Private experience and observational learning in pharmaceutical demand

Tanja Saxell[*†]

February 13, 2014

I quantify the roles of the physician's own experience and the past choices of other doctors in pharmaceutical demand. I develop a model of medical decision-making under uncertainty about the quality of the match between the patient and drug treatment. Unlike previous demand models, I take into account both private and social learning, and allow heterogeneity in product quality across individuals. I test whether information on the past choices of other doctors improves drug choices. Using rich data from the market for cholesterol drugs, I show that treatment patterns relying heavily on the past choices of other doctors can lead to over-prescribing in terms of efficiency. My results suggest that continuity of care, where a patient is repeatedly consulting the same doctor, is an efficient policy to limit such behavior.

Keywords: social and private learning, structural modeling, unobserved quality, asymmetric information, demand, information diffusion, physician behavior

[*]Government Institute for Economic Research and University of Helsinki, email. tanja.saxell@vatt.fi

[†]I am deeply grateful to my advisor Otto Toivanen for numerous useful suggestions and continuous encouragement. I also received very helpful comments from Gregory Crawford, Tuomas Takalo, James Burgess, Pauli Murto, Juuso Välimäki, Matti Liski and seminar participants at the Rising Star Session for Health Markets at the International Industrial Organization Conference, the 9th World Congress on Health Economics, the Annual Conference of the Royal Economic Society and the HECER Microeconomics and Industrial Organization Workshop. I am very thankful to the Social Insurance Institution of Finland for providing me data.

1 Introduction

Consumers may use their own experiences to learn about product quality. At the same time, they may follow the past choices of their peers to deduce what others believe about quality. Private and social learning are relevant in many markets, including stock trading, demand for restaurant services or houses and firms' investment decisions. Yet, there is very little previous work quantifying whether private and social learning help to reduce uncertainty around choices. In this paper, I explore these issues in pharmaceutical demand under uncertainty about the quality of the match between the patient and drug treatment. I will show that treatment patterns relying heavily on the past choices of other doctors can lead to over-prescribing in terms of efficiency. I analyze whether continuity of care, where a patient is repeatedly consulting the same doctor, is an efficient policy to limit such behavior. The policy is commonly used in primary care to promote the process of learning and to improve medical decision-making:

> However, there are other aspects to the doctor-patient relationship that have important implications on efficiency. The distinctive feature of general practice agency is that the doctor-patient relationship is usually long-term and more likely to be characterized by repeated transactions [...] In general practice repeated transactions are also potentially beneficial because the GP becomes more aware of the context of the patients' health problems, and has more information about the patients' medical history, social circumstances, values and preferences.
> Anthony Scott (2000), *Handbook of Health Economics*

I develop a model of medical decision-making to ask the following questions.[1] Is continuity of care preferable to providing information on the past behavior of other doctors through patient records? What are the implications of the long-term doctor-patient relationship on learning and the efficiency of drug choices? Does continuity matter for health care costs? In the model, each patient reacts differently to the drug treatment, and physicians may learn the individual match quality from their own experiences and the treatment choices of patient's previous doctors. I assume that the physicians of the patient may change the drug therapy, and the number of the doctor-patient consultations determines the physician's private experience, or the number of signals, about the match quality. I focus on the Finnish market for cholesterol drugs that are used to decrease the risk

[1] I consider physicians' decisions to continue the drug therapy of a patient in primary care. The model can be extended to allow multiple inside goods. This is very straightforward if the health effects of only one drug group, say patented products, are uncertain.

for cardiovascular diseases. Benefits from improvements in the drug treatment of high cholesterol can be substantial, as heart disease and stroke alone are among the most widespread and costly diseases. Still, many doctors claim that cholesterol drugs are over-prescribed to many low-risk patients.[2]

The model helps to understand why continuity of care can improve physician decision-making. Repeated consultations with the same physician are beneficial, because the physician becomes over time more familiar with the patient's disease and her perceptions on the distribution of health effects become more precise. The physician may thus learn whether the drug is on average good or bad for the patient which improves her treatment choices. If the physicians of the patient change, physicians try to learn the match quality from the treatment choices of previous doctors. As a result of social learning, physicians may start to follow or imitate the treatment choices of previous doctors. An inexperienced physician may belief that the drug treatment must perform well for the patient who has used the drug for many years. This optimism leads to over-prescribing when the drug is of low quality.

The model predictions are consistent with empirical evidence from health care and pharmaceutical markets. The extensive literature in medicine and economics (see e.g. Weiss and Blustein, 1996, Scott, 2000, King et al., 2008) has documented a positive association between continuity of care and improved health outcomes, such as lower mortality and hospitalization rates. Moreover, at least two observations indicate that the physician's personal experience and peer effects affect drug choices. First, my data from the cholesterol drug market confirms that prescriptions are highly responsive to changes in the length of the doctor-patient relationship.[3] Second, prescription behavior by inexperienced physicians is significantly affected by the choices of prominent physicians, or "opinion leaders" (Nair et al., 2010). In my data, the previous choices of peers affect prescribing behavior especially if a physician does not have much own experience of the patient.

A vast majority of the literature on demand for experience goods assumes that agents can only learn the quality of a product from their own experience (e.g. Crawford and Shum, 2005, Kim, 2010, Dickstein, 2011, Chan and Hamilton, 2006, Chernew et al. 2008) or that all information is public (e.g. Ackerberg, 2003, Ching, 2009). A few recent papers also look at the social learning of an agent who makes a once-in-a-lifetime decision (Cipriani

[2]See e.g. Franklin (2011), Adams (2011), Joelving (2011), BBC (2011).

[3]Specifically, I consider the choices of physicians working in the Finnish public primary care. In this market, the physicians of a patient change frequently for exogenous reasons, such as due to the shortage of physician labor. See section 2.1 for details.

and Guarino, 2012, Knight and Schiff, 2010, Zhang, 2010). My main contribution is that I take into account both private and social learning in demand. I modify the standard models of social learning (Chamley, 2004, Bikhchandani, Hirshleifer and Welch, 1992), by allowing agents to learn product quality also from their own experiences. With my framework, I can analyze how the own consumption experiences of an agent interact with information received from the past choices of peers in her learning process.[4] Furthermore, because private and social learning may induce divergent beliefs about quality, a demand model should capture them both in order to produce reliable estimates on product quality and on the effects of policy experiments on choices.[5] Finally, unlike the previous work on social learning, I allow heterogeneity (among patients) in quality.

I find that the average health effects of the cholesterol drug treatment are heterogeneous across patients. Particularly, the quality of the match is on average high for 72% of patients and low for the remainder. The estimates also imply that most of the uncertainty associated with quality vanishes when the patient has used the cholesterol drug treatment once. Even if quality was known, uncertainty regarding to health effects remains significant. These results have implications on efficiency.

The counterfactual experiments suggest that information on the patient's prescription history does not compensate for the lack of the long-term treatment relationship. If the patient had only one physician, the physician learns fast and better health outcomes realize. If quality is high (low), the long-term doctor-patient relationship increases (decreases) demand for cholesterol drugs. Information on the past choices of other doctors for a patient promotes learning about high quality, but not as efficiently as continuity of care. If quality is low, observing the patient's prescription history increases demand over the level of efficient prescribing.

The rest of the paper is organized as follows. Section 2 describes the dataset and provides descriptive evidence on the effects of physician's own experience and the past choices of other doctors on medical decision-making. Section 3 goes through the structural model and Section 4 discusses estimation and identification. Section 5 presents estimation results, the fit of the model and the results from the counterfactual experiments. Section 6 concludes.

[4] Traditional private and social learning models are special cases of my framework.

[5] If there is private information unobserved by the econometrician, but all information is assumed to be public, quality estimates become biased. Specifically, when quality is in reality high, quality estimate is downwards biased because private information slows down learning and decreases the probability of choosing the product. Low quality estimate is, on the other hand, upwards biased because social learning makes agents too optimistic about quality which increases the probability of choosing the product.

2 Market and data description

2.1 Cholesterol drug markets

Cardiovascular diseases (CVD), such as heart attacks, stroke and high blood pressure, affect millions of people globally. Heart disease and stroke alone are among the most common and costly health problems in Europe and the United States.[6] Patients who have experienced CVDs have to deal with high medical expenditures, lost wages and lower productivity.

I analyze the Finnish market for cholesterol drugs that are used to decrease the risk for cardiovascular events. I focus on statins (HMG-CoA reductase inhibitors) that is the most popular group of cholesterol drugs globally.[7] Statins decrease high serum LDL-cholesterol ("bad" cholesterol) and increase HDL-cholesterol ("good" cholesterol) by inhibiting an enzyme in the liver that has an important role in the production of cholesterol.[8] High morbidity to CVDs and a large volume of diagnoses of dyslipidemia, i.e. an abnormal amount of lipids, such as cholesterol and fat, in the blood, have made cholesterol drugs one of the world's largest selling drug groups.

Corresponding to the United States, the following active ingredients are on the Finnish statin market: Atorvastatin (Lipitor and Torvast), Fluvastatin (Lescol), Lovastatin (Mevacor, Altocor, Altoprev), Pravastatin (Pravachol, Selektine, Lipostat), Rosuvastatin (Crestor) and Simvastatin (Zocor, Lipex).[9] I focus on a physician's decision to continue the patient's statin therapy for several reasons. First, uncertainty is probably the highest in the health effects of statins in general. Second, clinical differences between statins in reducing cardiovascular events have been claimed to be small (National Institute for Health and Clinical Excellence, 2006) and thus it is quite natural to consider statins as a one group. I thereby ignore important questions regarding to a physician's or patient's choice between

[6] Around 12% of adults suffered from heart disease in $2009-2010$ in the United States (National Center for Health Statistics, 2011). Every year, there are around 152 000 strokes in the UK (British Heart Foundation, 2013).

[7] See e.g. Herper, M. (2010) *"Why You May Need Cholesterol Drugs"*, Forbes, and U.S. Food and Drug Administration (FDA), 2010.

[8] When cholesterol levels are too high, cholesterol can grow on the walls of blood vessels transporting blood from the heart to other body parts. Over time, these blood vessels can be blocked, preventing the heart from getting enough blood.See e.g. "What is cholesterol?" by the National Heart, Lung and Blood Institute that is a division of the National Institutes of Health in the USA.

[9] Within the group of an active ingredient, statins differ also in the form of drugs, package sizes, strengths and prices. I do not consider a combination preparations of a statin and an another active ingredient.

branded and generic products (see e.g. Scott-Morton, 1999, Ching, 2010a and 2010b) and between different active ingredients (see Crawford and Shum, 2005).[10]

A treatment decision by a physician is based on the benefits and adverse effects of statins. The statin therapy is initiated if the patient has a high risk for CVDs. The evaluation of the risk is based on several factors, including the patient's gender, age, blood pressure and cholesterol levels. In my model, the initial evaluation is captured by the physician's prior belief on the average health effect of cholesterol drugs for a particular patient. In the follow-up of the drug therapy, a physician evaluates the realization of the treatment goals and sustains the patient's treatment motivation. The main goal of cholesterol drug treatment is to decrease the total cholesterol level below 5 mmol/L (LDL-cholesterol below 3 mmol/L). If the patient experiences side effects, the physician decreases the dosage, experiments with an another statin or suspends the cholesterol drug therapy (the Finnish current care for dyslipidemia, 2011).[11] As patients respond differently to statins (the Finnish current care for dyslipidemia, 2011, Jousilahti, 2004), a physician may not know the efficacy and side effects for a single patient.[12] I take the uncertainty into account and let the physician to learn the average health effects of statins by observing realized health effects and the patient's past statin prescriptions.

Cholesterol drugs are also particularly interesting as there is no consensus on an appropriate level of cholesterol drug prescribing. Some doctors have claimed that there is a little evidence that statins reduce the CVDs of low-risk individuals. Doctors supporting the use of statins have said that they have prevented heart attacks and other CVDs.[13] In my model, physicians disagree on the health effects of statins, depending on their personal

[10] I also assume that the physician decides to end the patient's medical treatment. In practice, the final decision to end the therapy can be done either by the physician or the patient or both.

[11] Lifestyle changes, including exercising and changes in diet, are often adequate for a low-risk patient. However, patients are often unwilling to change their lifestyles, even after having a significant shock in their life. Perhaps 45% of smokers stop smoking after a myocardial infarction which is between 2 or 4 times of the success rate of antismoking clinics. Results are not as good for other cardiovascular risk factors related lifestyle, such as physical exercise or diet. Patients can become even less active after infarction. There is also some evidence that changes in self-reported fat intake in one year after infarction can be small. (Johnston, 1999)

[12] For example, statins are reported being useful for men, post menopausal women and patients who have arterial disease or diabetes. It has also been shown that statins decrease by 15% the mortality rate of patients who were 60 years and older and initially clinically asymptomatic. Genetic susceptibility and certain drug interactions can increase the risk of side effects. For example, approximately 5% of patients have been reported suffering muscular symptoms and an increase in the activity of serum muscular enzymes appears for $0.5 - 2.0\%$ of statin users, even though its clinical significance is often uncertain. (The Finnish current care for dyslipidemia, 2011)

[13] See e.g. Adams (2011), Joelving (2011), BBC (2011).

experience of the patient.

