485	-857.68^a	134.26
Simvastatin	-340.11^a	54.268	-356.23^a	105.25
Zetia	1769.0^a	202.40	2903.4^a	419.30
Crestor	142.09	183.90	620.50	425.20
Vytorin	5833.5^a	512.50	10332^a	1127.5
Discount Factor, δ	0.9873	Fixed	0.9873	Fixed
Risk Aversion Parameter, α	0.9900	Fixed	0.9900	Fixed
Number of Doctors	12,815		10,015	
Number of Observations	2,755,225		2,153,225	

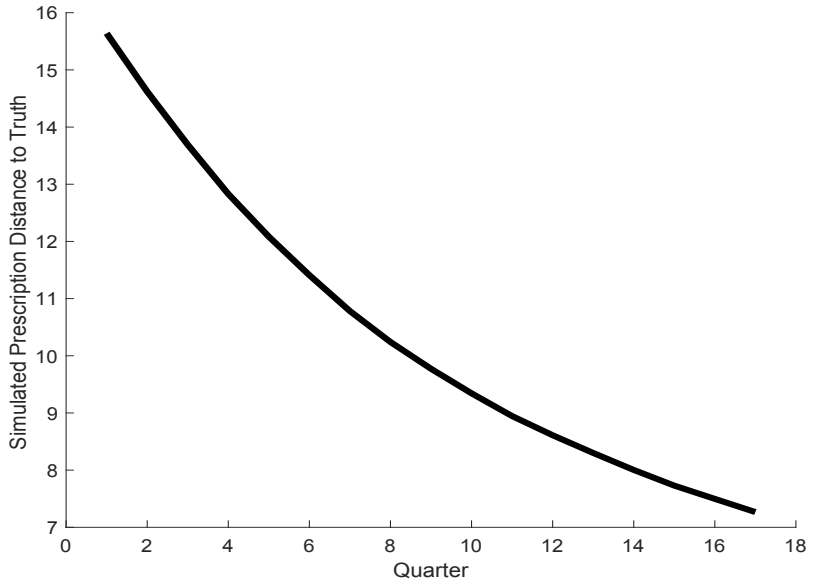
Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table provides nonlinear least-squares estimates of equation (11) under the restriction (10) for all physicians located in the U.S. Central States (columns 1 and 2, zipcodes begin with 5, 6, or 7) or the U.S. West (columns 3 and 4, zipcodes begin with 8 or 9). Standard errors are shown to the right of each point estimate. The reference drug is Lipitor and the risk aversion parameter α is from Crawford and Shum (2005).

Figure 15: Mean Euclidean Distance to the Truth, Initially-Available Products Only, All Spells



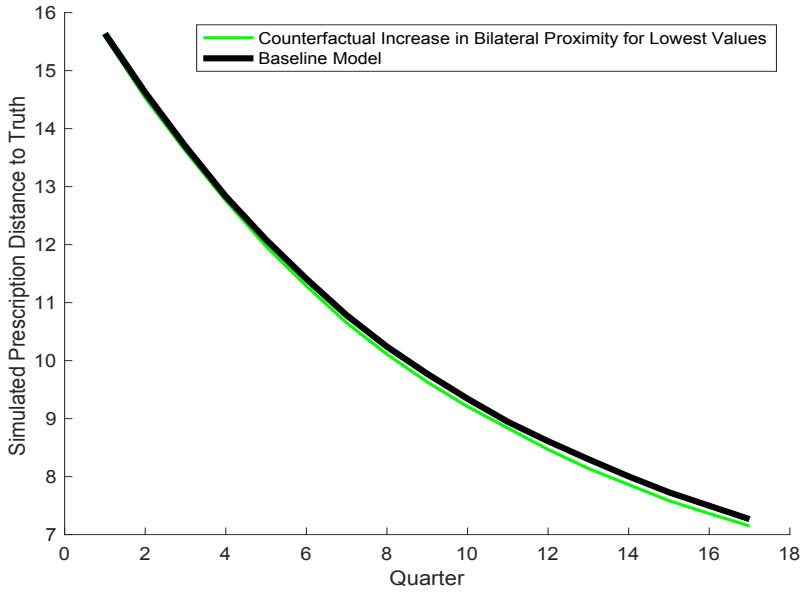
Notes: This figure plots the average (across doctors i) Euclidean distance between a) the vector of relative prescription shares $\ln \pi^T_{idtk} - \ln \pi^T_{id'tk}$ for doctor i that, as defined in (12), would be implied by doctor i choosing products d according to the true, unconditional product qualities β^T_{dk}, and b) the actual relative prescription shares $\ln \pi_{idt} - \ln \pi_{id't}$ of doctor i for each quarter t in the final spell. To distinguish the convergence mechanism from mechanical changes in the scale of the Euclidean distance measure resulting from product entry, only the four initially-available products are included. Vertical lines demarcate the six spells.

