Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Structural Model of Correlated Learning and Late-Mover Advantages: The Case of Statins

Andrew T. Ching, Hyunwoo Lim

To cite this article:
Andrew T. Ching, Hyunwoo Lim (2020) A Structural Model of Correlated Learning and Late-Mover Advantages: The Case of Statins. Management Science 66(3):1095-1123. https://doi.org/10.1287/mnsc.2018.3221

http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 66, No. 3, March 2020, pp. 1095–1123
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

A Structural Model of Correlated Learning and Late-Mover Advantages: The Case of Statins

Andrew T. Ching,[a] Hyunwoo Lim[b]

[a] Carey Business School, Johns Hopkins University, Baltimore, Maryland 21202; [b] School of Administrative Studies, York University, Toronto, Ontario M9N 2H9, Canada

Contact: andrew.ching@jhu.edu, http://orcid.org/0000-0002-0907-284X (ATC); hlim@yorku.ca, http://orcid.org/0000-0001-8664-5405 (HL)

Received: October 20, 2016
Revised: April 18, 2018
Accepted: August 11, 2018
Published Online in Articles in Advance: January 6, 2020

https://doi.org/10.1287/mnsc.2018.3221

Abstract. We propose a structural model of correlated learning with indirect inference to explain late-mover advantages. Our model focuses on a class of products with the following two features: (i) products that build on a common fundamental technology (e.g., computer processor, car, smartphone, etc.) and (ii) that consumers can observe some product attributes of a product (e.g., CPU clock speed, horsepower of a car engine, screen size of a smartphone, etc.), but when making their purchase decisions, consumers are not sure how efficiently the product can translate its observed attributes to performing tasks that they care about. For products that base on a similar technology, it is plausible that consumers use the information signals of one product's technological efficiency to help them update their belief about another product's technological efficiency within the same product category. As a result, a late entrant could benefit from the information spillover generated by an early entrant. We apply our framework to the statin market in Canada, where drugs rely on a similar mechanism to reduce the cholesterol level. In our model, patients/doctors can observe a statin's efficacy in reducing the cholesterol level, but they are uncertain about how effectively it can convert its cholesterol-reducing ability to reducing heart disease risks. Our estimation results show that the combination of correlated learning and informative and persuasive detailing explain the success of the two late entrants in the statin market: Lipitor and Crestor.

History: Accepted by Matthew Shum, marketing.

Keywords: correlated learning • late-mover advantages • clinical trials • detailing • efficiency ratio

1. Introduction

It is commonly believed that being the first to the market can often provide a product with significant competitive advantages. However, there are many incidents in which latecomers take over the first entrant. Golder and Tellis (1993) document that the failure rate of the first entrant is indeed quite high; it is around 47% out of 500 brands in 50 consumer product categories. Examples include disposable diapers, personal computers, chain restaurants, etc. This surprising fact suggests that being late to the market may also have some advantages. Previous research has proposed several theoretical explanations for late-mover advantages (e.g., free ride on information spillover in R&D, fewer constraints on innovation, etc.). But, up to this point, the literature has not developed and estimated any empirical structural model to quantify the significance of late-mover advantages. To our knowledge, we are the first to (a) propose a model that makes use of information spillover on the demand side (correlated learning) to explain late-mover advantages and (b) estimate the structural parameters of a model with late-mover advantages to quantify its economic significance.

Our model focuses on a class of products with the following two features: (i) products that build on a common fundamental technology (e.g., computer processor, car, smartphone, etc.) and (ii) that consumers can observe some product attributes of a product (e.g., CPU clock speed, horsepower of a car engine, screen size of a smartphone, etc.), but when making their purchase decisions, consumers are not sure how efficiently the product can translate its observed attributes to performing tasks that they care about (i.e., they are uncertain about its technical efficiency). For instance, even though consumers know the CPU clock speed of a computer, they may not know how long it would take this computer to process a certain amount of data. For products that base on a similar technology, it is plausible that consumers use the information signals of one product's technological efficiency to help them update their belief about another product's technological efficiency within the same product category. As a result, a late entrant could benefit from the information spillover generated by an early entrant. Such information spillover (correlated learning) may allow consumers to indirectly infer that a late entrant's product is superior if

some of its observed attributes are improved even without any direct evidence about how well its product can actually perform in completing tasks. In this paper, we develop a structural model of correlated learning and indirect inference to capture such late-mover advantages.

We apply our model to the statin market (a class of anticholesterol drugs) in Canada. Statin is the most popular class of anticholesterol drugs, and most patients take statins to reduce their bad cholesterol (LDL), hoping that it will reduce their heart disease risks. However, before a clinical study on reducing heart disease risks becomes available for a drug, sales representatives can only make a direct claim on its efficacy in lowering the bad cholesterol. Although a positive correlation between bad cholesterol and heart disease risks has been found in medical research, a drug that can reduce the bad cholesterol level effectively does not necessarily mean it can reduce heart disease risks.[1] This is because it might have some unknown side effects that could raise heart disease risks and counter its benefits of reducing the bad cholesterol. To claim that their drugs are effective in reducing heart disease risks, statin manufacturers have invested in postmarketing clinical trials to provide such direct evidence. However, postmarketing clinical trials on reducing heart disease risks often take several years to complete, require many participants, and carry large financial costs.

When Lipitor (atorvastatin) entered the market in 1997, there was no clinical evidence to show its efficacy in reducing heart disease risks until Q2 2003. Yet it was able to expand its market volume steadily and rapidly since its inception. Lipitor's success is puzzling because prior to its entry, three incumbent statins had already established clinical evidence about their ability in reducing heart disease risks. Assuming that physicians' ultimate goal is to lower patients' chances of having heart attacks or strokes, one would expect they prefer the older statins with direct clinical evidence of such efficacy.

Our model provides an explanation for this puzzle. Because statins use the same chemical mechanism to reduce the bad cholesterol (Zhou et al. 2006) (in other words, they use the same fundamental technology), it is plausible that physicians believe that all statins share a similar ability to convert *reduction in the bad cholesterol* to *reduction in heart disease risks*. Therefore, when seeing clinical evidence on this conversion ability from older statins, physicians may update their beliefs about Lipitor's conversion ability as well. In addition, because Lipitor is also more effective in lowering the bad cholesterol, physicians may then infer that Lipitor is more effective in reducing heart disease risks compared with its competitors even though there is not yet direct clinical evidence to prove this.

To quantify the importance of this information spillover story, we develop and estimate a structural demand model of correlated learning and indirect inference. We introduce a concept, "efficiency ratio," which measures how effective a drug can convert reduction in the bad cholesterol to reduction in heart disease risks. Although landmark clinical trials provide information about statins' efficiency ratios, physicians and patients might not actively search for clinical trial results and need to learn about this scientific information through different channels.[2] We, therefore, allow detailing[3] and news coverage (hereafter, we refer to it as "publicity") to play an informative role in delivering information embedded in clinical trials to physicians and patients. A pharmaceutical representative may inform or remind physicians of the drug's efficacies. Alternatively, a physician/patient may learn about a drug's efficacy or the release of an important clinical trial from news media (e.g., Ching et al. 2016).[4]

In addition to its informative role, it is well known that detailing plays a persuasive role in the prescription drug market (e.g., Leffler 1981, Narayanan et al. 2005, and Ching and Ishihara 2012). Hence, it is quite possible that the persuasive role of detailing is also responsible for the success of Lipitor. Our structural model takes this into account by allowing a persuasive detailing goodwill stock to enter physicians' utility function.

Our structural estimation results show that both informative and persuasive detailing are important. In fact, we find that the persuasive role of detailing is largely responsible for the rapid growth of Lipitor and Crestor. Despite this, we still find that Lipitor and Crestor benefit from correlated learning in a fairly important way. We find evidence that detailing and publicity are responsible for bringing the information about landmark clinical trials to physicians. Interestingly, we also find evidence that a firm uses its detailing to inform physicians about its rivals' landmark clinical trial outcomes; this indicates firms are aware of the existence of correlated learning/information spillover because, by informing physicians of their rivals' clinical findings, they could also benefit from it. Moreover, we find that both detailing and publicity play important roles in getting patients to adopt statins in the first place.