Two features of the Finnish market simplify my empirical analysis. The first is that a choice of a physician by a patient was very restricted in public primary care. During the observation period, the patient was not allowed to choose the health center. Within the health center, the patient's family physician was (exogenously) determined based on the patient's residential area (Finnish Medical Association, FMA, 2007).[14] However, due to the shortage of physician labor, patients were not often treated by their own family physicians.[15] I assume that a physician is exogenously determined for the patient in primary care.[16]

The second feature is that two characteristics of the Finnish statin market decrease variation in drug prices over time. First, drugs are subject to price cap regulation by the Pharmaceuticals Pricing Board that is subordinated to the Ministry of Social Affairs and Health in Finland. Second, the patents of Fluvastatin, Atorvastatin and Rosuvastatin remained effective during the whole observation period $2003 - 2006$. As patent protection limits competition, it is likely that the prices ceilings of the patented products were binding. In the empirical analysis, I follow much of the previous learning literature (e.g. Crawford and Shum, 2005) and assume that the drug prices are exogenous. The assumption simplifies the construction of the structural model as prices do not adjust with the observed behavior of physicians.[17]

2.2 Information transmission between physicians

In the model, I assume that a physician has personal experience about the patient-specific quality of the drug treatment. As MD Epstein (1999) illustrates in *the Journal of the American Medical Association*: "Clinical judgment is based on both explicit and tacit

[14]Family physician practices are widely adopted in many countries. For example in the USA, The American Academy of Family Physicians (AAFP) is one of the largest national medical organizations. See AAFP, http://www.aafp.org.

[15]For example in 2006, 9% of the appointments in health centers had a shortfall of physicians and almost the same share of working-age physicians were absent from their permanent jobs. In 46% of these cases, this was caused by staying abroad (FMA, 2006c). It has been estimated that 90% of family physicians treat other than their own patients every week (see FMA, 2005, 2006a, 2006c, 2007).

[16]To be more specific, I assume that the probability of getting a certain physician does not depend on the statin treatment or the health of the patient. This probability is needed to recover the choice probability for the outside good.

[17]In the financial market application of Cipriani and Guarino (2012), bid and ask prices (prices at which a trader can buy and sell) are endogenous because they reflect public information containing the history of trades and prices.

knowledge. Medical decision-making, however, is often presented only as a conscious application to the patient's problem of explicitly defined rules and objectively verifiable data. [...] Seasoned practitioners also apply to their practice a large body of knowledge, skills, values, and experiences that are not explicitly stated by or known to them. [...] While explicit elements of practice are taught formally, tacit elements are usually learned during observation and practice." In this section, I evaluate the validity of the assumption on private information further by discussing the information content of patient records and communication between physicians.

Patient records

A patient record documents and transfers information on a single patient's medication between physician. If all relevant information for medical decision-making is available in the record, a physician does not have any private information of the patient. To see whether this is the case, I next consider the information content of patient records.

The focus of patient records is on the patient's medical condition and medication.[18] To see what type of information is stored in patient records, consider an example of a patient record for a dispensary admission in Appendix B. The patient record provides a compact description of the patient's health status and the plan, the goal and the follow-up of the treatment. It also includes the name of the physician, the list of current medication and a brief justification for starting a medical treatment. In general, patient records may also contain information on whether medication is permanent and reasons for a physician's decision to end the patient's drug therapy.[19]

Patient records do not perfectly transfer all relevant information for medical decision-making between physicians. The case example demonstrates that the continuation of drug therapy is not justified (Appendix B). According to an interviewed specialist, this is a very common practice, at least in routine cases. Records do not include physician-specific factors, such as the physician's own preferences for medication and information on whether her medical decision-making is based on medical literature, advertising and treatment

[18]Patient records regarding to medication include entries about the need of pharmacotherapy and medical foundations, a prescription and given medical treatment, including the name, quantity, form, dosage, dosage form, the date and time of issue of a drug and the name of the physician who has given or prescribed the drug (The Ministry of Social Affairs and Health, 2005).

[19]Essential information in electronic patient documents are reported in the following guidebook and its updated versions (in Finnish): "Opas Ydintietojen, otsikoiden ja näkymien toteuttaminen sähköisessä potilaskertomuksessa", version 1.1, 28.2.2006.

recommendations. The physician's accumulated knowledge of the patient's preferences, values and circumstances is rarely recorded (see Guthrie et al., 2008). The specialist also claimed that a narrative text format complicates the interpretation of records that may impede information transmission. The registering of information takes the physician's time that may decrease her incentives to record all relevant information.

Communication

I evaluate next whether all relevant information for medical decision-making is transfered through communication. A physician who cares about her patient may want to consult her colleagues before deciding on the continuation of the treatment. Because communication is time-consuming, consultation does not probably happen in routine cases. On the other hand, the patient, who wants to get as good medical treatment as possible, may want to communicate all relevant information to her physicians. It is, however, unlikely that medical decision-making by physicians is exclusively based on information received from the patient (see e.g. Epstein, 1999).

The theoretical cheap-talk[20] literature (see for example Crawford and Sobel, 1982, Olszewski, 2004) has shown that the truthful information revelation of a consultant (a sender, here: other physicians or a patient) to a decision maker (a receiver, here: a physician) is only one of many possible outcomes, even if there is no disagreement between participants. If the preferences of the consultant are even slightly misalligned with the preferences of the decision maker, there is some information loss in all equilibria (Crawford and Sobel, 1982). If the consultation effort of the physician is unobserved to the patient, incentives for consultation may not be high.

Finally, if all physicians of a patient share the same information, they should have the same probability of choosing the medical treatment. As it turns out in the next section, this is not the case.

[20]In a typical cheap-talk game, the sender may, often costlessly, convey her private information through messages to the receiver. The receiver then takes an action that together with sender's signal affects the payoffs of both players.

2.3 Data

2.3.1 Sample selection

I use a rich dataset of all purchased cholesterol drug prescriptions in Finland from January 1 in 2003 to December 31 in 2006. The data is provided by the Social Insurance Institution of Finland which is responsible for the provision of public social security benefits to Finnish residents. The data identifies patients, their physicians and cholesterol drugs.[21]

I prepare my data for the empirical analysis in the following steps. First, to follow patients from the beginning of cholesterol drug therapy and to avoid left-censoring, I focus on "new" patients who did not have any prescriptions during the first 6 months of the observation period i.e., before July 2003.[22] Second, I ignore patients with multiple prescriptions or physicians within a day to simplify the analysis further. Third, I consider patients whose physicians are primarily working in public health centers. Ideally, I would like to concentrate on patients who have only used the services of public health centers but unfortunately the data does not include this information. As a proportion of physicians work for both the public and the private sectors[23], some patients in the sample may have used private health care services. Fourth, I concentrate on patients who belong to the working-age (15-64 years) population because the data does not allow me to distinguish the death of a patient from the ending of the statin treatment. Finally, for computational reasons, I draw a random sample of 10000 patients from the sample of new working-age patients whose physicians are working in primary care.

2.3.2 Descriptive evidence

In this section, I provide the descriptive analysis of the sample. The results in Table 1 demonstrate that the sample consists of very heterogeneous patients. Most of the patients in my sample were relatively old at the time of the last prescription (an average 51 years) and almost half of the patients were men. The number of diagnosis varies[24]

[21] Other characteristics than the primary job of a physician (public health center/public hospital/other) received from the survey conducted by the Finnish Medical Association (FMA) are from the registers of the Social Insurance Institution of Finland. The response rate of the yearly survey has been very high. For example, in 2006, the response rate of physicians who received the survey was 80% (FMA, 2006c).

[22] This six months' time window has been also used by Crawford and Shum (2005).

[23] In 2006, 19.6% of physicians, who were primarily working in health centers, had a sideline job (FMA, 2006c).

[24] The number of diagnosis is observed if the patient was on sick-leave.

in substantially around its mean (0.7).[25] A significant portion of patients (55%) were censored in the sample i.e., they had their last prescription within the last six months of the observation period.

Table 1: Descriptive statistics for the sample of patients[1]

	Mean	Std.Dev.	Only non-censored patients	At the time of the last prescription
Patient characteristics				
Age	55.03	7.20	No	Yes
Gender (1: male, 0: female)	0.49	0.50	No	Yes
Nbr of diagnosis	0.73	1.31	No	Yes
Censoring indicator (1: yes, 0: no)	0.52	0.50	No	Yes
Patient's medical treatment				
Treatment ending (1: yes, 0: no)	0.34	0.47	Yes	No
Nbr of prescriptions	1.93	1.17	Yes	Yes
Nbr of physicians	1.28	0.58	Yes	Yes
Prescriptions of a current physician	1.676	1.072	No	No
Visit a physician specialized in internal diseases	0.01	0.09	No	No
Visit a non-specialized physician	0.69	0.46	No	No
Total number of physician's prescriptions	1.65	1.07	No	No
Physician change (1: yes, 0: no)[2]	0.33	0.47	No	No
Active ingredient change (1: yes, 0: no)[2]	0.17	0.38	No	No
Price, eur	46.32	49.16	No	No
Number of observations	22 021			

[1] The relevant population consists of new working-age patients who have used statins and the services of public health centers. The size of the random sample is 10 000 patients.

[2] Note that here the number of prescriptions is at least 2 because the change in the value of the variable from the previous prescription is computed by using the difference between its current and lagged value.

Following Crawford and Shum (2005), I assume that the drug therapy of a non-censored patient ends after the last prescription in the data. If the patient is censored, the end of

[25] Information on the number of diagnosis is observed if a patient received sickness benefits from the Social Insurance Institution of Finland.

the therapy is not observed. If the censoring interval is too short, the estimation results may be biased. This is particularly true if the patient's drug treatment is prescribed at the end of the observation period and he has more than two prescriptions.[26] Dickstein (2011) used an alternative approach where the treatment episode of a patient ends at the last prescription if there was a gap of 90 days within the treatment history. A patient appearing in the data again after the gap is then treated as a new patient.

The cholesterol drug therapies of non-censored patients in the sample were on average relatively short, approximately 2 prescriptions (Table 1). The probability that the patient's therapy ends at any stage of therapy is 0.34. The average number of physicians per patient was 1.3 and the total number of prescriptions received from a particular physician was 1.65. Most of the patients (70%) were treated by a non-specialized physician. The average price of a prescription was 41 eur.

Table 2 presents the distribution of the total number of prescriptions and physicians at the time of the (non-censored) patient's last prescription. Most of the non-censored patients (52%) had only one prescription and 80% of the patients were in a permanent physician-patient relationship. Even though the distributions of the total number of prescriptions and physicians are skewed to the right, 48% of non-censored patients had more than one prescription and 20% were treated by more than one physician.

Table 2: The percentage share of non-censored patients in the sample conditional on the total number of prescriptions and physicians at the last prescription

	Physicians			
Prescriptions	1	2	3-	Total
1	51.91	.	.	51.91
2	18.55	8.37	.	26.93
3	6.77	4.80	1.45	13.02
4-	3.13	2.95	2.07	8.15
Total	80.36	16.12	3.52	100.00

I consider next the incidence of a physician change in the sample of patients. Table 1 illustrated that the breakdown of the physician-patient relationship was very common. The probability that the patient's physician changes from the previous prescription was

[26] As a robustness check, I used a one-year censoring interval and defined a patient to be "new" if he did not have prescriptions during the first year. Then, the probability that the patient is censored was somewhat higher (0.73) than with the original censoring interval. The probability that the patient's treatment ends was 0.40 which is fairly close to the corresponding probability with other definition (0.34).

33%. A high standard deviation also indicates significant diversity among patients in the incidence of a physician change.

Then, I analyze how the number of interactions between a physician and a patient affects prescriptions. I consider first how the probability of continuing the (non-censored) patient's statin therapy depends on the lagged number of physicians (Figure 1). I find that the continuation probability is 50% for patients who have only one physician, i.e. who do not have any physician switches. The choice probability decreases to 42% for patients having two physicians and further to 33% for patient with three physicians.