Figure 16: Simulated Mean Euclidean Distance to the Truth, Final Spell



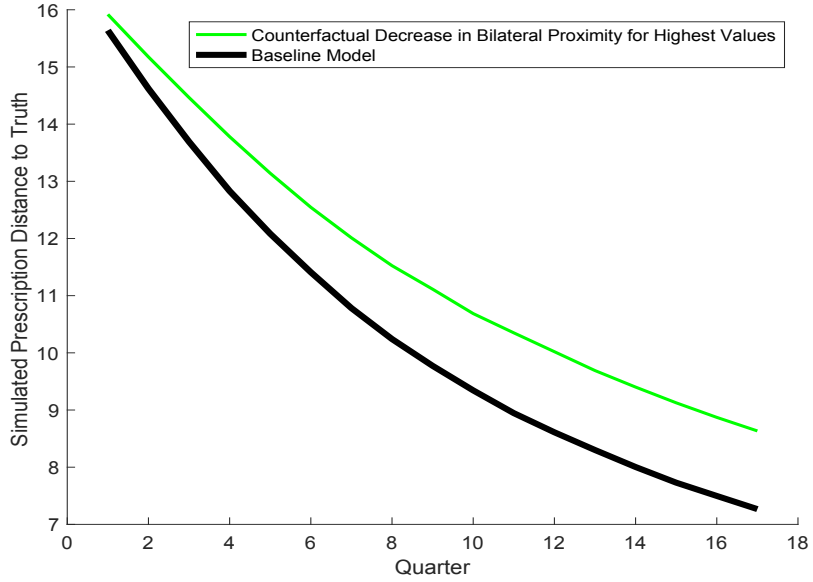
Notes: This figure plots the average (across doctors i) Euclidean distance between a) the vector of relative prescription shares $\ln \pi^T_{idtk} - \ln \pi^T_{id'tk}$ for doctor i that, as defined in (12), would be implied by doctor i choosing products d according to the true, unconditional product qualities β^T_{dk}, and b) the simulated relative prescription shares $\ln \pi_{idt} - \ln \pi_{id't}$ of doctor i for each quarter t in the final spell.

Figure 17: Simulated Mean Euclidean Distance to the Truth, Strengthen Weak Ties



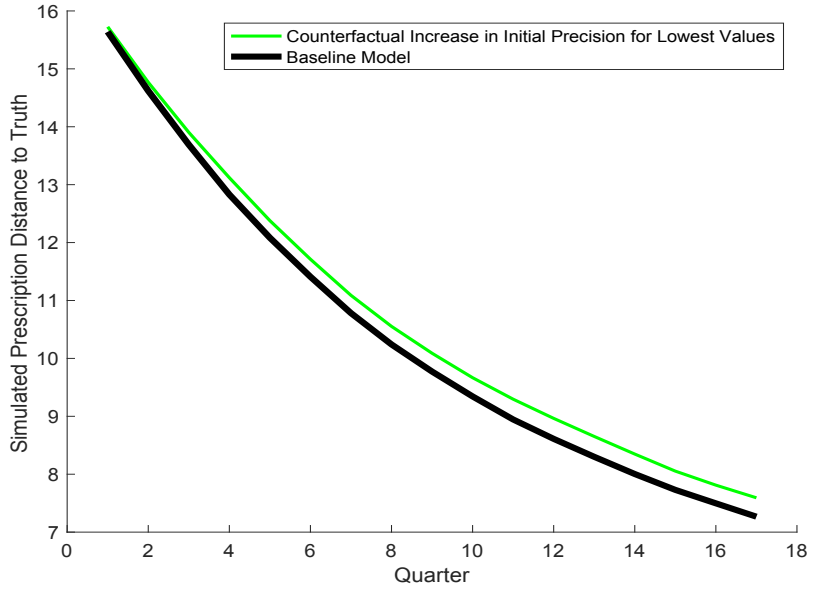
Notes: This figure plots the convergence path in Figure 16 along with the analogous curve (in green) that would result under a counterfactual network in which all first-quartile τ_{ij} values are replaced with the 25th-percentile value.