After estimating our model, we use it to conduct several counterfactual experiments. Because our model incorporates consumers' learning about clinical trials, the results can be used to forecast the returns of landmark clinical trials (measured by how much demand they can generate), which are usually sponsored by pharmaceutical firms. Such results are useful for managers who need to decide which clinical trials to fund. Note that Lipitor obtained its own landmark clinical trial results six years after its entry in 1997. How much of an impact did these landmark clinical trials have on Lipitor's sales? Were the landmark clinical trials worth the investment for Pfizer given that Lipitor was able to (imperfectly) free ride on clinical trials conducted by its

rivals? Our first counterfactual experiment finds that, without its own landmark trials, Lipitor's demand could have dropped by about 150,000 prescriptions in Q4 2004. We argue that, by projecting the impact in Canada to the world market, it is likely that Lipitor's investment in its landmark clinical trials is worthwhile.

Theoretically, it is possible that the landmark trials done by the early entrants could hurt themselves. This is because late entrants can benefit from it, which, in turn, increases the competitive pressure on the early entrants. Hence, it is not clear if early entrants should conduct their landmark clinical trials from the profit-maximizing viewpoint. In the second counterfactual experiment, we try to shed light on this question by removing the landmark trials for two early entrants: Mevacor and Pravachol. The results show that the changes in their demand would be very small. This suggests that the landmark trial investment done by these two early entrants may not have paid off.

To quantify the late-mover advantages, we conduct two more counterfactual experiments by removing correlated learning/information spillover. In other words, we only allow patients/physicians to learn about a drug's efficiency ratio from its own clinical trials. These experiments show that new and switching patients' demand for late entrants would drop quite significantly. The result suggests that, after controlling for persuasive detailing, correlated learning still plays an important role for the early success of Lipitor and Crestor.

The rest of this paper is organized as follows. Section 2 reviews the previous literature. Section 3 describes background information, including the market for statins. Section 4 discusses our data set. Section 5 describes our structural model. Section 6 presents the estimation results. Section 7 is the conclusion.

2. Literature Review

The existing literature on late-mover advantages can be divided into two streams: (i) incumbent inertia and technological discontinuities and (ii) knowledge spillover in R&D. The first stream is pioneered by Schumpeter (1961), known as creative destruction, and is later extended by Arrow (1962), Ghemawat (1986), Reinganum (1983), and Tang (1988).

Our research is more related to the second stream, in which Spence (1984), Ghemawat and Spence (1985), Lieberman (1987), and Guasch and Weiss (1980) argue that a late entrant could free ride on the earlier entrants by hiring the workers trained by them and, hence, incurring lower total cost of production. Gallant et al. (2018) develop and estimate a structural model of firms' entry decisions in which a firm's prior experience could lower its entry cost. However, they do not model information spillover across firms. Instead, they model knowledge spillover within a firm's own product portfolio. Our model complements this literature by focusing on information spillover on the demand side instead of the cost side.

Our paper also belongs to the correlated learning literature (e.g., Erdem 1998, Erdem and Sun 2002, and Janakiraman et al. 2009). But previous research focuses on learning about *one* product attribute, which directly enters consumers' utility function. In such a framework, it is difficult for correlated learning alone to explain why late entrants can succeed so quickly as we see in Lipitor because of imperfect information spillover. Instead of assuming consumers derive utility directly from all product attributes, our model assumes that a product uses a "production" technology to convert some "input" attributes to an "output" that consumers care about and derive utility from. Consumers observe the input attributes of the product, but they do not observe the output when they make their purchase decisions, and they are uncertain about the production efficiency for this product. Hence, they need to form expectation about its output based on their belief about its production function parameters and the observed input attributes. We assume correlated learning happens at the production-function level.

Although some papers have developed learning models to study the pharmaceutical market (e.g., Crawford and Shum 2005; Narayanan et al. 2005; Chintagunta et al. 2009; Ching 2010a, b; and Chan et al. 2013), the vast majority of them do not model clinical evidence as a source of quality signals at all. An exception is the study by Ching and Ishihara (2010), who use qualitative information on comparison clinical studies (which say whether drug A is better than drug B) as signals about product attributes, which directly enter consumers' utility. In this study, we treat clinical trial results more seriously than previous research. We make use of the quantitative information in clinical trials instead of just the qualitative information as in Ching and Ishihara (2010). More importantly, we treat the information reported in landmark clinical trials as "observable" signals not only to the agents in the model, but also to researchers. This greatly simplifies the estimation procedure by avoiding the integration of unobserved signals when forming the likelihood, a computationally intensive procedure that is typically needed in previous works of estimating structural learning models (Ching et al. 2013). In addition, the clinical trial data also helps identify the parameters of the model as we discuss later. We should also note that Ching and Ishihara (2010) ignore the possibility of correlated learning. However, many me-too drugs belonging to the same therapeutic class often share a similar technology or mechanism to tackle an illness. It is plausible that our proposed framework can be applied to other classes of drugs and other products as well (e.g., smartphone, computer, electric vehicles, etc.).

Our model is also related to Chan et al. (2013), Liu et al. (2017), and Shapiro (2018). Chan et al. (2013) propose a learning model incorporating multidimensional attributes and positive state dependence. They investigate physicians' learning on the effectiveness and side effects of drugs separately through patients' reported reasons for switching in the erectile dysfunction category. Similar to their study, we develop a multidimensional model. However, the sources of identification are very different. They rely on physician-level survey data, whereas we rely on the content of clinical trials. Liu et al. (2017) model a different type of detailing spillover effect when drugs are sold as a bundle (e.g., HIV combination therapy). However, they do not consider consumer uncertainty and learning about drugs' efficacies. Shapiro (2018) estimates the market expansion effect of direct-to-consumer advertising, which triggers patients to see doctors, who then decide which drug to prescribe. Our model also has such a market expansion effect because of informative detailing and publicity. But on top of it, we explicitly model how information signals about one drug can improve doctors'/patients' prior belief about other drugs via correlated learning, a mechanism that is not studied in Chan et al. (2013), Liu et al. (2017), or Shapiro (2018).[5]

3. Background

There are two main types of cholesterol: LDL ("bad" cholesterol) and HDL ("good" cholesterol). The medical literature has shown that high LDL is a risk factor for heart disease. Hereafter, we follow the tradition and use the term "cholesterol" for LDL. Although the main purpose of statins is to reduce heart disease risks, a drug company cannot make the direct claim that its statin can achieve this goal until it obtains direct evidence from a clinical trial to support the claim. This is because the public health agency is worried that some unknown side effects of the drug could counter its benefits of lowering the cholesterol level. However, the information on the effectiveness of a statin in reducing heart disease risks is usually unavailable when a statin is marketed because it takes a few years to obtain such direct evidence. To obtain direct scientific evidence, pharmaceutical firms invest in very expensive post-marketing clinical trials, which are called landmark clinical trials. Each landmark clinical trial reports a drug's efficacies in (i) reducing heart disease risks and (ii) lowering LDL. By examining these two reported efficacies, physicians can learn about the efficiency ratio of a statin. Landmark clinical trials are important sources of information about statins' effectiveness in reducing heart disease risks. In particular, because heart attack/stroke is a very rare event, it is difficult for a physician to tell how well a drug can reduce heart disease risks by only observing the physician's own patients' feedback. Therefore, in our model, we assume landmark clinical trials are the only source of quality signals about drugs' efficiency ratios.