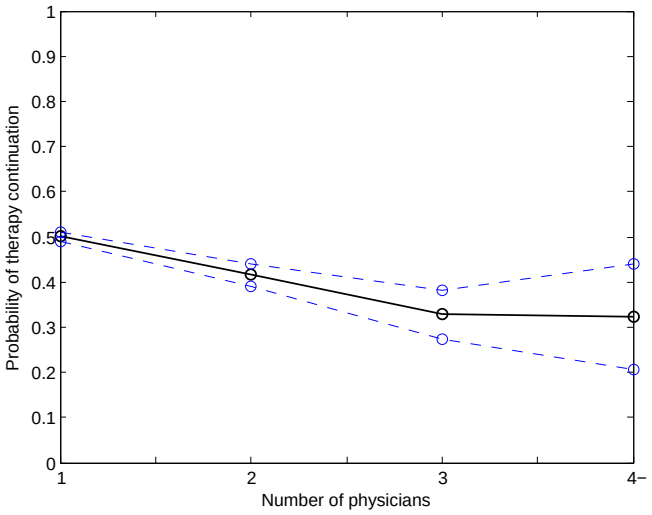

Figure 1: The probability of treatment continuation and its 95% confidence intervals by the number of physicians for non-censored patients, sample averages

I investigate next whether the decreasing pattern between the choice probability and the number of physicians is driven by the phase of the patient's therapy. To see if this is the case, I estimate the following linear probability model for the continuation of the (non-censored) patient's statin therapy,

$$\mathrm{a}_{it} = \alpha + X_{i(t-1)}\beta + e_{it},\ t > 1,$$

where a_{it} is an indicator variable that gets value 1 if the statin therapy of patient i

is continued at time, or prescription, t and 0 otherwise[27], $X_{i(t-1)}$ is a vector of lagged explanatory variables and e_{it} is the error term.

The results presented in Table 3 suggest that the continuation probability increases by 13% when the number of previous physicians increases by one. The lagged length of the doctor-patient relationship has an opposite effect on the continuation probability. These findings may suggest that physicians do not share the same information about the health effects of the cholesterol drug treatment for a patient.

[27] To be more precise, $a_{it} = 0$ only once when the patient's statin therapy ends.

Table 3: Descriptive regressions for the probability of therapy continuation in the sample of non-censored patients

Variable[1]	Model (1)	Model (2)
Constant	0.672***	0.704***
	(0.167)	(0.169)
Own experience:		
prescriptions/current physician		-0.126***
		(0.013)
Nbr of physicians	0.129***	
	(0.017)	
Prescription nbr	-0.156***	-0.046***
	(0.008)	(0.010)
Gender	0.0288**	0.0277*
	(0.0109)	(0.0109)
Age	0.000	0.000
	(0.001)	(0.001)
Nbr of diagnosis	0.006	0.006
	(0.004)	(0.004)
Cost, eur	0.000***	0.000***
	(0.000)	(0.000)
Reimbursement	-0.000***	-0.000***
	(0.000)	(0.000)
Fixed effects:		
physician, ATC-code, hospital district	yes	yes
N	10031	10031
adj. R^2	0.093	0.100

[1] Explanatory variables are lagged by a one prescription.

[2] Variables are for cholesterol drug prescriptions.

[2] Standard errors in parentheses.

[3] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To get further evidence on peer effects and the role of private experience in demand, Table 4 illustrates how medical spending in the sample depends on the length of the physician-patient relationship, after controlling for observed characteristics. When the number of physicians increases by one, the total costs of the therapy at any stage decreases by 7 euros which is 15% of the average costs of statins in the sample. Table 4 also shows that the more the physician has experience of the patient, the less the previous choices of peers

- measured by the number of cholesterol drug prescriptions provided by other doctors to a single patient - affect an average medical spending at any phase of the therapy.[28] When the physicians of a patient change frequently relative to the stage of the drug therapy, the effect of physician's own experience on the total costs becomes small. These results are consistent with "asymmetric peer effects" where inexperienced physicians rely on experienced doctors to decrease uncertainty around their prescription decisions (see e.g. Nair et al., 2010). Still, the findings remain very indicative without putting any structure in the model that helps to isolate the effects of personal experience and social learning on medical decision-making.

[28] I measure the physician's own experience with the number of interactions with the patient.

Table 4: Descriptive regressions for treatment costs in the sample of patients

Explained variable	Total cost,	Total cost, eur[1]	Cost, eur	Total cost, eur[1]
Constant	-92.48***	-102.5***	12.52***	-168.2***
	(17.54)	(18.04)	(2.173)	(23.10)
Nbr of physicians	-7.106*			
	(3.312)			
Own experience:				
prescriptions/current physician		60.63***	0.549***	23.25***
		(2.952)	(0.119)	(5.537)
Other physicians' experience:				
prescriptions/previous physicians		60.20***	0.572***	
		(4.542)	(0.152)	
Own experience*others' experience		-3.718	-0.126**	
		(3.499)	(0.044)	
Physicians/prescriptions				86.71***
				(13.73)
Own experience*				
physicians/prescriptions				-31.65*
				(12.88)
Nbr of prescriptions	57.93***			55.60***
	(3.014)			(4.182)
Reimbursement	0.028***	0.028***	0.019***	0.029***
	(0.001)	(0.001)	(0.000)	(0.001)
Prescription date	0.002***	0.002***	-0.000***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Min prescription date	-0.002***	-0.002***	-0.000	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Age, years	0.020	0.018	-0.015	0.035
	(0.092)	(0.092)	(0.009)	(0.091)
Gender	1.722	1.775	0.165	1.499
	(1.531)	(1.513)	(0.149)	(1.486)
Nbr of diagnosis	-0.0260	-0.0343	-0.151**	-0.0927
	(0.479)	(0.489)	(0.059)	(0.463)
Fixed effects:				
physician, ATC-code, hospital district	yes	yes	yes	yes
N	22183	22183	22183	22183
adj. R^2	0.715	0.716	0.974	0.723

[1] Total (cumulative) costs at a given stage of the therapy.

[2] Variables are for cholesterol drug prescriptions.

[2] Standard errors in parentheses.

[3] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3 A theoretical model of pharmaceutical demand

3.1 Overview

In this section, I present a structural model of medical decision-making with private experience and observational learning. In each period during the drug therapy, the patient (he) is randomly matched to a physician (she). After an initial treatment choice, the physician investigates the patient and gets private information about the quality of the match between the patient and the drug treatment. During the course of the patient's therapy, the physician may learn quality from from her own experience and the previous choices of other doctors for this particular patient.[29]

Consider patient i who comes for the first time to a public health center to seek drug treatment for her medical condition. After entrance, a physician is randomly assigned to the patient. As the sensitivity of patients to cholesterol drugs differ, the physician does not know ex-ante the average health effects, or quality, of the drug treatment for this particular patient. To form the prior belief on quality, the physician evaluates the patient's risk for CVDs based on the patient's observed characteristics. The physician takes the prior belief and her privately observed idiosyncratic preferences into account when she decides whether to initiate the cholesterol drug therapy.

In the follow-up of the drug therapy at time (or prescription number) t, patient i comes again to the health center where he is randomly matched physician l. First, the physician performs a diagnostic procedure, physical examination and tests for the patient to privately evaluate the efficacy and side effects of the drug treatment. This evaluation is modeled by an experience signal x_{ilt}. Simultaneously, she looks at patient records to see how long the patient has been using the drug. Conditional on the prior, the past choices of other doctors indexed by $l_1, ..., l_{t-1}$, $h_{it} = \{a_{il_1 1}, ..., a_{il_{t-1}(t-1)}\}$, and all private experience signals that the physician has received during the course of the patient's drug therapy up to and including time t, she updates her belief about its quality.

Recall that in previous social learning models (Cipriani and Guarino, 2012, Knight and Schiff, 2010, Zhang, 2010) agents can receive only one experience signal. Based on this posterior belief and her private preference shocks for the drug treatment and the outside good, v_{il1t} and v_{il0t} respectively, the physician makes a decision on the continuation of the

[29] A relatively easy extension of the model is to enrich the choice set of physicians that could include other medical treatment alternatives, such as non-patented products, with the known (to physicians) but possibly random quality. An extension that allows several inside goods with uncertain qualities comes at the cost of computation.

patient's therapy. Further decisions follow until any physician decides to end the drug therapy. The timing of events is summarized by Figure 2.

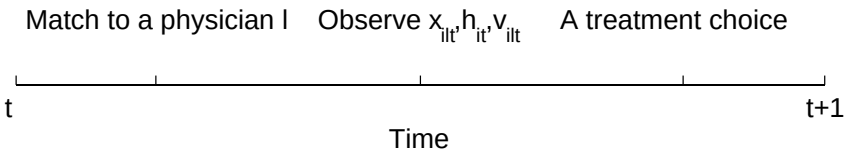

Figure 2: The timing of events in period t during the follow-up of the therapy: 1.) a patient is first matched to a physician, 2.) the physician observes a new signal x_{ilt} and the past choices of other doctors h_{it} and private idiosyncratic preference shocks, v_{il1t} and v_{il0t}, 3.) the physician makes a treatment choice on all her private signals received up to and including time t, public information h_{it} and private preference shocks.

In the long-term treatment relationship, the physician learns about the average health effects of the drug treatment from her own experience. If the relationship breaks down, a physician attempts to infer quality from the past choices of other doctors. The less the physician has own experience of the patient, the more the past choices of peers affect her prescription behavior. If the patient has used the drug treatment long, an inexperienced physician may perceive that the drug must be effective. When the drug is of high quality, observing the past choices of other doctors improves learning. On contrary, the optimism on quality leads to over-prescribing when the drug is of low quality.

To keep the model tractable and to avoid the salient computational burden, I assume that a physician maximizes her expected per-period utility. The assumption of myopic behavior is often made in the structural learning literature (e.g. Coscelli and Shum, 2004, Ching, 2009, Chernew et al., 2008) and it abstracts away incentives to experiment with the drug treatment to get new information about quality in the next period (see e.g. Crawford and Shum, 2005).[30]

Following e.g. Crawford and Shum (2005) and Dickstein (2011), the model does not take into account learning across patients.[31] This type of learning could be incorporated to the

[30] My future plan is to estimate a dynamic version of the model.

[31] For learning across patients, see Kim (2010) and Coscelli and Shum (2004). Note also that Crawford

model by using the entry of a new active ingredient, Rosuvastatin. This extension comes again with the cost of computation and tractability because physicians and the econometrician have to keep track on the posteriors of all doctors. Because many cholesterol drugs have been on the market since the end of the 1980s or the early 1990s, learning about the distribution of health effects across patients does not probably have a significant role in my application.

In the following sections, I present the model in detail. I first formulate a deterministic process governing the assignment of a physician for a patient.[32] Because the physician is not forward-looking in her treatment continuation choices, the assignment, or matching, probability does not affect her behavior. Then, I describe a therapy continuation choice under uncertainty and the information structure, including the distribution of signals (health effects) and the patient-specific quality. Finally, I derive the posterior belief of the physician about quality, conditional on her private experience and the patient's prescription history.

3.2 The theoretical model

3.2.1 Physician and patient matching

In each period until the therapy ends, patient i is assigned to a physician. The physician is either "new" i.e., she does not have the previous treatment relationship with the patient, or is any of the previously drawn "old" physicians $1, ..., N_{it}$. The number of old physicians at time $t+1$ increases by one, $N_{i(t+1)} = N_{it} + 1$, if the new physician treats the patient at time t, and otherwise it remains unchanged, $N_{i(t+1)} = N_{it}$.

I assume that the patient is assigned to the new physician with probability κ_i and to the old physician with probability $(1 - \kappa_{it}) \times \frac{1}{N_{it}}$. This specification implies that each old physician is randomly selected for the patient from the pool of the previously drawn physicians with the same probability $\frac{1}{N_{it}}$.[33]

I assume the following functional form for the matching probability of patient i:

and Shum (2005) allow the possibility of non-rational expectations, because in their model physicians' *prior* beliefs for one particular drug, Omeprazole, can evolve over time, which captures common changes in priors, for example, due to advertising. However, posteriors may also vary through a different type of mechanism, namely based on the previous medication decisions of a particular physician or other doctors.

[32] The assignment probability is used to recover the probability of the outside good (see Section 4.1).

[33] Note that only 3.5% of patients had more than 2 physicians in my data (see Table 2).

$$\kappa_i = P_i(d_{it} = 1) = \frac{e^{y_i}}{1 + e^{y_i}}. \quad (1)$$

In the above expression, y_i is $N(\theta^y, \sigma_y^2)$- distributed patient level random coefficient. The variance of the random coefficient, σ_y^2, measures the magnitude of heterogeneity in matching probabilities across patients. The heterogeneity is potentially important because the probability of a physician change can differ between patients, for example, by residential area.