Figure 18: Simulated Mean Euclidean Distance to the Truth, Weaken Strong Ties



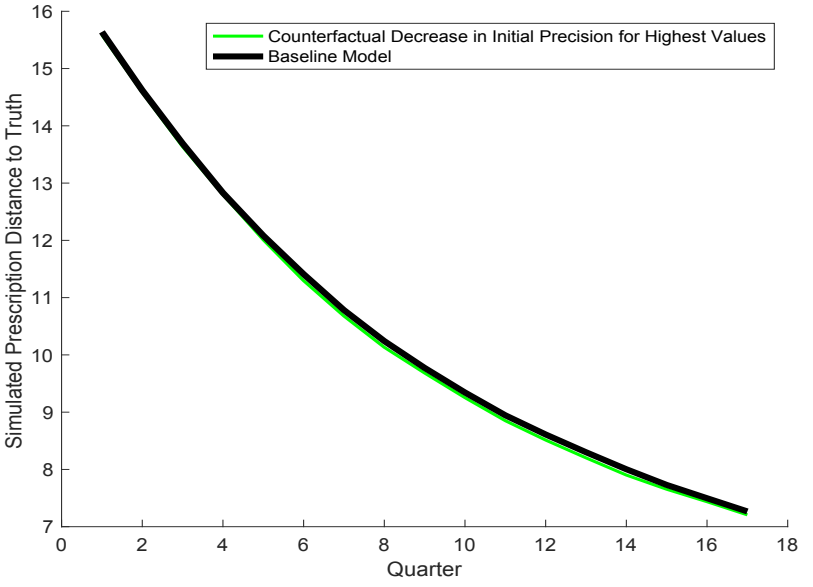
Notes: This figure plots the convergence path in Figure 16 along with the analogous curve (in green) that would result under a counterfactual network in which all top-quartile τ_{ij} values are replaced with the 75th-percentile value.

Figure 19: Simulated Mean Euclidean Distance to the Truth, Increase Precision for Low-S_{i0}



Notes: This figure plots the convergence path in Figure 16 along with the analogous curve (in green) that would result under a counterfactual distribution of initial precision levels in which all first-quartile S_{i0} values are replaced with the 25th-percentile value.

Figure 20: Simulated Mean Euclidean Distance to the Truth, Decrease Precision for High-S_{i0}



Notes: This figure plots the convergence path in Figure 16 along with the analogous curve (in green) that would result under a counterfactual distribution of initial precision levels in which all top-quartile S_{i0} values are replaced with the 75th-percentile value.

Table 14: Convergence, Network Position, and Initial Beliefs Precision

Dependent Variable: λ_i	(1)	(2)	(3)
Network Connectedness $\overline{\tau}_i$	-7.95***	-13.0***	-18.1***
	(0.75)	(4.3)	(4.6)
Beliefs' Precision S_{i0}			0.041***
			(.013)
Zipcode FE	N	Y	Y
R^2	0.0435	0.6872	0.6923
Number of Observations	2,477	2,477	2,477

Notes: *a* denotes 1% significance, *b* denotes 5% significance, *c* denotes 10% significance. This table provides least-squares estimates of (18). Standard errors are shown below each point estimate.