In general, statins do not relieve any acute symptoms for patients, and that makes it difficult for patients to experience their ability to reduce heart disease risks. Moreover, because patients do not feel any direct discomfort from the discontinuation of statin treatment, a significant proportion of patients discontinues statin treatment in each period (Neslin et al. 2009). It should also be pointed out that most statins actually have minimal side effects. The only potential side effect is muscle pain, and medical research cannot find a statistically significant difference between the treatment and control groups for all statins but Baycol.[6] This has led to this statement from the *Guardian*, "Cholesterol-lowering statins have almost no side-effects."[7]

Table 1 shows the manufacturer of each statin, when it entered the market, and when each of them started facing generic entry. During our 10-year sample period, seven statins entered the market and one exited. Interestingly, they all entered at different points of time. Mevacor is the first statin, and it entered in Q3 1988. After two years, Zocor and Pravachol entered the market in Q3 and Q4, respectively, in 1990. Lescol then entered the market in Q1 1994. Lipitor, being the fifth entrant, entered the market in Q1 1997, three years after Lescol. Baycol then entered one year later (Q1 1998) but was withdrawn in Q3 2001 because of its potentially serious side effect mentioned earlier. Crestor entered the market in Q1 2003, five years after Baycol's entry.

Table 1. Summary Information on Statins

Brand	Molecule	Entry date	Generic entry	Manufacturer
Mevacor	lovastatin	Q3/1988	Q2/1997	Merck & Co.
Zocor	simvastatin	Q3/1990	Q1/2003	Merck & Co.
Pravachol	pravastatin	Q4/1990	Q3/2000	Bristol-Myers Squibb
Lescol	fluvastatin	Ql/1994	NA	Novartis
Lipitor	atorvastatin	Ql/1997	NA	Pfizer
Baycol	cerivastatin	Ql/1998	NA	Bayer
Crestor	rosuvastatin	Q1/2003	NA	AstraZeneca

Notes. For Lescol and Lipitor, the patent expiration date is beyond our sample period. Baycol was withdrawn from the market in Q3/2001 before its patent expired. NA, not applicable.

During our sample period, three early entrants (Mevacor, Zocor, and Pravachol) face generic entry.

4. Data

This research makes use of four different data sources: (i) product-level quarterly prescription volume and detailing data for the Canadian statin market from IMS Canada, (ii) product-level quarterly prescription switching rates between statins and discontinuing rates from Ontario Health Insurance Program (OHIP), (iii) landmark clinical trials obtained from published medical journals and a meta-analysis that summarizes statins' efficacy in lowering the bad cholesterol, and (iv) news articles covering statins collected from Factiva.

4.1. Prescription Volume and Detailing

The product-level data obtained from the marketing research firm IMS Canada consists of quarterly observations of prescription volumes and detailing costs for each statin across Canada from Q2 1993 ($t = 1$) to Q4 2004 ($t = 47$). The market is defined as the national market for quarter t. The observation is defined as a molecule-quarter combination.

In Figure 1, we plot the quarterly prescription volumes for seven statins in Canada. The prescription volume for Lipitor reached around 2.8 million by 2004, and the earlier arrivals, Zocor and Pravachol, had 900,000 and 500,000 quarterly prescriptions, respectively. In 2002, Lipitor achieved estimated annual global sales of $7.4 billion and became the best-selling product in the prescription drug market. Lipitor entered the market in 1997, and it did not have any direct evidence about its efficacy in reducing heart disease risks until May 2003. Yet, as shown in Figure 1, Lipitor was able to rise rapidly and surpass its rivals in merely six quarters after its entry. It becomes, by far, the most popular statin well before its first landmark clinical trial is released in Q2 2003. Given that the ultimate goal of taking statins is to reduce the chance of having heart disease or strokes, the success of Lipitor seems puzzling.

Previous research has documented that marketing activities have an influence on physicians' learning. Because detailing is considered a major promotional activity in the pharmaceutical industry, it is important for us to consider detailing data for each drug to see if it can explain the success of Lipitor. Figure 2 graphs the evolution of the quarterly detailing spending for five major statins.[8] Note that Mevacor (Q2 1997), Pravachol (Q3 2000), and Zocor (Q1 2003) stopped detailing when generic substitutes for their own products were introduced in the market. The detailing of Lipitor and Crestor started in Q1 1997 and Q1 2003, respectively, immediately following their launch, and their initial levels are high. It seems likely that detailing can partially explain the success of Lipitor and Crestor. Hence, in our structural model, we control for the effects of detailing by explicitly modeling its informative and persuasive roles.

4.2. Switching- and Discontinuing-Rate Data

Our data on switching and discontinuing rates are obtained from OHIP. It consists of the quarterly number of patients who continue using the same statin (c_{jt}), patients who switch to other statins (s_{jt}), and patients who discontinue statin medication (d_{jt}) at time t among the patients who use statin j at time $t-1$ in Ontario from Q2 1993 ($t = 1$) to Q4 2004 ($t = 47$).[9] From this data set, we obtain the switching rate $S_{jt} = s_{jt}/(c_{jt} + s_{jt} + d_{jt})$ and the discontinuing rate $D_{jt} = d_{jt}/(c_{jt} + s_{jt} + d_{jt})$. Figure 3 shows

Figure 1. (Color online) Quarterly Number of Total Prescriptions for Statins

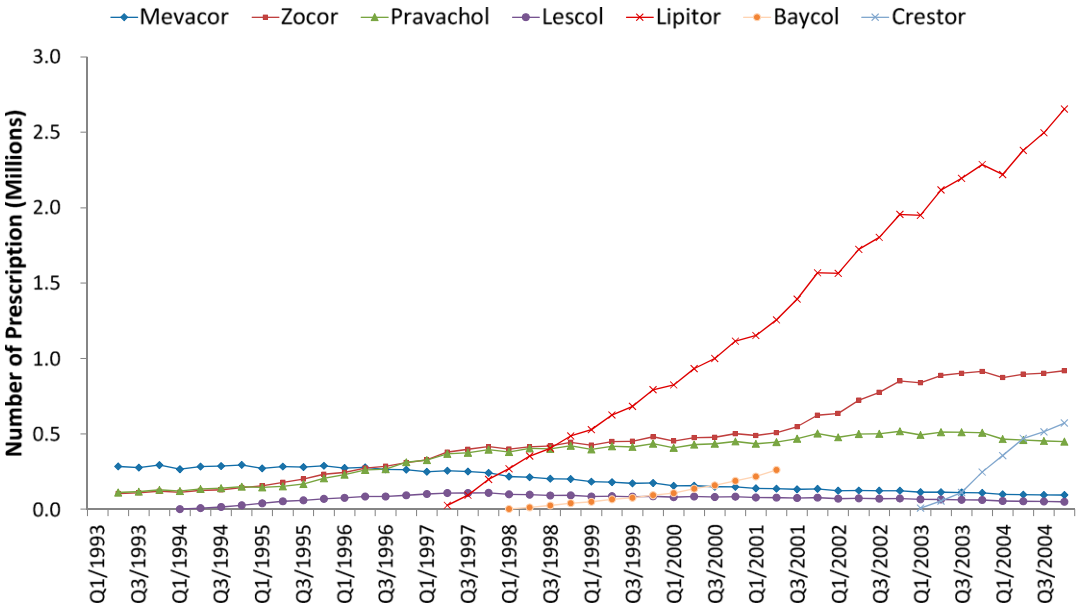
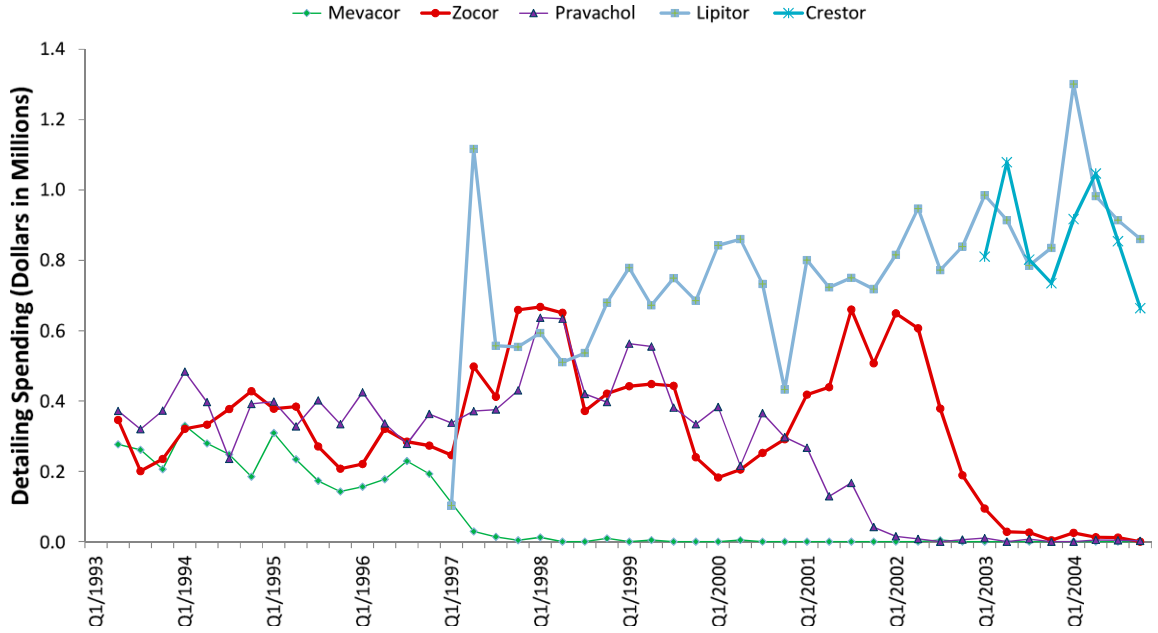