3.2.2 A therapy continuation choice under uncertainty

Assume that physician l is drawn for patient i at time t. The physician decides whether to continue the drug therapy of patient i, $a_{ilt} = 1$, or end the therapy for good, $a_{ilt} = 0$, conditional on her information at that time, I_{ilt}. In the perfect Bayesian equilibrium, the physician chooses to continue the medical therapy if the expected utility from the medical treatment exceeds the utility from the outside option (the non-purchase option),

$$a_{ilt} = 1 \Leftrightarrow E(u_{il1t}|I_{ilt}) \geq u_{il0t}. \quad (2)$$

I assume that the per-period utility received from the medical treatment, u_{il1t}, depends on the quality signal, or health effects, x_{ilt}, and a vector of control variables, $\mathbf{Z}_{il1t}$. The controls include, for example, the (average) price of statins, observed patient level characteristics and the time trend capturing general market level changes over time due to advertising. These controls are observed by both physicians and the econometrician. Because patient records do not contain information on preference shocks, I assume that the physician's idiosyncratic, Type 1 extreme value distributed tastes for the drug treatment and the outside option, v_{il1t} and v_{il0t}, are her private information. Following the previous literature (e.g. Crawford and Shum, 2005), I assume a Constant Absolute Risk Aversion (CARA) sub-utility specification for the health effects. To be more specific, I consider the following utility function,

$$u(x_{ilt}, \mathbf{Z}_{il1t}, v_{il1t}) = -e^{-r \cdot x_{ilt}} + \mathbf{Z}_{il1t}\boldsymbol{\alpha} + v_{il1t}, \quad (3)$$

where $r > 0$ is the risk aversion coefficient.

I assume that the utility of the outside good for the physician l of patient i at time t, u_{il0t}, is a function of a vector of observed characteristics, $\mathbf{Z}_{il0t}$, and the physician's private preference shock, v_{il0t},

$$u(\mathbf{Z}_{il0t}, v_{il0t}) = \mathbf{Z}_{il0t}\boldsymbol{\beta} + v_{il0t}. \tag{4}$$

To ensure identification in the discrete choice model, I make a typical restriction that the constant of the outside option is zero. Recall that the utility of the outside good varies with the patient's observed characteristics (see Chan and Hamilton, 2006, for a similar approach). For example, cholesterol drugs prevent coronary events in the long-run after the patient's drug therapy has ended.[34] I control this with the number of prescriptions.

3.2.3 Health effects

The quality of the match between the patient and the drug treatment (referred as "quality"), θ_i, is without loss of generality either high θ_1 or low θ_0 with prior probabilities $p_i(\theta_1)$ and $1 - p_i(\theta_1)$, respectively. [35] The variance of random quality, $\text{Var}(\theta_i) = E(\theta_i^2) - (E(\theta_i))^2 = p_i(1 - p_i)(\theta_1^2 + \theta_0^2 - 2\theta_1\theta_0)$, measures prior uncertainty regarding to quality. The prior is uninformative when it equals 1/2.

The prior probability is common knowledge for physicians but it may vary across patients, depending on the patient's observed characteristics. I assume that each physician has the following prior belief that the treatment has high quality for patient i:

$$p_i(\theta_1) = \frac{e^{\gamma_0 + \mathbf{Z}_i^p \boldsymbol{\gamma}_1}}{1 + e^{\gamma_0 + \mathbf{Z}_i^p \boldsymbol{\gamma}_1}}, \tag{5}$$

where $\mathbf{Z}_i^p$ is a vector of patient level characteristics at the time of the first prescription.

In the follow-up of the patient's drug therapy at time $t > 1$, the physician observes an experience signal, or health effects associated with the use of cholesterol drugs. I assume that health effects are independent and normally distributed conditional on the true quality,

$$x_{ilt}|\theta_i \sim N(\theta_i, \sigma^2), \tag{6}$$

[34] The literature has explained this with the stabilization of existing plaque and the slowing of the progression of coronary artery disease (Ford et al., 2007).

[35] The model could be generalized to allow a continuous quality level but the computation of the posterior probability for quality θ conditional on information at time t I_t, $f(\theta|I_t)$, becomes more difficult than in the binary case as it would involve integration over quality levels θ.

where σ^2 measures uncertainty regarding to health effects. The distributions of signals and priors are common knowledge and θ_1, θ_0, σ^2, γ_0 and $\boldsymbol{\gamma}_1$ are parameters to be estimated.[36]

Because prior beliefs are heterogeneous across patients, the unconditional (mixture) density of health effects, $f(x_{ilt})$, depends on the observed characteristics of the patient. This means that the sensitivity of patients on the efficacy and side effects of statins may differ for example by their gender and age, as the medical literature suggests (see Section 1).

3.2.4 A physician's information set

Because signals are private information to physicians, a physician's information set for the patient at time t, I^{θ}_{ilt}, includes her own private experience of the patient and the previous therapy continuation choices of other physicians. Formally, $I^{\theta}_{ilt} = \mathbf{x}_{ilt} \cup h_{it} \setminus \{a_{ilt'}, t' < t\}$ where $\mathbf{x}_{ilt}$ is the set of signals that physician l has received up to (and including) time t and $h_{it} \setminus \{a_{ilt'}, t' < t\}$ is the patient's prescription history, $h_{it} = \{a_{il_1 1}, ..., a_{il_{t-1}(t-1)}\}$, without the physician l's actions, $\{a_{ilt'}, t' < t\}$. Because the preference shocks of physician l are her private information, the final information set of physician l at time t for patient i is given by $I_{ilt} = I^{\theta}_{ilt} \cup \mathbf{v}_{ilt}$ where $\mathbf{v}_{ilt}$ is the set of preference shocks that physician l has received up to (and including) time t.

3.2.5 The expected utility

The expected utility of physician l associated with the continuation of the drug therapy for patient i conditional on her information at time t, I_{ilt}, can be written as:

$$\begin{aligned} E(u_{il1t}|I_{ilt}) &= E_{\theta_i|I} E_{x|\theta_i,I}(-e^{-r x_{ilt}}) + \mathbf{Z}_{il1t}\boldsymbol{\alpha} + v_{il1t} \\ &= E_{\theta_i|I}(-e^{-r\theta_i + \frac{1}{2}r^2\sigma^2}) + \mathbf{Z}_{il1t}\boldsymbol{\alpha} + v_{il1t} \\ &= -\lambda_{ilt} e^{-r\theta_1 + \frac{1}{2}r^2\sigma^2} - (1-\lambda_{ilt}) e^{-r\theta_0 + \frac{1}{2}r^2\sigma^2} + \mathbf{Z}_{il1t}\boldsymbol{\alpha} + v_{il1t}. \end{aligned} \tag{7}$$

$\lambda_{ilt} = Pr(\theta_1|I_{ilt})$ is the posterior probability that quality is high. The first equality follows from the law of iterated expectations and the second one from the moment generating function of the normal distribution.

The expected utility of the risk averse physician decreases with uncertainty about the effect of the drug therapy on the patient's health, σ^2. The risk aversion parameter increases the

[36] The model could be extended to allow unobserved heterogeneity. In this case, the mean and variance of a signal can differ depending on the type of the patient that is observed by his physicians.

expected utility through quality parameters θ_1 and θ_0 and decreases it through the risk premium $\frac{1}{2}r^2\sigma^2$. Clearly, the latter effect starts to dominate when either σ^2 or the risk aversion parameter r is large enough, namely $r > \frac{2\theta_k}{\sigma^2}$, $k \in \{0, 1\}$.

3.2.6 Public and private beliefs

In this section, I describe how the physician updates her beliefs about the quality of the drug treatment. I find that the posterior belief about quality, λ_{ilt}, is a function of the prior and the physician's private and public beliefs. The private belief is the probability of quality, conditional on physician's accumulated private experience of the patient, $\mathbf{x}_{ilt}$. The public belief is the probability of quality, conditional on the past choices of other doctors. I show that the private experience affects the private belief through a sum of signals. It turns out that this property decreases the computational burden of the model substantially. Even though the physician does not observe the private information of other doctors, she tries to infer quality from their past therapy continuation choices.

The posterior belief

Let $P_i(\theta_1|\mathbf{x}_{ilt})$ denote the private belief of physician l that quality is high for patient i at time t conditional on her private experience $\mathbf{x}_{ilt}$. I denote by $q_{ilt} = P(\theta_1|l, h_{it})$ the corresponding public belief that is conditional on the previous therapy continuation decisions of other physicians $l' \neq l$.

Conditional on health effects $\mathbf{x}_{ilt}$ and the past choices of other doctors for patient i, physician l updates her beliefs about the quality of the treatment for patient i using Bayes' rule and the iid nature of the health effects,

$$\begin{aligned} \lambda_{ilt} &= P_i(\theta_1|l, h_{it}, \mathbf{x}_{ilt}) \\ &= \frac{P(h_{it}|l, \theta_1) f(\mathbf{x}_{ilt}|\theta_1) p_i(\theta_1)}{P(h_{it}|l, \theta_1) f(\mathbf{x}_{ilt}|\theta_1) p_i(\theta_1) + P(h_{it}|l, \theta_0) f(\mathbf{x}_{ilt}|\theta_0) p_i(\theta_0)}. \end{aligned} \tag{8}$$

In the above expression, $P(h_{it}|l, \theta)$ is the probability of other doctors' treatment continuation choices for the patient and $f(\mathbf{x}_{ilt}|\theta)$ is the probability of health effects, conditional on the true quality of the drug, $\theta \in \{\theta_0, \theta_1\}$.

The posterior can be linked to the prior, private and public beliefs as follows:

$$\begin{aligned}\lambda_{ilt} &= \frac{q_{ilt}f(\mathbf{x}_{ilt}|\theta_1)}{q_{ilt}f(\mathbf{x}_{ilt}|\theta_1)+(1-q_{ilt})f(\mathbf{x}_{ilt}|\theta_0)} \\ &= \frac{q_{ilt}P_i(\theta_1|\mathbf{x}_{ilt})/p_i(\theta_1)}{q_{ilt}P_i(\theta_1|\mathbf{x}_{ilt})/p_i(\theta_1)+(1-q_{ilt})P_i(\theta_0|\mathbf{x}_{ilt})/p_i(\theta_0)},\end{aligned} \tag{9}$$

where the first equality follows from (8). To see this, multiply and divide (8) by $1/P(l,h_{it})$ and note that $q_{ilt}=\frac{P_i(h_{it}|l,\theta_1)p_i(\theta_1)}{P(l,h_{it})}$ where $P(l,h_{it})$ is the probability of the public medication history of the patient without the physician l's actions. The second equality in (9) follows from the first one by dividing and multiplying the first equality by $1/f(\mathbf{x}_{ilt})$ and by observing that $\frac{f(\mathbf{x}_{ilt}|\theta)}{f(\mathbf{x}_{ilt})}=\frac{P(\theta|\mathbf{x}_{ilt})}{p(\theta)}$ for $\theta\in\{\theta_0,\theta_1\}$.

The posterior belief is determined by the prior, $p_i(\theta_1)$, and private and public beliefs, $P_i(\theta_1|\mathbf{x}_{ilt})$ and q_{ilt}. When the public (private) belief is uninformative (equals 1/2), the posterior belief depends only on the private (public) and prior beliefs. When the physician puts weight only on her prior and private experience, the model corresponds to a traditional structural learning model where agents learn only from their private experience (see e.g. Coscelli and Shum, 2004, Crawford and Shum, 2005, Ackerberg, 2003). Recall also that the posterior is an increasing function of private and public beliefs. Hence the higher these beliefs are, the more confident the physician becomes that the quality of the medical treatment is high.

The last step is to derive the evolution of private and public beliefs.

The private belief

First, I describe how the physician learns from her private experience. Assume that the physician has seen the patient S times in the follow-up of the therapy and has observed health $x_{il1},...,x_{ilS}$. Denote by $f(x_{il1},...,x_{ilS}|\theta)$ the joint probability of health effects $x_{il1},...,x_{ilS}$ conditional on θ for $\theta\in\{\theta_0,\theta_1\}$. By using the normality and independence of health effects, the physician updates her private belief about θ_1 for patient i according to Bayes' rule:

$$
\begin{aligned}
P_i(\theta_1|x_{il1},...,x_{ilS}) &= \frac{f(x_{il1},...,x_{ilS}|\theta_1)p_i(\theta_1)}{f(x_{il1},...,x_{ilS}|\theta_1)p_i(\theta_1)+f(x_{il1},...,x_{ilS}|\theta_0)p_i(\theta_0)} \\
&= \frac{\prod_{s=1}^{S}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_{ils}-\theta_1)^2}{2\sigma^2}}p_i(\theta_1)}{\prod_{s=1}^{S}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_{ils}-\theta_1)^2}{2\sigma^2}}p_i(\theta_1)+\prod_{s=1}^{S}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_{ils}-\theta_0)^2}{2\sigma^2}}p_i(\theta_0)} \\
&= \frac{1}{1+e^{\sum_{s=1}^{S}\frac{-(x_{ils}-\theta_0)^2+(x_{ils}-\theta_1)^2}{2\sigma^2}}\frac{p_i(\theta_0)}{p_i(\theta_1)}} \\
&= \frac{1}{1+e^{\frac{1}{2\sigma^2}(-2(\theta_1-\theta_0)X_{ilS}+S(\theta_1^2-\theta_0^2))}\frac{p_i(\theta_0)}{p_i(\theta_1)}}. \qquad (10)
\end{aligned}
$$

The posterior[37] depends on signals $x_{il1},...,x_{ilS}$ only through their sum $X_{ilS}=\sum_{s=1}^{S}x_{ils}$, which is also normally distributed given the true quality,

$$X_{ilS}|\theta_i \sim N(S\theta_i, S\sigma^2). \qquad (11)$$

The result generalizes to continuous, normally distributed quality, $\theta_i \sim N(\theta,\sigma^2)$.