Appendix (For Online Publication)

A.1 Expected Utility

This section includes a detailed derivation of (1). In particular, plugging in the definition of $U_{idt}(\varepsilon_{dt}(\nu_i), \beta)$ and rearranging terms, we can show the result as follows,

$$\begin{aligned}
\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i), \beta_{idt}, \sigma_{idt}) &\equiv \mathbb{E}_{\beta}[U_{idt}(\varepsilon_{dt}(\nu_i), \beta)|\beta_{idt}, \sigma_{idt}] \\
&= \int U_{idt}(\varepsilon_{dt}(\nu_i), \beta) dG_{idt} \\
&= \int -\exp(-\alpha(\beta + \varepsilon_{dt}(\nu_i))) dG_{idt} \\
&= -\exp(-\alpha\varepsilon_{dt}(\nu_i)) \int \exp(-\alpha\beta) dG_{idt} \\
&= -\exp(-\alpha\varepsilon_{dt}(\nu_i)) \exp(-\alpha\beta_{idt} + \frac{1}{2}\alpha^2\sigma_{idt}^2) \\
&= -\exp\left(-\alpha\beta_{idt} + \frac{\alpha^2\sigma_{idt}^2}{2} - \alpha\varepsilon_{dt}(\nu_i)\right),
\end{aligned}$$

where the last step uses the fact that for $Y \sim N(\mu, \sigma^2)$, $\mathbb{E}[\exp(aY)] = \exp(a\mu + \frac{1}{2}a^2\sigma^2)$.

A.2 Choice Probabilities

This section provides a detailed proof for Result 1. First, apply the following monotone transformation $h(\cdot)$ of $\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i), \beta_{idt}, \sigma_{idt})$ to the last line above, where

$$h(z) = -\log(-z)/\alpha.$$

Notice that,

$$\begin{aligned}
h(\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i), \beta_{idt}, \sigma_{idt})) &= h\left(-\exp\left(-\alpha\beta_{idt} + \frac{\alpha^2\sigma_{idt}^2}{2} - \alpha\varepsilon_{dt}(\nu_i)\right)\right) \\
&= \alpha^{-1}\log\left(\left[\exp(-\alpha\varepsilon_{dt}(\nu_i))\exp\left(-\alpha\beta_{idt} + \frac{\alpha^2\sigma_{idt}^2}{2}\right)\right]^{-1}\right) \\
&= \alpha^{-1}\log\left(\exp(\alpha\varepsilon_{dt}(\nu_i))\exp\left(\alpha\beta_{idt} - \frac{\alpha^2\sigma_{idt}^2}{2}\right)\right) \\
&= \alpha^{-1}\left(\alpha\varepsilon_{dt}(\nu_i) + \alpha\beta_{idt} - \frac{\alpha^2\sigma_{idt}^2}{2}\right) \\
&= \beta_{idt} - \frac{\alpha\sigma_{idt}^2}{2} + \varepsilon_{dt}(\nu_i).
\end{aligned}$$

Given the assumption above that $\varepsilon_{dt}(\nu_i)$ follows a Gumbel distribution, the decision rule

$$\widehat{d}^*(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = \underset{d \in D}{\operatorname{argmax}} \left\{\beta_{idt} - \frac{\alpha\sigma_{idt}^2}{2} + \varepsilon_{dt}(\nu_i)\right\}$$

corresponds to the standard multinomial choice problem, for which is it well known (e.g. Train 2009, chapter 3) that agent i chooses product d for client ν_i with probability

$$\begin{aligned}\pi_{idt}(\boldsymbol{\beta}_{it},\boldsymbol{\sigma}_{it}) &\equiv \Pr\left\{\beta_{idt}-\frac{\alpha\sigma_{idt}^2}{2}+\epsilon_{dt}(\nu_i)>\beta_{id't}-\frac{\alpha\sigma_{id't}^2}{2}+\epsilon_{d't}(\nu_i),\ \forall d'\neq d\right\}\\ &=\frac{\exp(\beta_{idt}-\alpha\sigma_{idt}^2/2)}{\sum_{d'\in\mathcal{D}}\exp(\beta_{id't}-\alpha\sigma_{id't}^2/2)},\end{aligned}$$

where the last line is (3).