Figure 2. (Color online) Quarterly Detailing Spending for Leading Statins



that the switching rates between statins are less than 5% for almost all quarters for all drugs. Such low switching rates indicate the presence of large switching costs in the statin market. Moreover, switching rates became higher when Lipitor and Crestor were introduced in 1997 and 2003, respectively. Figure 4 shows that discontinuing rates are around 15% on average, which suggests that the cost of refilling prescriptions is high.

Note that prescription volume and detailing data are for the whole Canadian market, and the switching- and discontinuing-rate data are for Ontario only. Unfortunately, we are not able to obtain the switching- and discontinuing-rate data for other provinces. But the population in Ontario is more than one third of the population in Canada; they should serve as a reasonable proxy of average rates in Canada.

4.3. Clinical Trials

In general, there are two types of clinical trials in the statin market: (i) landmark clinical trials, which focus

Figure 3. (Color online) Quarterly Switching Rate for Leading Statins

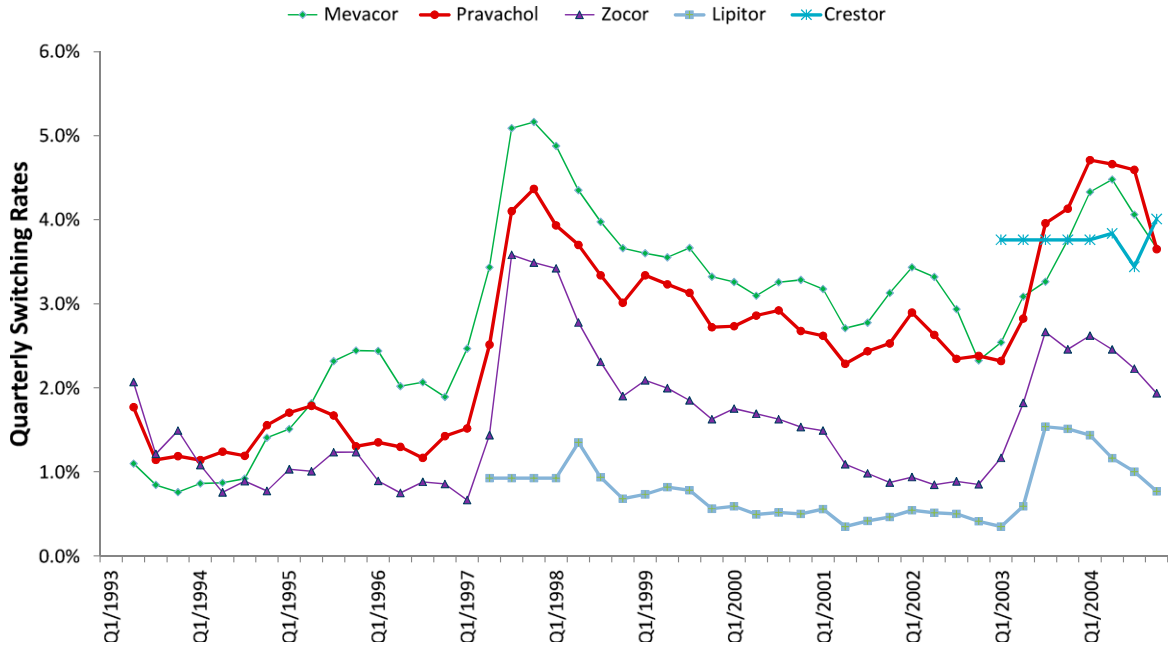
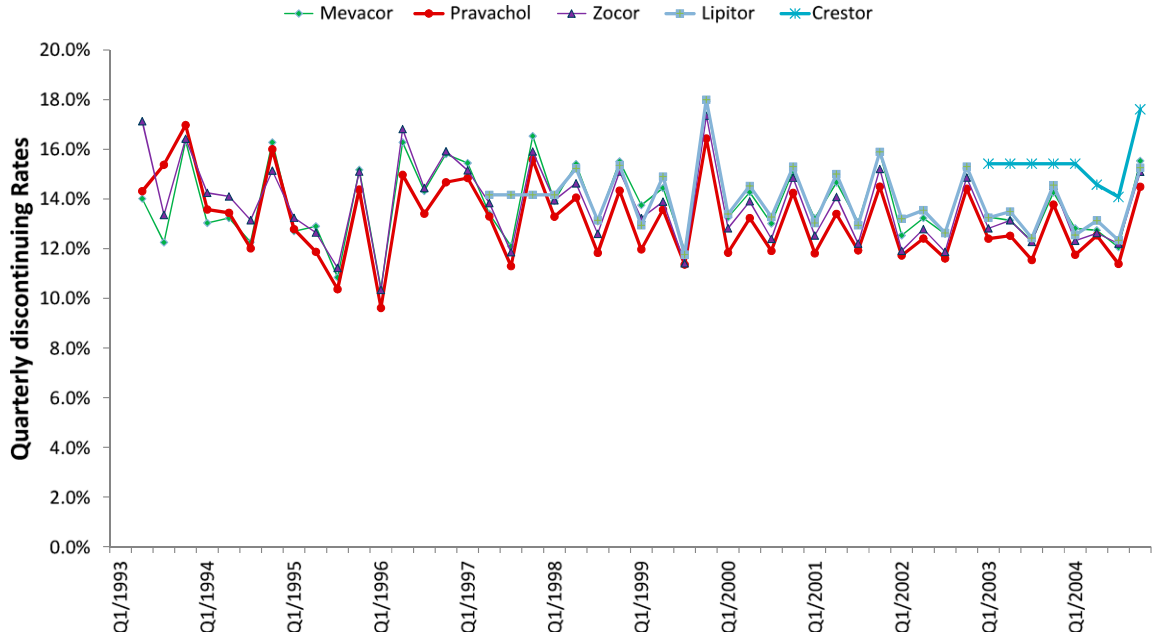


Figure 4. (Color online) Quarterly Discontinuing Rate for Leading Statins



on investigating a drug's efficiency ratio as we mentioned previously, and (ii) nonlandmark clinical trials, which just investigate drugs' efficacy of lowering LDL. Nonlandmark trials usually have less than 200 subjects, and their follow-up periods only last for several weeks.[10] The small number of subjects and short follow-up periods in nonlandmark clinical trials suggest that physicians can also learn about a statin's efficacy of lowering LDL from their own patients' experiences fairly quickly.