A physician learns the true quality through her own experience when the number of signals is large enough. Assume that quality is high.[38] In this case, the joint probability for signals converges to zero more slowly than the corresponding probability for low quality. To see this, examine the denominator in (10) that can be rewritten as

$$1+e^{\frac{1}{2\sigma^2}(-S(\theta_1-\theta_0)^2-2(\theta_1-\theta_0)\sigma\sum_{s=1}^{S}e_{ils})}\frac{p_i(\theta_0)}{p_i(\theta_1)} \qquad (12)$$

when $x_{ils}=\theta_1+\sigma e_{ils}$ for $e_{ils}\sim N(0,1)$. Because the expected value of e_{ils} is zero, the denominator approaches one when the number of signals S increases.

At the patient population level, the weights of the exponential terms increase when the priors of patients, $p_i(\theta_1)$, $\forall i$, decrease. This delays private learning about high quality and increases variation in private posteriors across patients. Note also that for high enough signal realizations i.e., $X_{ilS}>\frac{S((\theta_1)^2-(\theta_0)^2}{2(\theta_1-\theta_0)}$, the private posterior decreases with the uncertainty parameter σ^2, making physicians less likely to continue the drug therapy.

[37]Note that this is a valid probability distribution as the posterior of signals given the true state is restricted between zero and one.

[38]Private learning on low quality is analogous.

The public belief

Next, I consider the social learning of the physician from the past choices of other doctors. After observing the action of physician $-l$, a_{i-lt}, the physician l (and all other physicians except physician $-l$) updates her posterior belief about high quality by using the following Bayes formula:

$$q_{il(t+1)} = \frac{P(a_{i-lt}|h_{it}, \theta_1) q_{ilt}}{P(a_{i-lt}|h_{it}, \theta_1) q_{ilt} + P(a_{i-lt}|h_{it}, \theta_0)(1 - q_{ilt})}. \tag{13}$$

The public posterior belief at time $t+1$ is determined by the (conditional) choice probabilities for high and low qualities and the public belief of physician l at time t. Given that the public beliefs correspond to priors at the beginning of the therapy, $q_{il1} = p_i(\theta_1)$, the final step is to compute the probability of a physician $-l$'s choice, conditional on the patient's prescription history and true quality, $Pr(a_{i-lt}|h_{it}, \theta)$ for $\theta \in \{\theta_0, \theta_1\}$. This is done in two steps.

First assume that physician l observes the physician $-l$'s signals, but not her preference shocks. Let's define a threshold for the difference of private valuations $v_{i-l0t} - v_{i-l1t}$ for which physician $-l$ is indifferent between the continuation and ending of the drug therapy,

$$W_{i-l1t} - W_{i-l0t} = \overline{v}_{i-l0t} - \overline{v}_{i-l1t},$$

where $W_{i-l1t} = E(u_{i-l1t}|I_{i-lt}) - v_{i-l1t}$ is the expected mean utility of the treatment and $W_{i-l0t} = u_{i-l0t} - v_{i-l0t}$ is the corresponding mean utility from the outside good.

Conditional on her signals, the public belief and control variables, a physician's optimal action is to continue the drug therapy if and only if the difference in private valuations is less or equal to the threshold, $v_{il0t} - v_{il1t} \leq \overline{v}_{il0t} - \overline{v}_{il1t}$. If physician l observes that physician $-l$ continued the therapy, she infers that the realization of the difference in private valuations must have been less or equal to this threshold. The larger the threshold, the larger the probability that the drug therapy is chosen.[39]

With the assumption on the distribution of $v_{i-l0t} - v_{i-l1t}$, the conditional choice probability $P(a_{i-lt}|X_{i-lt}, h_{it})$ can be recovered from the thresholds $\overline{v}_{il0t} - \overline{v}_{il1t}$ for all X_{i-lt}. Equivalently, when private valuations are Type 1 extreme value distributed, the conditional probability that physician $-l$ chooses the drug therapy is

[39] See Goeree et al., 2005 for theoretical work with one private signal.

$$
\begin{aligned}
P(a_{i-lt}=1|X_{i-lt},h_{it}) &= P(E(u_{i-l1t}|I_{ilt}) \geq u_{i-l0t}|X_{i-lt},h_{it}) \\
&= \frac{e^{W_{i-l1t}}}{e^{W_{i-l0t}}+e^{W_{i-l1t}}}. \qquad (14)
\end{aligned}
$$

As physician l does not observe the physician $-l$'s private experience, the second step is to compute the choice probability, conditional on the patient's prescription history and quality. The conditional choice probabilities for θ_0 and θ_1 are calculated by using the law of iterated expectations,

$$
P(a_{i-lt}=1|h_{it},\theta) = \int \frac{e^{W_{i-l1t}}}{e^{W_{i-l0t}}+e^{W_{i-l1t}}} dF(X_{i-lt}|\theta) \text{ for } \theta \in \{\theta_0,\theta_1\}. \qquad (15)
$$

where I average out the effect of the sum of signals on the physician's behavior. Without the property that the private belief depends on signals through their sum, the computation of the conditional choice probability would involve S integrals, instead of one. I compute the choice probability numerically by using Simpson's method with 100 uniform grid points.

When physician $-l$ decides to continue the drug therapy of patient i, the public belief of physician l at time $t+1$, $q_{il(t+1)}$, increases from q_{ilt} and hence she becomes more optimistic about quality. To see this, note first that the sum of signals X_{i-lt} is higher under θ_1 than θ_0. The expected utility associated with the continuation of the drug therapy for physician $-l$, $E(u_{i-l1t}|I_{i-lt})$, is increasing with the posterior belief λ_{i-lt}. The higher the sum of signals X_{i-lt} is, the more confident the physician becomes that quality is high i.e., $\frac{\partial \lambda_{i-lt}}{\partial X_{i-lt}} \geq 0$. Therefore, $P(a_{i-lt}=1|X_{i-lt},h_{it})$ in (14) is at least as high when quality is θ_1 than θ_0. Because $F(X_{i-lt}|\theta_1)$ has first-order stochastic dominance over $F(X_{i-lt}|\theta_0)$ for $\theta_1 > \theta_0$, $P(a_{i-lt}=1|h_{it},\theta_1) \geq P(a_{i-lt}=1|h_{it},\theta_0)$. As a result, the public posterior of physician l increases from the previous period i.e., $q_{il(t+1)} \geq q_{ilt}$.

4 The econometric model and identification

In this section, I present the simulated likelihood function of the structural learning model and discuss identification. I use the following data to compute the simulated likelihood function: 1.) the total number of physician visits for patient i, T_i, where the statin therapy of patient i was continued in periods $1,...,T_i-1$ and the outside option was chosen in

period T_i if the patient is non-censored, 2.) the number of patient i's "old" physicians at time t, N_{it}, 3.) an indicator variable if a previously chosen physician l is drawn for patient i again among N_{it} old physicians, d_{ilt}^{old}, 4.) a vector of control variables affecting utilities received from the statin therapy and the outside good, $\mathbf{Z}_{ilt}$, 5.) the censoring indicator, c_i, and 6.) the characteristics of patient i at the beginning of the therapy, $\mathbf{Z}_i^p$, that affect the prior probability.

4.1 The likelihood function

The likelihood contribution of censored patient i contains the following probabilities for each period $t \in \{1, ..., T_i - 1\}$ and physician $l \in \{1, ..., N_{it} + 1\}$ who is drawn for the patient at the beginning of period t: 1.) the probability that physician l is matched to patient i and 2.) the probability that physician l chooses the statin therapy for patient i conditional on the sum of signals and the patient's prescription history, $p_{ilt} = Pr(a_{ilt} = 1 | X_{ilt}, h_{it})$. Because health effects x_{ilt}, preference shocks v_{ilkt}, $k \in \{0, 1\}$, and random coefficients y_i are unobserved by the econometrician, their effects to the likelihood contribution of patient i must be integrated out.

The likelihood contribution of censored patient i is

$$L_i^c \equiv E(\tilde{L}_i^c) = E \prod_{t=1}^{T_i-1} \prod_{l=1}^{N_{it}} \underbrace{[\frac{1-\kappa_i}{N_{it}} p_{il1t}]^{d_{ilt}^{old}}}_{\text{a previously drawn doctor}} \underbrace{[\kappa_i p_{i(N_{it}+1)1t}]^{1-d_{ilt}^{old}}}_{\text{a new doctor}}, \tag{16}$$

which consists of the likelihood contributions of the patient's previously drawn and new doctors. For example, $\frac{1-\kappa_i}{N_{it}}$ is the probability that old physician l is drawn for the patient at the beginning of period t and p_{il1t} is the probability that the treatment of patient i is continued at time t by this physician l.

The data does not contain any information on the identity of the physician who decided to end the therapy. To tackle this problem, I first form the joint probability that a certain physician is drawn for the patient and the same physician chooses to end the drug therapy. Then I sum these joint probabilities over the physicians of the patient to recover the probability that any physician ends the therapy at time T_i.

Formally, the likelihood contribution for the observed data of non-censored patient i is

$$L_i^{nc} = E(\tilde{L}_i^c \cdot [\frac{1-\kappa_i}{N_{iT_i}} \sum_{l=1}^{N_{iT_i}} p_{il0T_i} + \kappa_i p_{i(N_{iT_i}+1)0T_i}]), \tag{17}$$

where $\frac{1-\kappa_i}{N_{iT_i}} p_{il0T_i}$ is the joint probability that an old physician l is drawn and she decides to end the treatment and $\kappa_i p_{i(N_{iT_i}+1)0T_i}$ is the corresponding joint probability for new physician $N_{iT_i}+1$.

Because expectations over signals in the likelihood function contributions are difficult to compute numerically, I use their simulated counterparts $L_i^{c,s}$ and $L_i^{nc,s}$. For example, for non-censored patients,

$$L_i^{nc,s} = \frac{1}{S} \sum_{s=1}^{S} (\tilde{L}_i^{c,s} \cdot [\frac{1-\kappa_i^s}{N_{iT_i}} \sum_{l=1}^{N_{iT_i}} p_{il0T_i}^s + \kappa_i^s p_{i(N_{iT_i}+1)0T_i}^s]), \tag{18}$$

where S is the number of simulations. To compute the simulated likelihood function contribution for each patient, I draw S realization of random coefficients y_i^s governing physicians switching probabilities and $T_i \times S$ realizations of signals and preference shocks to get choice probabilities for each period and patient.[40]

Finally, the simulated log-likelihood function is

$$\log L^s(\theta) = \sum_{i=1}^{N} [c_i \log L_i^{c,s}(\theta) + (1-c_i) \log L_i^{nc,s}(\theta)]. \tag{19}$$

In general, simulation error increases the variance of the he maximum simulated likelihood (MSL) $\hat{\theta}_{MSL}$ estimator compared to the maximum likelihood (ML) estimator. This simulation error disappears asymptotically when the number of simulations increases at a rate higher than $\sqrt{N}$. As the estimation of the model is computationally intensive, I set the number of simulations per patients to ten.[41] Obviously, simulation error may be an issue when the number of simulations is small and therefore estimation results must be interpreted with this caveat. To get appropriate standard errors, I use the standard formula for the simulated estimate of the asymptotic variance which relies on the BHHH estimate for the information matrix. I estimate the model by using the derivative free simplex method (see e.g. Cameron and Trivedi, 2005).

[40] Note that only one physician makes a treatment choice each period and therefore in total $T_i \times S$ simulations of signals and preference shocks are needed for each patient.

[41] For example, Crawford and Shum, 2006, had 30 simulations per patient. I plan to experiment with the number of simulations to see how the results would change.