A.3 Expected Period Payoff

This section provides a detailed proof for Result 1. For simplicity, consider first the case in which i serves a unit-measure continuum of clients $R_{it}=1$ at t. Then,

$$\begin{aligned}W_{it}(\boldsymbol{\beta}_{it},\boldsymbol{\sigma}_{it}) &= E_\varepsilon[U_{it}(\boldsymbol{\varepsilon_t}(\nu_i),\boldsymbol{\beta}_{it},\boldsymbol{\sigma}_{it})]\\ &= E_\varepsilon[\max_{d\in D}\{\mathcal{U}_{idt}(\varepsilon_{dt}(\nu_i),\beta_{idt},\sigma_{idt})\}]\\ &= E_\varepsilon\left[\max_{d\in D}\left\{-\exp\left(-\alpha\beta_{idt}+\frac{\alpha^2\sigma_{idt}^2}{2}-\alpha\varepsilon_{dt}(\nu_i)\right)\right\}\right]\\ &= E_\varepsilon\left[-\exp\left(-\max_{d\in D}\left\{\left(\alpha\beta_{idt}-\frac{\alpha^2\sigma_{idt}^2}{2}+\alpha\varepsilon_{dt}(\nu_i)\right)\right\}\right)\right],\end{aligned}$$

where the last line relies on the strict monotonicity of the function $h(x)=-\exp(-x)$. Defining the following random variable,

$$g_{it}=\max_{d\in D}\left\{\alpha\beta_{idt}-\frac{\alpha^2\sigma_{idt}^2}{2}+\alpha\varepsilon_{dt}(\nu_i)\right\}$$

notice that the distribution of g_{it} may be characterized as follows,

$$\begin{aligned}G_{it}(x) &\equiv \Pr[g_{it}\leq x]\\ &= \Pr\left[\max_{d\in D}\left\{\alpha\beta_{idt}-\frac{\alpha^2\sigma_{idt}^2}{2}+\alpha\varepsilon_{dt}(\nu_i)\right\}\leq x\right]\\ &= \prod_{d\in\mathcal{D}}F\left(\frac{x}{\alpha}-\beta_{idt}+\frac{\alpha\sigma_{idt}^2}{2}\right)\\ &= \prod_{d\in\mathcal{D}}\exp\left(-\exp\left(\frac{-\frac{x}{\alpha}+\beta_{idt}-\frac{\alpha\sigma_{idt}^2}{2}}{\theta}\right)\right)\\ &= \exp\left(-\exp\left(-\frac{x}{\alpha\theta}\right)\times\sum_{d\in\mathcal{D}}\exp\left(\frac{1}{\theta}\left(\beta_{idt}-\frac{\alpha\sigma_{idt}^2}{2}\right)\right)\right)\\ &= \exp\left(-\exp\left(\frac{-x+H_{it}}{\alpha\theta}\right)\right),\end{aligned}$$

where $H_{it}\equiv\theta\alpha\ln\sum_{d\in\mathcal{D}}\exp\left(\frac{1}{\theta}\left(\beta_{idt}-\frac{\alpha\sigma_{idt}^2}{2}\right)\right)$. Thus, G_{it} is a Gumbel distribution with shape parameter $\theta\alpha$ and location parameter H_{it}, implying

$$\begin{aligned}W_{it}(\boldsymbol{\beta}_{it},\boldsymbol{\sigma}_{it}) &= E_\varepsilon[-\exp(-g_{it})]\\ &= -E[\exp(-g_{it})]\\ &= -E[\exp(zg_{it})|_{z=-1}]\\ &= -\Gamma(1+\theta\alpha)\times\exp(-H_{it}),\end{aligned}$$

where the last line uses the moment generating function of the Gumbel distribution. Adding the definition of H_{it}, we thus arrive at

$$W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = -\Gamma(1+\theta\alpha) \times \left(\sum_{d \in \mathcal{D}} \exp\left(\frac{1}{\theta} \left(\beta_{idt} - \frac{\alpha \sigma_{idt}^2}{2} \right) \right) \right)^{-\theta\alpha}.$$