On the contrary, the landmark trials have (a) much higher numbers of subjects, ranging from 1,351 to 20,536, and (b) much longer follow-up periods of at least three years. This is because the occurrence of heart disease events is very low even for high-risk patients. Hence, one has to follow a large number of patients for a long period of time to document statistically significant results. For instance, the WOSCOPS trial shows only 5.5% and 7.9% of the subjects in the treatment (Pravachol) and control groups, respectively, experienced a definite nonfatal myocardial infarction or death from coronary heart disease after a follow-up period of 4.8 years. Such a low percentage of occurrence is the reason why the study needs 6,595 subjects to document this statistically meaningful result. Because it is so rare for physicians to observe a heart attack, it seems implausible that they would try to learn a statin's efficacy in reducing heart disease risks directly from their patients' experiences.[11] This is also why pharmaceutical companies are willing to invest in landmark clinical trials to provide doctors with such information.

Every statin is approved as a cholesterol-lowering drug, and the manufacturer is required by public health agencies to prove its statin's ability in lowering LDL through nonlandmark clinical trials before gaining the approval. Hence, the information about a statin's efficacy in reducing LDL is available at the time when it enters the market. Law et al. (2003) conducted a meta-analysis summarizing the results of clinical trials that investigate the effectiveness of statins on reducing LDL.

Table 2 shows the mean absolute LDL reduction of each statin by strength.[12] By taking the average of the reported mean LDL reductions across the strengths of each drug, we create a drug-specific LDL reduction efficacy variable, which we use in our empirical analysis.[13]

Each landmark clinical trial studies one statin and uses patients with different conditions. The follow-up periods of all trials last for several years. Even though the clinical end points and patients' conditions are slightly different across landmark clinical trials, every landmark clinical trial reports two main clinical end points for the drug being studied: (i) mean absolute LDL reduction and (ii) percentage reduction in major vascular events, which include strokes.

The medical literature assumes that there is an overall positive and linear relationship between reduction in

Table 2. Statins' Mean Cholesterol Reduction by Strength (mmol/L)

	Daily dose (mg)					
Drug	5	10	20	40	80	Mean
Mevacor	NA	1.02	1.40	1.77	2.15	1.59
Zocor	1.08	1.31	1.54	1.78	2.01	1.66
Pravachol	0.73	0.95	1.17	1.38	1.60	1.28
Lescol	0.46	0.74	1.02	1.30	1.58	1.16
Lipitor	1.51	1.79	2.07	2.36	2.64	2.22
Crestor	1.84	2.08	2.32	2.56	2.80	2.44

Note. NA, not applicable.

LDL and reduction in the risk for major cardiovascular events as demonstrated in several meta-studies; for example, see figure 3 in Cholesterol Treatment Trialists' Collaborators (2005) and figure 1 of Delahoy et al. (2009). To capture this relationship, we introduce the term efficiency ratio, which measures how efficiently a statin can convert reduction in LDL to reduction in heart disease risks. Based on the two end point measures reported in a landmark trial, one can easily compute its implied efficiency ratio. Table 3 lists the 14 landmark clinical trials used in this research. We include information on which drugs are being studied, their subject sizes, and implied efficiency ratios.

4.4. News Coverage (Publicity)

The news coverage data are obtained from Ching et al. (2016), which explains the details. We briefly describe the data here. The data are collected from Factiva, and it bases on news articles covering statins that contain the word "statin" or words related to statin (such as the chemical names or brand names) from 1986 to 2004. We restrict our attention to sources to which Canadian patients should have access. For each article, we extract its headline, source, content, and publication date. We first map the information of each article into two multidimensional variables: (a) a *general* publicity variable ($publicity_t^g$) if it has sentences that discuss statins in general without referring to any particular statin by brand or chemical name and (b) a *drug-specific* publicity variable ($publicity_{jt}$) if it has sentences that refer to one or more statins by either brand or chemical name. Note that an article may contain information that can be mapped onto both variables; it can provide general information about statins at the beginning and then later provide drug-specific information. Ching et al. (2016) find that general publicity can affect the overall demand for statins, and drug-specific publicity influences both total demand for statins and which particular statin to use.[14]

We classify both general and drug-specific publicity into three dimensions: lowering the bad cholesterol (*lc*), reducing heart disease risks (*rh*), and side effects (*se*). Hereafter, we use lc_t^g, rh_t^g, and se_t^g to represent the three dimensions of the general publicity variable in time t. For the drug-specific publicity variable, we use lc_{jt}, rh_{jt}, and se_{jt} to represent its three dimensions, where j is an index for drug. For each dimension of both drug-specific and general publicity, we use a two-step Likert scale (+1, −1) to assess its tone. We assign "+1" ("−1") if the article contains sentences that favor (do not favor) the focal drug.

In our empirical analysis, the length of a period is a quarter. Because there are usually more than one news stories published/broadcasted in each quarter, we need to aggregate the outcomes of the news that appeared in the same quarter to obtain a quarterly observation. We use the following procedure to do the aggregation. Let $publicity_{t,l}^g, publicity_{jt,l}$ denote the publicity variables associated with article l that is published in quarter t. Also, let L_t be the total number of news stories that appeared in quarter t. Then the values of $publicity_t^g, publicity_{jt}$ are obtained by simply summing $publicity_{t,l}^g, publicity_{jt,l}$ across the news stories that appeared in quarter t. For example, $publicity_t^g = \sum_{l=1}^{L_t} publicity_{t,l}^g$.

Figure 5 shows the general publicity flow variables. Although there are some bad news articles about statins' side effects, especially in 2001 when Baycol was removed from the market, most news articles report that statins are effective in lowering cholesterol levels and reducing heart disease risks. Table 4 presents a descriptive summary of general and drug-specific publicity variables with entry quarter for each drug.

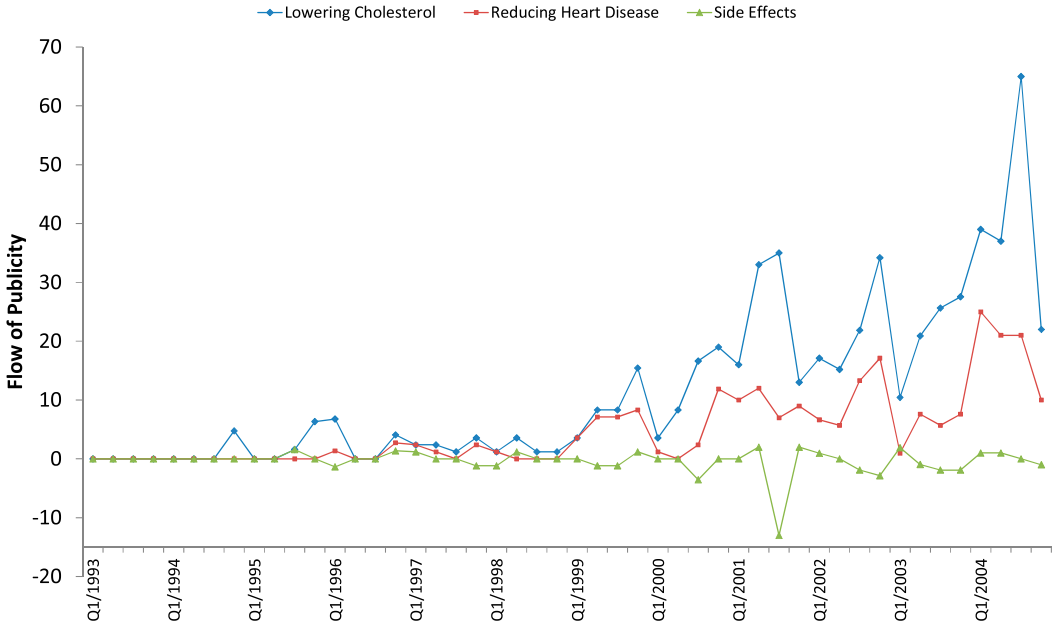
4.5. Potential Market Size

To study market expansion, our model includes an outside good (i.e., we allow patients with high cholesterol to choose treatments other than statins or no treatment at all). We, therefore, need to measure the potential market size for statins, which includes high-cholesterol patients who are on statins and other anticholesterol drugs and those who choose not to take any drugs.