4.2 Identification

In this section, I briefly consider the structural assumptions of the demand model and the variation in the data that help identify the parameter vector $\Theta = (\theta_0, \theta_1, \sigma^2, \gamma_0, \boldsymbol{\gamma}_1, \boldsymbol{\alpha}, \theta_y, \sigma_y^2, \eta)$. To a large extent, identification relies on similar arguments that have been presented in the previous literature on demand for experience goods (see e.g. Crawford and Shum, 2005).

Market shares at the beginning of the therapy identify the parameters of the prior distribution, γ_0 and $\boldsymbol{\gamma}_1$, because the treatment choice of the physician is then governed by her prior belief. Because the private learning of the physician decreases uncertainty associated with the quality of the medical treatment, choice probabilities at the end of the long-term drug therapy identify parameters for unobserved quality, θ_0 and θ_1. This is particularly true if the patient is in a long-term treatment relationship with his physician. The identification of quality parameters can be also seen from the expected utility of the drug treatment (equation (7)). After fixing the parameters of the prior distribution, γ_0 and $\boldsymbol{\gamma}_1$, and the variance of signals, σ^2, changes in the posterior belief λ_{ilt} with the number of prescriptions identify the quality parameters. Heterogeneity in the choices of physicians both across patients and over time identify the standard deviation of signals. Because quality has two possible values θ_0 and θ_1, it is not possible to separately identify the quality parameters and the risk aversion coefficient, r. I normalize the risk aversion parameter to one which is close to the parameter estimate of Crawford and Shum (2005).[42]

5 Results

In this section, I present results from the estimation of the structural learning model and describe the fit of the model. Because the risk of cardiovascular diseases increases with age and is higher for men than for pre-menopausal women, I allow the prior probability to depend the log of age at $t = 1$ and gender. The prior depends also on an indicator variable for whether the patient was treated by an internal disease specialist at the time of the first physician visit. It is likely that the patient, who used the services of the specialist, is more severely ill and gains more from cholesterol drugs.

I allow the utilities associated with the statin treatment and the outside good to depend on several observed variables. First, I let the utility from therapy continuation to depend

[42] An alternative is to interpret parameters θ and σ^2 relative to risk aversion coefficient r, e.g. $\hat{\theta}_1 = r\theta_1$, where $\hat{\theta}_1$ is the estimated parameter.

on for the average price of statins at time t. I also control for a time trend in months since January 2003 because market level changes, such as advertising, might as well affect the utility from statins. Because the patient's health might deteriorate when he becomes older, I let the utility without cholesterol drugs to depend on age at time t. As cholesterol drugs prevent coronary events in the long-run after the patient's drug therapy has ended, I allow the outside good utility to vary with the number of prescriptions.[43]

Discussion of the results and the fit of the model

Table 5 presents the parameter estimates and their standard errors. The first set contains the key parameters of the model: quality levels θ_0 and θ_1 and the standard deviation of health effects, σ (see 6). Figure 3 presents the conditional and unconditional distributions of signals, $f(x_{ilt}|\theta_0)$, $f(x_{ilt}|\theta_1)$ and $f(x_{ilt})$, for the estimated parameters and the average of priors $p_i(\theta_1)$.

[43] Alternatively, the controls of the outside good could be included in a vector of inside good controls.

Table 5: Parameter estimates for the learning model in the sample of patients

Parameter	Estimate	Std.Err.
Signal (x_{ilt}) parameters		
Low quality (θ_0)	-0.220	0.001
High quality (θ_1)	1.338	0.002
Std. Dev. (σ)	1.049	0.003
Prior parameters		
Constant (γ^0)	-0.003	0.001
log(Age in years at t=1)	0.120	0.000
Gender	0.093	0.001
Visit an internal disease		
specialist at t=1 (1: yes, 0: no)	0.067	0.443
Prior mean and std	0.717	0.012
Physician matching probability		
Random coefficient		
Constant (θ^y)	-0.049	0.001
Std. Dev. (σ_y)	1.057	0.004
Physician switching probability,		
mean and average std	0.491	0.217
Control variables		
Patient's deductible, eur	-0.021	0.000
Time trend in months/10	-0.028	0.000
Outside good controls		
Patient's age/10 years	-0.089	0.000
Number of prescriptions/10	0.107	0.000
Number of observations	22 021	
Number of patients	10 000	
Number of simulations[1]	10	
Simulated log-likelihood function	30 555	

[1] The number of simulations per patient and physician visit.

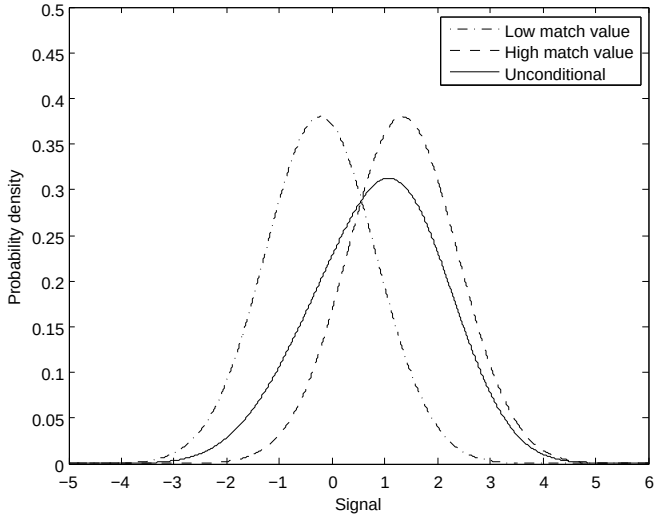

Figure 3: The conditional and mixture probability densities of signals, $f(x_{ilt}|\theta_0)$, $f(x_{ilt}|\theta_1)$ and $f(x_{ilt})$, for estimated parameters and the average prior in the sample of patients

The results demonstrate substantial uncertainty and heterogeneity among patients in the quality and health effects of the statin treatment. The parameter estimate for high quality θ_1 (1.34) is in absolute terms over 6 times higher than the estimate of low quality, θ_0 (-0.22). The variance estimate of signals, σ^2, implies that physicians face significant uncertainty about the health effects of statins even if quality was known. To be more precise, the variance of signals is 5 times higher than the low quality estimate $\hat{\theta}_0$ and 82% of the value of the high quality estimate $\hat{\theta}_1$.

Heterogeneity in health effects implies that information and learning may significantly improve medical decision-making by a physician. Without uncertainty about quality, the incentives of the physician to continue the patient's therapy may be much higher when quality is high rather than low. A high uncertainty in health effects decreases the expected utility of a risk-averse physician, slows down her learning[44] and diminishes her incentives to continue the patient's statin therapy.

The second set of parameters in Table 5 includes estimates for the physician's prior belief that the quality of the statin treatment is high, $p_i(\theta_1)$. As expected, the physician has

[44]This can be seen from the denominator of equation (15) in which iid physician l's shocks e_{ils}, $s \in \{1, ..., S\}$, for patient i get more weight when standard deviation σ increases.

a higher prior probability if her patient is older and male and thus has a higher risk of CVDs compared with other patients. Quite intuitively, the prior belief is higher if the patient has visited an internal disease specialist at the time of the first prescription.

Depending on the characteristics of patients, the prior probability varies across patients from 65% to 75% and has a mean of 72% with a small standard deviation. At the beginning of the therapy, the physician beliefs that quality is more likely to be high than low. Because the average prior belief is fairly uninformative, the posterior belief of the physician λ_{ilt} is mostly determined by her private and public beliefs. This, coupled with a relatively large variance of signals, σ^2, implies that the learning of the physician from her private experience may take some time.

Third, I report the parameters of the random coefficient y_i that affects the probability that the patient is assigned for a new physician, κ_i. The set of parameters for the random coefficient includes the constant, θ_y, and the standard deviation, σ_y. The results suggest that the estimated standard deviation $\hat{\sigma}_y$ (1.06) is much higher than the estimated mean $\hat{\theta}^y$ (-0.05). These findings imply that the probability of getting a new physician varies substantially (0-99%) around its mean (49%). The (average) standard deviation of κ_i is 0.19 that is 32% of the estimated mean of κ_i. Heterogeneity in assignment probabilities across patients can arise for several reasons, including differences between municipalities in their ability to recruit permanent physician labour.

The final set of variables includes control variables affecting utilities associated with the statin therapy and the outside option. The price of statins has a very small, negative effect on the expected utility from the statin treatment. A physician can be insensitive to changes in average prices because a significant part of expenses is covered by the national health insurance. Over time, the expected utility of the physician from the statin treatment decreases. This may reflect changes in advertising by pharmaceutical firms over a product's life cycle and other market level changes. Physicians whose patients are older, and hence have a higher risk of having more severe diseases, are less likely to end the statin therapy as their patients gain less from the outside alternative. The utility associated the outside good increases with the number of prescriptions. This may happen because the statin therapy is likely to have long-term effects on the patient's health even after the statin therapy has ended.

Finally, I consider the model fit by comparing average predicted and observed choice probabilities. For each physician-patient pair, I compute the predicted probability of choosing the statin treatment, conditional on the sum of signals and the patient's prescription history, $P(a_{ilt} = 1|X_{ilt}, h_{it})$. I then compare the corresponding observed choice probabilities

to these predicted probabilities, as presented in Figure 4. The model fits the data relatively well even though it slightly over-predicts the observed average choice probability at the beginning of the treatment and under-predicts after that. At the aggregate level, the model fits the data reasonably well: the average observed probability of choosing the statin therapy is 79% which is close to the predicted probability, 81%. The average predicted probability of getting a new physician is lower (49%) than the corresponding observed probability (60%).

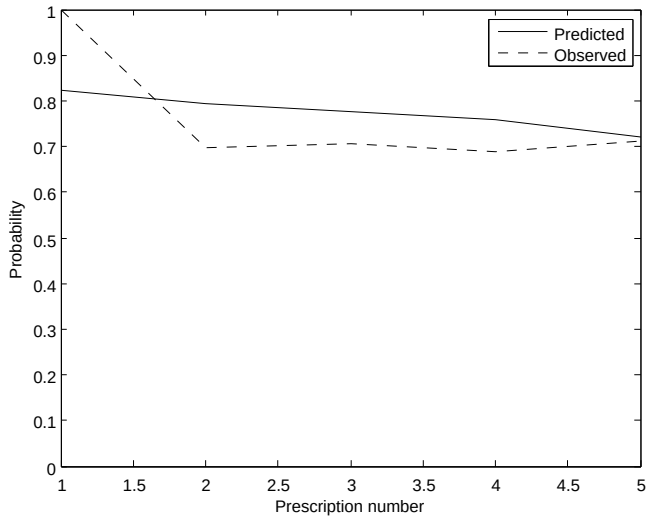

Figure 4: Difference between observed and predicted choice probabilities by the number of prescriptions in the sample, an average over patients, physician visits and simulations

6 Counterfactual experiments

After estimating the parameters, I quantify the roles of private and observational learning in medical decision-making. The main objective is to evaluate to the length of the doctor-patient relationship affects the process of learning and the efficiency of medical decision-making. To be more specific, I evaluate whether the policy promoting continuity of care is preferable to providing information on the past choices of other doctors.

I first investigate what happens if the patient had only one physician. In this case, the physician learns only from her private experience. Next, I investigate whether information

on the past choices of other doctors compensates for the lack of continuity of care. To do this, I compare treatment outcomes and costs in the long-term treatment relationship with the policy where the patient has a different physician every period. A physician has then a one-shot opportunity to investigate the patient to get information on the health effects of cholesterol drugs but she observes the patient's treatment history. To understand the role of peer effects in demand, I study how the behavior of the physician changes if information on the past choices of other doctors was not available. In this experiment, the physician has to rely only on her private experience and the prior belief. Finally, I evaluate the consequences of the policy where the physician does not learn. In this case, the physician decides about the continuation of the patient's therapy without investigating him. I compare the results with the baseline scenario predicted by the estimated model. To perform the policy experiments, I simulate 10 prescription paths for each patient in the observed sample of 10 000 patients used in the estimation of the model.[45]

I begin by describing the development of posterior beliefs over time and dispersion among patients under different policy experiments. I then investigate how treatment adherence, expected utilities and costs change when the length of the treatment relationship and the amount of available information were changed. I measure adherence by the predicted length of the drug therapy and the probability of choosing the statin therapy conditional on the information of the physician, $P(a_{ilt} = 1|I^{\theta}_{ilt})$ (see Dickstein, 2011, for the similar approach).