Noting that the above expected period payoff result relies on the assumption that i serves a continuum of clients, rather than the actual volume of clients served, a doctor serving a measure $R_{it} > 1$ of clients obtains a level of expected utility proportional to that above,

$$W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = -R_{it} \times \Gamma(1+\theta\alpha) \times \left(\sum_{d \in \mathcal{D}} \exp\left(\frac{1}{\theta} \left(\beta_{idt} - \frac{\alpha \sigma_{idt}^2}{2} \right) \right) \right)^{-\theta\alpha},$$

which follows due to the *ex ante* homogeneity of clients in the model.

notice that the distribution of g_{it} may be characterized as follows,

$$\begin{aligned}
G_{it}(x) &\equiv \Pr[g_{it} \leq x] \\
&= \Pr\left[\max_{d \in D}\left\{\alpha\beta_{idt} - \frac{\alpha^2\sigma_{idt}^2}{2} + \alpha\varepsilon_{dt}(\nu_i)\right\} \leq x\right] \\
&= \prod_{d\in\mathcal{D}} F\left(\frac{x}{\alpha} - \beta_{idt} + \frac{\alpha\sigma_{idt}^2}{2}\right) \\
&= \prod_{d\in\mathcal{D}} \exp\left(-\exp\left(\frac{-\frac{x}{\alpha} + \beta_{idt} - \frac{\alpha\sigma_{idt}^2}{2}}{\theta}\right)\right) \\
&= \exp\left(-\exp\left(-\frac{x}{\alpha\theta}\right) \times \sum_{d\in\mathcal{D}} \exp\left(\frac{1}{\theta}\left(\beta_{idt} - \frac{\alpha\sigma_{idt}^2}{2}\right)\right)\right) \\
&= \exp\left(-\exp\left(\frac{-x + H_{it}}{\alpha\theta}\right)\right),
\end{aligned}$$

where $H_{it} \equiv \theta\alpha \ln \sum_{d\in\mathcal{D}} \exp\left(\frac{1}{\theta}\left(\beta_{idt} - \frac{\alpha\sigma_{idt}^2}{2}\right)\right)$. Thus, G_{it} is a Gumbel distribution with shape parameter $\theta\alpha$ and location parameter H_{it}, implying

$$\begin{aligned}
W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) &= E_\varepsilon[-\exp(-g_{it})] \\
&= -E[\exp(-g_{it})] \\
&= -E[\exp(zg_{it})|_{z=-1}] \\
&= -\Gamma(1+\theta\alpha) \times \exp(-H_{it}),
\end{aligned}$$

where the last line uses the moment generating function of the Gumbel distribution. Adding the definition of H_{it}, we thus arrive at

$$W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = -\Gamma(1+\theta\alpha) \times \left(\sum_{d\in\mathcal{D}} \exp\left(\frac{1}{\theta}\left(\beta_{idt} - \frac{\alpha\sigma_{idt}^2}{2}\right)\right)\right)^{-\theta\alpha}.$$

Noting that the above expected period payoff result relies on the assumption that i serves a continuum of clients, rather than the actual volume of clients served, a doctor serving a measure $R_{it} > 1$ of clients obtains a level of expected utility proportional to that above,

$$W_{it}(\boldsymbol{\beta}_{it}, \boldsymbol{\sigma}_{it}) = -R_{it} \times \Gamma(1+\theta\alpha) \times \left(\sum_{d\in\mathcal{D}} \exp\left(\frac{1}{\theta}\left(\beta_{idt} - \frac{\alpha\sigma_{idt}^2}{2}\right)\right)\right)^{-\theta\alpha},$$

which follows due to the *ex ante* homogeneity of clients in the model.