Table 3. Landmark Clinical Trials for Statins

Title	Publication date	Drugs studied	Number of subjects	Follow-up period (years)	Efficiency ratio
4S	December 1994	Zocor	4,444	5.2	0.19
WOSCOPS	November 1995	Pravachol	6,595	4.8	0.26
CARE	October 1996	Pravachol	4,159	4.8	0.22
Post-CABG	January 1997	Mevacor	1,351	4.2	0.21
AFCAPS/TexCAPS	May 1998	Mevacor	6,605	5.3	0.31
LIPID	November 1998	Pravachol	9,014	5.6	0.19
GISSI Prevention	December 2000	Pravachol	4,271	1.9	0.29
LIPS	June 2002	Lescol	1,677	3.1	0.24
HPS	July 2002	Zocor	20,536	5.0	0.21
PROSPER	November 2002	Pravachol	5,804	3.2	0.13
ALLHAT-LLT	December 2002	Pravachol	10,355	4.8	0.13
ASCOT-LLA	May 2003	Lipitor	10,305	3.2	0.28
ALERT	June 2003	Lescol	2,102	5.1	0.12
CARDS	August 2004	Lipitor	2,838	3.9	0.32

Figure 5. (Color online) Quarterly Flow of General Publicity



To estimate the percentage of Canadians with a high-cholesterol problem, we follow Ching et al. (2016) and use data from Canadian Heart Health Surveys from 1986 and 1992. We multiply them by the total Canadian population for each age group in a given quarter (obtained from Statistics Canada) and use the result as a proxy for the total number of potential patients for statins. To convert the total population with high cholesterol levels to the number of prescriptions, we assume that each patient receives a prescription once per 90 days (see Cosh 2010 and Ching et al. 2016).

4.6. Preliminary Evidence for Correlated Learning

If correlated learning/information spillover is present, it is plausible that a firm may use its own detailing to inform physicians about its rival's landmark clinical trials. Hence, one empirical implication is that the impact of detailing for one drug may change upon the release of a landmark trial for another drug. To see if the data support this implication (and, hence, support correlated learning), we regress the number of prescriptions of a drug on brand dummies, brand-specific detailing stock, and clinical trial stock and interactions between individual landmark clinical trial and detailing stock. More specifically, we allow the interaction term to be heterogeneous across drugs. The coefficients of interest are the interaction terms because they provide evidence on how the release of a landmark trial for one drug may change the marginal impact of detailing for other drugs. A significant interaction term for drugs other than the one studied in the trial would be consistent with the presence of correlated learning/information spillover.

Table 4. Summary of Publicity Variables

			General publicity			
Dimension		Number of quarters	Mean	Standard deviation	Minimum	Maximum
Lowering cholesterol levels		47	11.59	14.07	0	65.00
Reducing heart disease risks		47	4.84	6.37	0	25.00
Side effects		47	−0.37	2.20	−13	2.00
Drug	Entry quarter	Number of quarters	Drug-specific publicity in reducing heart disease risks			
			Mean	Standard deviation	Minimum	Maximum
Mevacor	Q3/1988	47	0.84	1.77	−1.58	7.13
Zocor	Q3/1990	47	2.01	3.15	−3.00	14.25
Pravachol	Q4/1990	47	1.97	2.68	0	12.67
Lescol	Q1/1994	44	0.07	0.25	0	1.00
Lipitor	Q1/1997	32	2.98	4.28	0	16.15
Baycol	Q1/1998	14	0.07	0.27	0	1.00
Crestor	Q1/2003	8	0.84	1.07	0	2.85

We let STK_detail_{jt} denote the detailing goodwill stock for drug j in quarter t:

$$STK\_detail_{jt} = \delta_d \cdot STK\_detail_{jt-1} + detail_{jt}, \quad (1)$$

where δ_d is the quarterly carryover rate for detailing and $detail_{jt}$ is detailing spending (in thousands of dollars) for drug j in quarter t. We set $\delta_d = 0.95$.[15]

When constructing the clinical stock variable, we weigh each clinical trial by the number of patients. Following Azoulay (2002), we assume that clinical stocks do not depreciate over time. We let $STK\_clinical_{jt}$ denote the goodwill stock of clinical trials for drug j in quarter t:

$$STK\_clinical_{jt} = STK\_clinical_{jt-1} + n_{jt}, \quad (2)$$

where n_{jt} is the number of participants in the landmark trials for drug j released in quarter t (it equals zero if there is no landmark clinical trial for drug j released in quarter t).

We focus on the number of prescriptions for new patients and switching patients (i.e., patients who decide not to use the drug that they used last period). The demand from new patients and switching patients, d_{jt}^{ns}, can be expressed as[16]

$$d_{jt}^{ns} = d_{jt} - d_{jt-1} \cdot (1 - S_{jt} - D_{jt}), \quad (3)$$

where S_{jt} and D_{jt} are the switching and discontinuing rates of drug j at time t, respectively. We are able to obtain $\{d_{jt}^{ns}\}$ because we observe S_{jt}, D_{jt}, and d_{jt}.

Moreover, with our sample size, we cannot estimate a model that includes brand-specific interactions between detailing stock and every landmark clinical trial (there are 14 landmark trials in 40 quarters). We, therefore, decide to run a series of regressions that include one set of interactions per landmark clinical trial. The results are reported in Table 5. The set of drugs we consider includes Lipitor, Zocor, Mevacor, Pravachol, and Lescol.[17]

In our reduced-form regressions, we consider the set of landmark clinical trials that were released after Lipitor's entry date. However, because PROSPER and ALLHAT-LLT are released in the same quarter and so are ASCOT-LLA and ALERT, we only include the dummies for PROSPER and ASCOT-LLA in the regressions. One should keep in mind that the impact of PROSPER includes ALLHAT-LLT and that of ASCOT-LLA also includes ALERT. All these clinical trials provide positive news about the drug being investigated (see Table 3). Each specification in Table 5 corresponds to one clinical trial, and we present the regressions according to the chronological order of the clinical trials.

We discuss specification (1) first. AFCAPS/TexCAPS is a landmark trial for Mevacor. We focus on the coefficients for the interaction between detailing stock and the dummy variable that indicates whether AFCAPS/TexCAPS is available in quarter t.

In terms of boasting the effectiveness of detailing, Lipitor benefits most from AFCAPS/TexCAPS with a positive and statistically significant estimate. Its effects for Zocor and Mevacor are not statistically significant. Both Pravachol and Lescol have negative and statistically significant interaction terms. Judging from the point estimates, the ranking of the interaction terms are Lipitor > Zocor > Mevacor > Pravachol > Lescol. Interestingly, it coincides with their ranking based on the efficacy of lowering cholesterol (see Table 2). Moreover, these patterns generally hold across model specifications that use different landmark clinical trials.

Why are some interaction terms insignificant and some negative? What could explain this asymmetric effect? On the one hand, if correlated learning is important, we expect every statin to benefit from a landmark clinical trial regardless of which statin it studies. On the other hand, if a statin (e.g., Lipitor) benefits more from the information spillover compared with other statins, its enhanced position will negatively affect the demand for other statins. Depending on how much information spillover other drugs can benefit from, this competitive effect may lead to zero or negative net effect in their interaction terms.