The speed of learning

Figure 5 describes the development of the average posterior belief over patients, physicians and simulations, conditional on high quality. At the beginning of the therapy, a physician is fairly pessimistic about the effect of the drug treatment on patient health since the average prior for low quality is 28%. Most of the uncertainty regarding to quality vanishes after the first physician visit. At this stage of the therapy, the physician has observed how well the first prescription decreased the patient's cholesterol levels and whether any side effects realized. In the long-term treatment relationship, the physician learns quality fast, by the eighth physician visit. In short-term relationships, physicians become more

[45]When the number of predicted prescriptions is less than the observed one, I use the observed characteristics of patients. Otherwise, I assume that patients come to seek treatment for high cholesterol once a year. The time trend increases by 12 months, the patient's age by a one year and the number of prescriptions by one in period $t+1$ from the previous period t. An exception is the average price of statins at time t which I replace with the average over time, products and patients.

optimistic on quality during the course of the patient's therapy, but learning is slower than in the long-term relationship. The bottom half of Figure 5 presents the standard deviation of posterior beliefs. At the first prescription, variation in posteriors arises because prior beliefs are heterogeneous across patients. Reflecting high variation in health effects, the standard deviation increases to 0.2 at the second prescription. As expected, learning diminishes the variances of the posteriors gradually.

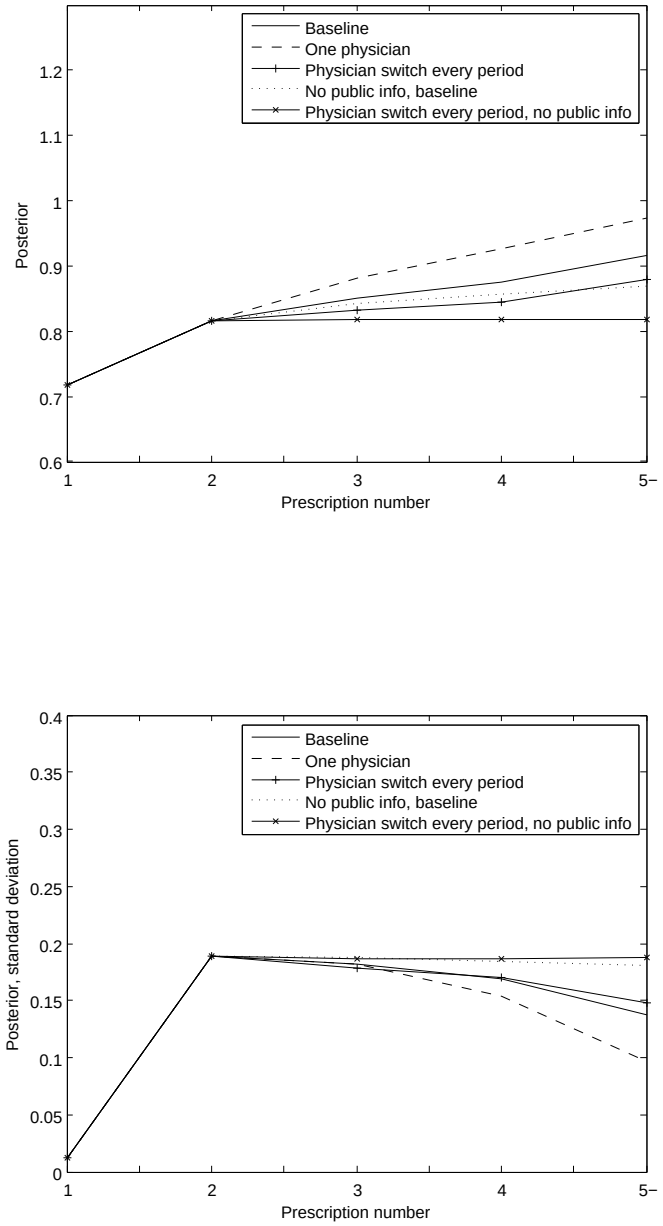

Figure 5: The mean (higher figure) and variance (lower figure) of the posterior belief $\lambda_{ilt} = Pr(\theta_1|I_{ilt})$ given that true quality is high ($\theta_i = \theta_1$) in the sample of patients

The top of Figure 6 illustrates the development of the average posterior when the patient-specific quality is low. In this case, large differences in average posteriors between different

scenarios arise. In the long-term treatment relationship, the physician learns again fast. If physicians change frequently, the average posterior starts to increase after a few prescriptions. Again, the physician becomes more optimistic about quality when other doctors have chosen the drug treatment for the patient previously. The bottom part of Figure 6 shows that heterogeneity in posteriors at the aggregate level is higher among patients when quality is low rather than high. The standard deviation of posteriors are fairly similar in the counterfactual experiments. In particular, a high variation in the posteriors remains also in the permanent treatment relationship, even though the posterior belief is decreasing over time.[46]

[46]Note that the exponential term in equation (10) is $e^{S(\theta_0-\theta_1)^2-2(\theta_1-\theta_0)\sigma\sum_{s=1}^{S}e_{ils}}$ if $\theta_i = \theta_0$. When $2(\theta_1-\theta_0)\sigma\sum_{s=1}^{S}e_{ils}$ is high relative to constant term $S(\theta_0-\theta_1)^2$, there can be much variation in the posterior beliefs of physicians among patients.

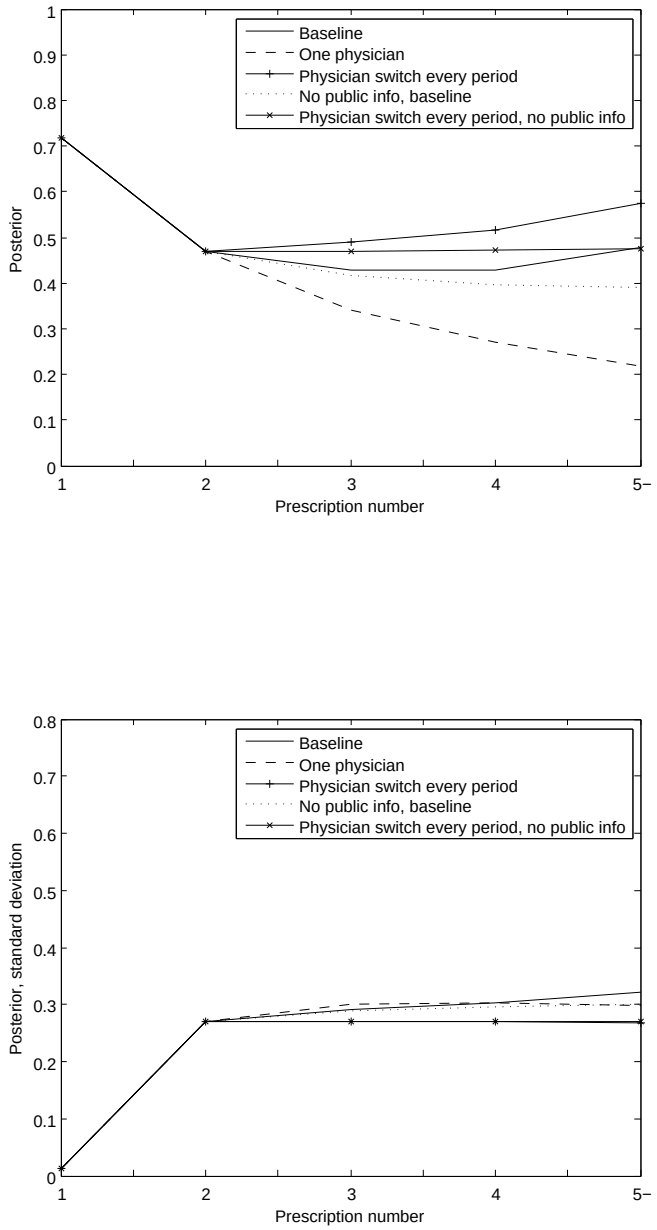

Figure 6: The mean (higher figure) and variance (lower figure) of the posterior belief $\lambda_{ilt} = Pr(\theta_1|I_{ilt})$ given that true quality is low ($\theta_i = \theta_0$) in the sample of patients

Overall, the results suggest that the long-term doctor-patient relationship promotes the process of learning about quality. The physician becomes optimistic about quality when

she observes the past choices of other doctors. When quality is high, information on the prescription history improves learning, but not as efficiently as the long-term relationship. When quality is low, such information slows down learning. These results have implications on prescriptions, costs and efficiency.

The length of the doctor-patient relationship

I first examine how the long-term doctor-patient relationship affects outcomes and costs. Table 6 presents averages for the expected utility, the adherence of the treatment and the total costs, conditional on quality. The results suggest that continuity of care promotes learning and improves medical decision-making by a physician. Consider first the patient with high quality of the match with cholesterol drugs. In this case, the long-term physician-patient relationship leads to the highest expected utility and the treatment adherence among evaluated experiments. Still, the treatment adherence and the total costs increase only slightly from the estimated benchmark. When quality is low, I find that continuity of care decreases treatment length by 5% and the total costs of the drug therapy by 5% compared with the estimated benchmark. This is so, since the physician learns fast that the treatment does not suit well for the patient.

Table 6: Counterfactual simulations in the sample of patients

Outcome variable	Baseline[2]	One physician[3]	Physician change every period	No public info[4]	No public info, physician change every period[4]	No learning[5]
High quality θ_1:						
Expected utility	-2.612	-2.594	-2.624	-2.618	-2.632	-2.704
Expected utility, ≥ 5 prescriptions	-2.791	-2.765	-2.808	-2.802	-2.824	-2.873
Treatment length	4.327	4.420	4.420	4.272	4.199	4.013
Probability of statin therapy	0.770	0.774	0.766	0.766	0.763	0.751
Total cost/100 eur	2.883	2.945	2.844	2.846	2.797	2.674
Low quality θ_0:						
Expected utility	-2.815	-2.844	-2.798	-2.823	-2.808	-2.692
Expected utility, ≥ 5 prescriptions	-3.034	-3.104	-3.006	-3.056	-3.031	-2.865
Treatment length	3.406	3.243	3.503	3.361	3.439	3.967
Probability of of statin therapy	0.706	0.692	0.715	0.703	0.710	0.748
Total cost/100 eur	2.270	2.160	2.335	2.240	2.292	2.644

[1] These values are calculated by using the observed sample of 10 000 patients and 10 simulated prescription sequences per patient.

[2] The baseline scenario is predicted by the model estimates.

[3] In this experiment, the physician-patient relationship is permanent.

[4] Public information on the previous treatment continuation choices of other physicians is not available.

[5] The physician does not learn and hence the posterior belief equals to the prior belief, i.e. $\lambda_{ilt} = p_i(\theta_1)$.

I next investigate the consequences of the policy where a new physician treats the patient in every period (Table 6). When the physician does not have much own experience of the patient, she relies more on the past choices of other doctors. Consider first the patient with high quality of the match with statins. In this case, continuity of care does not much improve drug choices or change treatment outcomes compared to the policy where treatment relationships are short-term but the prescription history is observed.

Consider then the patient with the low quality of the match in Table 6. In this case, the length of the treatment relationship has more pronounced effect on treatment outcomes and costs. This happens because social learning increases the optimism of the physician about the quality and can lead to over-prescribing. The results show that the policy with the short-term relationship increases the adherence by 3% and the total costs by 8% from the experiment with continuity of care. Table 6 demonstrates that the physician would be

slightly better-off, in terms of efficiency, without information on the prescription history when physicians change frequently. Specifically, when treatment relationships are short-term, providing information on the past choices of other doctors increases the adherence by 1% and the total costs by 2% from the policy without such information. Again in terms of efficiency, even worse outcomes arise if learning is not possible.

The results have several policy implications. Continuity of care helps the physician to find out sooner the health effects of the drug treatment. This reduces the costs of uncertainty and improves her medical decision-making, as suggested by the existing reduced-form literature (Weiss and Blustein, 1996, Scott, 2000, King et al., 2008). The second conclusion is that information on the patient's prescription history does not compensate for the lack of the long-term treatment relationship. When the treatment suits well for the patient, prescription records promote learning, but not as efficiently as continuity of care. If a physician does not have much own experience, treatment patterns based on the observed medication history of the patient may hinder learning and lead to over-prescribing for a fraction of patients.

7 Conclusions

I quantified the roles of private experience and the past choices of other doctors in pharmaceutical demand. I constructed a structural model of demand for pharmaceuticals under uncertainty about the quality of the match between the patient and the drug treatment. I analyzed whether continuity of care is preferable to providing information on the past behavior of other doctors through patient records.

Using rich data from the market for cholesterol drugs, I found that prescriptions are highly responsive to the length of the doctor-patient relationship. I illustrated that the number of interactions between the physician and the patient have important implications on the efficiency of drug choices. My analysis suggested that treatment patterns relying heavily on the past choices of other doctors may lead to over-prescribing for a fraction of patients. I also showed that the long-term treatment relationship can limit over-prescribing and improve medical decision-making.

The structural model can be extended to allow the other important features of the pharmaceutical market. The first extension is to make physicians forward-looking in their decision-making, creating incentives for experimentation to get more information. Second, the model can be broadened to incorporate several inside goods. The framework

can be also applied in other experience good markets, such as financial markets, where traders are investing in assets with uncertain returns.

A typical example of a patient story for one dispensary visit:

The reason of entry
A patient comes with the referral of physician X due to atrial fibrillation

At issue a 65 years old retired gymnastics teacher. In an anamnesis 2003 acute coronary thrombosis, angioplasty RCA. Discovered then also a decreasing diverticulum of an aorta ad 50mm, controls in fall. In the Doppler-ultrasound-research of neck veins in 2005 was discovered in left arteria carotis interna stenosis less 50%. Discovered year 2007 COPD. The patient smoked over 30 years, quit 6 years ago. In a tolerance test 8/07, no coronary ischaemia.

The patient has visited in the health center of X due to dizziness. Discovered elevated blood pressure, irregular beat. Hear enzymes and other laboratory values normal, pro-BNP over 500. Patient's medication at this moment Pravachol 20mg x 1, Linatil 20mg x 1, Carvedilol 12.5mg x 2. Started Marevan due to atrial fibrillation, aiming to do cardioversion.

Today taken INR, only 1.3. Hence cardioversion cannot be done now. Pulse also fairly fast 80-90/min, RR-level 180-170/110-100. Carvedilol ad raised 25mg + 12.5mg. INR-controls will continue in the side of outpatient treatment. Phone contact after a month.

Physician X

References

[1] Ackerberg, D. (2003) Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination. *International Economic Review,* Vol. 44, pp. 1007-1040.

[2] Adams, S. (2011) Millions Taking Statins 'Needlessly'. *The Telegraph.*

[3] Anderson, L. – Holt, C. (1997) Information Cascades in the Laboratory. *American Economic Review*, Vol. 87, pp. 847-862.

[4] Athey, S. – Stern, S. (2002) The Impact of Information Technology on Emergency health care Outcomes. *RAND Journal of Economics*, Vol. 33, pp. 399-432.

[5] Avery, C. – Zemsky, P. (1998) Multidimensional Uncertainty and Herd Behavior in Financial Markets. *American Economic Review*, Vol. 88, pp. 724-748.

[6] BBC (2011) Questions over statin prescribing.

[7] Bikhchandani, S. – Hirshleifer, D. – Welch, I. (1992) Theory of Fads, Fashion, Custom, and Structural Change as Informational Cascades. *Journal of Political Economy*, Vol. 100, pp. 992-1026.

[8] Cameron, A. – Trivedi, P. (2005) Microeconometrics - Methods and Applications. *Cambridge University Press*, New York, USA.

[9] Celen, B. – Kariv, S. (2004) Distinguishing Informational Cascades from Herd behavior in the Laboratory. *American Economic Review*, Vol. 94, pp. 484-498.

[10] Chan, T. – Hamilton, B. (2006) Learning, Private Information, and the Economic Evaluation of Randomized Experiments. *Journal of Political Economy*, Vol. 114, pp. 997-1040.

[11] Chamley, C. (2004) Rational Herds - Economic Models of Social Learning. *Cambridge University Press*, the United States of America.

[12] Chaudhry, B. – Wang, J. – Wu, S. – Maglione, M. – Mojica, W. – Roth, E. – Morton, S. – Shekelle, P. (2006) Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care. *Annals of Internal Medicine,* Vol. 144, pp. 742-752.

[13] Chernew, M. – Gowrisankaran, G. – Scanlon, D. (2008) Learning and the Value of Information: Evidence from Health Plan Report Cards. Journal of Econometrics, Vol. 144, pp. 156-174.

[14] Chevalier, J. – Ellison, G. (1999) Career Concerns of Mutual Fund Managers. *Quarterly Journal of Economics*, Vol. 114, pp. 389-432.

[15] Ching, A. (2009) A Dynamic Oligopoly Structural Model for the Prescription Drug Market after Patent Expiration. Forthcoming in *International Economic Review*.

[16] Ching, A. (2010a) A Dynamic Oligopoly Structural Model for the Prescription Drug Market After Patent Expiration. *International Economic Review*, Vol. 51, pp. 1175-1207.

[17] Ching, A. (2010b) Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs After Patent Expiration. *International Journal of Industrial Organization*, Vol. 28, pp. 619-638.

[18] Cipriani, M. – Guarino, A. (2012) Estimating a Structural Model of Herd behavior in Financial Markets. Forthcoming in *American Economic Review*.

[19] Coscelli, A. – Shum, M. (2004) An Empirical Model of Learning and Patient Spillovers in New Drug Entry. *Journal of Econometrics*, Vol. 122, pp. 213-246.

[20] Crawford, G. S. – Shum, M. (2005) Uncertainty and Learning in Pharmaceutical Demand. *Econometrica*, Vol. 73, pp. 1137-1173.

[21] Crawford, V. – Sobel, J. (1982) Strategic Information-Transmission. *Econometrica*, Vol. 50, pp. 1431-1451.

[22] Dickstein, M. (2011) Efficient Provision of Experience Goods: Evidence from Antidepressant Choice. Job Market Paper, Harvard University.

[23] Drehmann, M. – Oechssler, J. – Roider, A. (2005) Herding and Contrarian Behavior in Financial Markets: An Internet Experiment. *American Economic Review*, 95, pp. 1403-1426.

[24] Epstein, R. M. (1999) Mindful Practice. *Journal of American Medical Association*, Vol. 282, pp. 833-839.

[25] Finnish Medical Association (FMA) (2006a) Physician Situation in Health Centers 2006 (In Finnish: Terveyskeskusten lääkäritilanne, 2006).

[26] Finnish Medical Association (FMA) (2006b) Ennalta Ehkäisevä Terveydenhuolto Vaarantuu: Omalääkärijärjestelmä Uhkaa Romuttua. *FMA Newsletter*, 2.5.2005.

[27] Finnish Medical Association (FMA) (2006c) Physician Survey. www.laakariliitto.fi/files/laakarikysely2006.pdf (In Finnish)

[28] Ford, I. – Murray, H. – Packard, C. – Shepherd, J. – Macfarlane, P. – Cobbe, S. (2007) Long-Term Follow-up of the West of Scotland Coronary Prevention Study. *New England Journal of Medicine*, Vol. 357, pp. 1477-1486.

[29] Franklin, C. (2011) America's Epidemic of Over-Prescribing. *Guardian.*

[30] Goeree, J. – Palfrey, T. – Rogers, B. (2005) Social Learning with Private and Common Values. *Economic Theory, Research Articles*, Vol. 28, pp. 245-264.

[31] Golman, G. (1990) The Tacit Dimension of Clinical Judgment. *The Yale Journal of Biology and Medicine*, Vol. 63, pp. 47-61.

[32] Graham, J. (1999) Herding among Investment Newsletters: Theory and Evidence. *The Journal of Finance*, Vol. 54, pp. 237-268.

[33] Guthrie, B. – Saultz, J. W. – Freeman, G. K. – Haggerty, J. L. (2008) Continuity of Care Matters. *BMJ,* Vol. 337, pp. -.

[34] Haggerty, J. – Reid, R. – Freeman, G. – Starfield, B. – Adair, C. – McKendry, R. (2003) Continuity of Care: a Multidisciplinary Review. *BMJ*, Vol. 327, pp. 1219-1221

[35] Hillestad, R. – Bigelow, J. – Bower, A. – Girosi, F. – Meili, R. – Scoville, R. – Taylor, R. (2005) Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, And Costs. *Health Affairs*, Vol. 24, pp. 1103-1117.

[36] Hong, H. – Kubik, J. – Solomon, A. (2000) Security Analysts' Career Concerns and Herding of Earnings Forecasts. *RAND Journal of Economics*, Vol. 31, pp. 121-144.

[37] Joelving, F. (2011) Statin Benefits Questioned in Low-Risk Patients. *Reuters Health.*

[38] Johnston, D. (1999) Lifestyle Changes After a Myocardial Infarction. Editorial in *Heart*, Vol. 82, pp. 543-544.

[39] Jousilahti, P. (2004) Current Care for Dyslipidemias is Renewed (in Finnish: Dyslipidemioiden Käypä hoito uudistuu). *Kansanterveys-lehti*, National Institute for Health and Welfare.

[40] Kim, M. (2010) Physician Learning and New Drug Diffusion. Job Market Paper, University of Chicago.

[41] King, M. – Jones, L. – Richardson, A. – Murad, S. – Irving, A. – Aslett, H. – Ramsay, A. – Coelho, H. – Andreou, P. – Tookman, A. – Mason, C. – Nazareth, I. (2008) The Relationship Between Patients' Experiences of Continuity of Cancer Care and Health Outcomes: a Mixed Methods Study. *British Journal of Cancer*, pp. 529-536.

[42] Klaukka, T., Helin-Salmivaara, A., Huupponen, R., Idanpää-Heikkilä, J. (2011) Simvastatiinin asema vahvistui vuonna 2005. *Finnish Medical Journal*, 18/2006, Vol. 61, pp. 1998-2001.

[43] Knight, B. – Schiff, N. (2010) Momentum and Social Learning in Presidential Primaries. *Journal of Political Economy*, Vol. 118, pp. 1110-1150.

[44] Ministry of Social Affairs and Health in Finland (2005) National Guide for Pharmacotherapy in Social and health care. *Handbooks of the Ministry of Social Affairs and Health, Finland*, 32, pp. 1-117.

[45] Nair, H. S. – Manchanda, P. – Bhatia, T. (2010) Asymmetric Social Interactions in Physician Prescription Behavior: The Role of Opinion Leaders. *Journal of Marketing Research*, Vol. 47, pp. 883-895.

[46] National Audit Office of Finland (2011) Sosiaali- ja Terveydenhuollon Valtakunnallisten IT- hankkeiden Toteuttaminen. *Valtiontalouden tarkastusviraston tuloksellisuustarkastuskertomukset*, 217/2011, pp. 1-274.

[47] National Center for Health Statistics (2011) Health, United States, 2011: With Special Feature on Socioeconomic Status and Health. Hyattsville, MD. 2012.

[48] National Health Service (2013) Cholesterol-Lowering Medicines, Statins - Side Effects.

[49] National Institute for Health and Clinical Excellence (2006) Statins for the Prevention of Cardiovascular Events. *NICE Technology Appraisal*, Vol. 94, pp. 1-45.

[50] National Institute of Health and Welfare (2011) Cardiovascular Diseases.

[51] Novack DH, Suchman AL, Clark W, Epstein RM, Najberg E, Kaplan C.(1997) Calibrating the Physician. Personal Awareness and Effective Patient Care. Working Group on Promoting Physician Personal Awareness, American Academy on Physician and Patient. *The Journal of American Medical Association*, Vol. 278, pp. 502-509.

[52] Olszewski, W. (2004) Informal Communication. *Journal of Economic Theory*, Vol. 117, pp. 180-200

[53] Park, A., Sabourian, H. (2010) Herding and Contrarian Behavior in Financial Markets. Forthcoming in *Econometrica.*

[54] Pisacano, N. J. (2007) History of the Specialty. *American Board of Family Medicine.*

[55] Scott, A. (2000) Economics of General Practice. Chapter 22 in *Handbook of Health Economics*, Volume 1, Edited by A.J. Culyer and J.P. Newhouse, Elsevier Science B.V.

[56] Scott-Morton, F. (1999) Entry Decisions in the Generic Drug Industry. *RAND Journal of Economics*, Vol. 30, 421-440.

[57] Smith, L. – Sorensen, P. (2000) Pathological Outcomes of Observational Learning. *Econometrica*, Vol. 68, No. 2, pp. 371-398.

[58] Sturmberg, JP. (2001) Continuity of Care Revisited. *Family Practice*, Vol. 18, pp. 461-462.

[59] U.S. Food and Drug Administration (FDA) (2010) Controlling Cholesterol with Statins. *FDA Consumer Health Information*, pp. 1-2.

[60] Vives, X. (2008) Information and Learning in Markets: The Impact of Market Microstructure. *Princeton University Press,* the United States of America.

[61] Weiss, L.J. – Blustein, J. (1996) Faithful Patients: The Effect of Long-Term Physician-Patient Relationships on the Costs and Use of Health Care by Older Americans. *American Journal of Public Health*, Vol. 86, pp. 1742-1747.

[62] Winblad, I. – Reponen, J., Hämäläinen, P. – Kangas, M. (2006) Informaatio- ja kommunikaatioteknologian käyttö. Tilanne ja kehityksen suunta. Stakes, Reports, 7/2006, pp. 1-83.

[63] Wold heart federation (2011) Cardiovascular disease risk factors. http://www.world-heart-federation.org/cardiovascular-health/cardiovascular-disease-risk-factors/

[64] Zhang, J. (2010) The Sound of Silence: Observational Learning in the U.S. Kidney Market. *Marketing Science*, Vol. 29, pp. 315-335.