Now let us focus on why the order of the interaction terms is broadly consistent with that of drugs' effectiveness in lowering cholesterol. This pattern begs a theory to explain it. We hypothesize that doctors may be uncertain about to what extent a statin can translate its cholesterol-lowering ability to lowering heart disease risks. A landmark clinical trial provides us with information to update the belief about this link for all drugs even though the information spillover may be imperfect. Positive information from the trial would lead doctors to revise the strength of this link upward. Consequently, a statin that is more effective in lowering cholesterol could end up benefiting more from another statin's landmark clinical trial. Our structural model captures this intuition.

We should also point out another interesting pattern. The size of the interaction terms tend to decline as we move from older to more recent landmark trials (i.e., as we move from specification (1) to (8)). Such a pattern is consistent with the Bayesian learning theory. In general, when an agent accumulates more information signals, the later signals would have less impact on the agent's information set (see Ching et al. 2013, 2017).

5. Model and Estimation

In this section, we propose a structural demand model incorporating physicians' correlated learning about clinical trial outcomes. We also discuss the identification issues and how to construct the likelihood function. It is important to emphasize that we only model drug choice for new patients and existing patients who decide to switch. We do not model why consumers decide to quit statins or why the vast majority of

Table 5. Preliminary Evidence for Correlated Learning

	(1)		(2)		(3)		(4)	
Variables	Estimates	Standard error	Estimates	Standard error	Estimates	Standard error	Estimates	Standard error
Lipitor	**50,551.16**	10,535.49	**69,405.80**	8,964.80	**105,966.20**	7,025.45	**118,377.50**	8,176.97
Zocor	–2,040.35	6,668.62	–502.93	6,437.79	4,713.42	5,900.29	–3,901.97	6,634.14
Mevacor	**–19,332.84**	7,620.78	**–19,384.58**	7,233.03	**–11,906.08**	5,916.81	**–28,900.97**	5,964.95
Pravachol	**–23,779.38**	7,344.16	**–23,313.41**	7,088.09	**–18,683.57**	6,642.86	**–33,235.43**	7,450.66
Lescol	1,420.78	6,823.54	–1,851.94	6,412.28	**–11,020.88**	5,585.86	**–16,956.12**	6,156.08
STK_Detail_{jt} (in thousand dollars)	**1,519.20**	185.47	**1,478.52**	168.58	**1,245.33**	133.88	**1,829.82**	139.54
$STK\_Clinical_{jt}$	**2.15**	0.23	**2.24**	0.24	**1.51**	0.26	**1.23**	0.36
STK_Detail_{jt} × *Lipitor* × *AFCAPS*	**1,416.81**	198.45						
STK_Detail_{jt} × *Zocor* × *AFCAPS*	119.81	125.34						
STK_Detail_{jt} × *Mevacor* × *AFCAPS*	–259.21	325.13						
STK_Detail_{it} × *Pravachol* × *AFCAPS*	**–480.69**	130.99						
STK_Detail_{jt} × *Lescol* × *AFCAPS*	**–775.64**	179.19						
STK_Detail_{jt} × *Lipitor* × *LIPID*			**1,265.78**	163.14				
STK_Detail_{jt} × *Zocor* × *LIPID*			132.64	118.26				
STK_Detail_{jt} × *Mevacor* × *LIPID*			–238.78	379.41				
STK_Detail_{jt} × *Pravachol* × *LIPID*			**–545.25**	130.31				
STK_Detail_{jt} × *Lescol* × *LIPID*			**–666.31**	173.43				
STK_Detail_{jt} × *Lipitor* × *GISSI*					**1,207.84**	106.87		
STK_Detail_{jt} × *Zocor* × *GISSI*					**497.01**	118.83		
*STK Detail^* × *Mevacor* × *GISSI*					54.87	548.94		
STK_Detail_{jt} × *Pravachol* × *GISSI*					–134.16	157.49		
STK_Detail_{jt} × *Lescol* × *GISSI*					–312.07	204.60		
STK_Detail_{jt} × *Lipitor* × *LIPS*							**708.43**	108.28
STK_Detail_{jt} × *Zocor* × *LIPS*							**492.72**	180.14
STK_Detail_{jt} × *Mevacor* × *LIPS*							1,377.55	884.96
STK_Detail_{jt} × *Pravachol* × *LIPS*							170.22	298.22
STK_Detail_{jt} × *Lescol* × *LIPS*							51.64	336.07
R^2	0.955		0.954		0.953		0.935	
Adjusted R^2	0.953		0.952		0.950		0.932	
Number of observations	211		211		211		211	

Notes. Estimates shown in bold are significant at the 5% level. Definitions of variables are as follows: STK_Detail_{jt}, cumulative stock of detailing for drug j at quarter t, and its carryover rate is 95%. $STK\_Clinical_{jt}$, cumulative clinical oucomes for drug j at quarter t, and its carryover rate is 100%. *Mevacor*, *Zocor*, *Pravachol*, *Lescol*, and *Lipitor* denote a brand dummy for each drug. *AFCAPS*, *LIPID*, *GISSI*, and *LIPS* are dummy variables that indicate observation periods after each clinical trial.

existing patients keep using the same brand/drug even though another drug is superior based on expected utility; we leave these topics for future research.

5.1. Bayesian Learning Model

Consider a situation in which physician k needs to decide which drug to prescribe for patient i. The utility of patient i who consumes drug j at time t is given by

$$U_{ijt} = \omega \cdot q_j^h + \lambda_j + \epsilon_{ijt}, \tag{4}$$

where q_j^h denotes drug j's efficacy in reducing heart disease risks, λ_j captures other time-invariant brand-specific preference (e.g., price difference across brands,[18] drug j's efficacy in reducing LDL, etc.), and ϵ_{ijt} is an independent and identically distributed (i.i.d.) extreme value distributed random shock.

Note that physicians/patients are uncertain about q_j^h. We, therefore, assume that a physician makes prescribing decisions to maximize the expected utility for the physician's patients. Let $I(k,t)$ denote physician k's information set at time t. For simplicity, we assume that a physician does not receive utility from persuasive detailing (which we introduce in Section 5.2). Hence, physician k's expected utility of prescribing drug j to patient i at time t is

$$E\left[U_{ijt}|I(k,t)\right] = \omega \cdot E\left[q_j^h|I(k,t)\right] + \lambda_j + \epsilon_{ijt}, \tag{5}$$

Table 5. Preliminary Evidence for Correlated Learning (Continued)

	(5)		(6)		(7)		(8)	
Variables	Estimates	Standard error	Estimates	Standard error	Estimates	Standard error	Estimates	Standard error
Lipitor	**118,758.70**	8,290.16	**119,225.90**	8,584.49	**118,557.60**	8,631.76	**119,245.70**	8,590.44
Zocor	–5,479.44	6,640.60	–8,337.09	6,859.05	–12,056.87	6,835.29	**–15,855.79**	6,572.21
Mevacor	**–30,470.27**	5,829.98	**–34,651.05**	5,899.36	**–37,809.74**	5,660.18	**–40,506.29**	5,081.96
Pravachol	**–35,190.41**	7,371.81	**–40,772.01**	7,557.44	**–45,701.84**	7,439.80	**–55,103.45**	7,085.52
Lescol	**–17,751.28**	6,141.47	**–18,203.34**	6,286.96	**–19,463.76**	6,190.17	**–19,945.32**	5,909.42
$STK\ Detail_{jt}$ (in thousand dollars)	**1,891.49**	137.56	**2,023.55**	138.86	**2,149.10**	132.19	**2,260.77**	113.49
$STK\_Clinical_{jt}$	**1.25**	0.36	**1.41**	0.34	**1.45**	0.30	**1.83**	0.24
$STK\_Detail_{jt} \times Lipitor \times HPS$	**671.29**	107.78						
$STK\_Detail_{jt} \times Zocor \times HPS$	**524.90**	188.01						
$STK\_Detail_{jt} \times Mevacor \times HPS$	1,510.35	950.21						
$STK\_Detail_{jt} \times Pravachol \times HPS$	174.28	329.47						
$STK\_Detail_{jt} \times Lescol \times HPS$	119.84	362.98						
$STK\_Detail_{jt} \times Lipitor \times PROSPER$			**524.19**	111.79				
$STK\_Detail_{jt} \times Zocor \times PROSPER$			**439.36**	198.91				
$STK\_Detail_{jt} \times Mevacor \times PROSPER$			1,705.66	1,055.23				
$STK\_Detail_{jt} \times Pravachol \times PROSPER$			139.36	340.45				
$STK\_Detail_{jt} \times Lescol \times PROSPER$			176.43	406.43				
$STK\_Detail_{jt} \times Lipitor \times ASCOT$					**432.26**	115.35		
$STK\_Detail_{jt} \times Zocor \times ASCOT$					**546.73**	222.71		
$STK\_Detail_{jt} \times Mevacor \times ASCOT$					2,100.46	1,256.25		
$STK\_Detail_{jt} \times Pravachol \times ASCOT$					254.17	364.43		
$STK\_Detail_{jt} \times Lescol \times ASCOT$					327.87	489.29		
$STK\_Detail_{jt} \times Lipitor \times CARDS$							**679.37**	212.93
$STK\_Detail_{it} \times Zocor \times CARDS$							**865.96**	546.95
$STK\_Detail_{jt} \times Mevacor \times CARDS$							2,670.29	3,304.06
$STK\_Detail_{jt} \times Pravachol \times CARDS$							377.21	815.98
$STK\_Detail_{jt} \times Lescol \times CARDS$							639.67	1,303.26
R^2	0.934		0.929		0.927		0.924	
Adjusted R^2	0.931		0.925		0.923		0.919	
Number of observations	211		211		211		211	

Notes. Estimates shown in bold are significant at 5% level. Definitions of variables are as follows: STK_Detail_{jt}, cumulative stock of detailing for drug j at quarter t, and its carryover rate is 95%. $STK\_Clinical_{jt}$, cumulative clinical oucomes for drug j at quarter t, and its carryover rate is 100%. *Mevacor*, *Zocor*, *Pravachol*, *Lescol*, and *Lipitor* denote a brand dummy for each drug. *HPS*, *PROSPER*, *ASCOT*, and *CARDS* are dummy variables that indicate observation periods after each clinical trial.

where $E[\cdot|\, I(k,t)]$ denotes the expected value given physician k's information set at time t. The demand system is obtained by aggregating this discrete choice model of an individual physician's behavior.

Recall that we define the efficiency ratio (β_j) as a measure on how efficiently a drug can convert reduction in LDL to reduction in heart disease risks. Let q_j^c be the efficacy in lowering LDL of drug j. Then q_j^h can be expressed as follows:

$$q_j^h = q_j^c \cdot \beta_j. \tag{6}$$

We also assume that physicians have complete information about q_j^c but are uncertain about β_j.[19] Consequently, physician k's expectation about q_j^h can be expressed as follows:

$$E\left[q_j^h|I(k,t)\right] = q_j^c \cdot E\left[\beta_j|I(k,t)\right]. \tag{7}$$

In Section 6.2, we consider a version of the model in which physicians may forget q_j^c, and detailing can help remind them about it.

How do physicians learn about β_j's? We model physicians' learning process by adopting the Bayesian learning framework (DeGroot 1970). Physicians construct their initial prior belief before they learn about the results of landmark clinical trials. As discussed earlier, because all

statins use a similar mechanism to lower the cholesterol level, their initial prior belief about β_j's may be correlated across j. In other words, information about β_j can be useful for updating β_{-j} and vice versa. Because of this intrinsic correlated prior belief, physicians may infer q_j^h indirectly from the clinical trial evidence on β_{-j}. More specifically, we assume the initial prior is normally distributed and allow the off-diagonal elements (ρ) in the variance–covariance matrix for the initial prior beliefs to be nonzero. As a first step, we assume the initial priors to be the same across drugs.

As we explained earlier, it is unlikely for physicians to learn about the efficacy in heart disease risks of statins from their patients' experiences, and we assume landmark clinical trials (which are specifically designed to prove the efficacy of drugs in heart disease risks) are the only sources of information about this dimension of efficacy. Physicians are assumed to update their beliefs on β_j of each drug when they are exposed to landmark clinical trial results. Note that, even when two landmark clinical trials test the same drug, they typically select patients with different conditions (e.g., one trial may use participants with diabetes, and another trial uses participants who previously had a heart attack). Hence, β's obtained from different landmark clinical trials should not be the same even if the number of patients goes to infinity. We assume that the true efficiency ratio associated with the condition studied by trial l ($\bar{\beta}_{jl}$) is capture by

$$\bar{\beta}_{jl} = \beta_j + \epsilon_{jl}, \tag{8}$$

where β_j is the true mean level of the efficiency ratio for drug j and $\epsilon_{jl} \sim N(0, \sigma_\epsilon)$ and is i.i.d. But the exact signals revealed by clinical trial l could differ from $\bar{\beta}_{jl}$ because of sampling errors. More precisely, the signal from clinical trial l for drug j ($\tilde{\beta}_{jl}$) can be expressed as

$$\tilde{\beta}_{jl} = \bar{\beta}_{jl} + \zeta_{jl}, \tag{9}$$

where ζ_{jl} is an i.i.d. signal noise and $\zeta_{jl} \sim N(0, \sigma_{\zeta l}^2)$. Let σ_ζ^2 be signal variance for one patient and N_l be the number of patients who participate in landmark clinical trial l. As long as the individual signals are i.i.d. across patients, it can be shown that $\sigma_{\zeta l}^2 = \sigma_\zeta^2 / N_l$. This implies that the more participants a clinical trial has, the more confidence physicians have about its results. Moreover, we can combine the two equations and write

$$\tilde{\beta}_{jl} = \beta_j + \nu_{jl}, \tag{10}$$

where $\nu_{jl} = \epsilon_{jl} + \zeta_{jl}$. This implies that $\nu_{jl} \sim N(0, \sigma_{\nu l}^2)$, where $\sigma_{\nu l}^2 = \sigma_\epsilon^2 + \sigma_\zeta^2 / N_l$. Note that, unlike most of the previous research, we are able to observe quality signals by using the information from the landmark clinical trials. As we explain later, the clinical trial data helps us identify the correlated learning parameters and simplify the estimation procedure.

To explain how physicians update their beliefs through learning about clinical trials, let us provide a simplified example that can be easily generalized. Our model extends the correlated learning models by Erdem (1998), Erdem and Keane (1996), Marcoul and Weninger (2008), and Janakiraman et al. (2009). In this example, we assume that there are two statins ($j = 1, 2$) and there is only one landmark clinical trial that investigates drug 1's efficacy in reducing heart disease risks. Let β_{jt} be the expected perceived efficiency ratio and $\sigma_{\beta jt}^2$ be the perceived variance of drug j conditional on physician k's information set at time t. The variance–covariance matrix for prior beliefs of physician k at time t becomes

$$V\left[\beta_j | I(k,t)\right] = \begin{pmatrix} \sigma_{\beta 1t}^2 & \pi_t \\ \pi_t & \sigma_{\beta 2t}^2 \end{pmatrix}. \tag{11}$$

If the physician learns about clinical trial l for drug 1, the physician will update the physician's posterior mean on the efficiency ratio of drug 1 as follows:

$$\beta_{1t+1} = \beta_{1t} + \frac{\sigma_{\beta 1t}^2}{\sigma_{\beta 1t}^2 + \sigma_{\nu 1l}^2} \cdot \left(\tilde{\beta}_{1l} - \beta_{1t}\right). \tag{12}$$

The physician will also update the physician's prior variance on the efficiency ratio of drug 1 at time t as follows:

$$\sigma_{\beta 1t+1}^2 = \frac{\sigma_{\beta 1t}^2 \sigma_{\nu 1l}^2}{\sigma_{\beta 1t}^2 + \sigma_{\nu 1l}^2}. \tag{13}$$

With correlated prior beliefs on the efficiency ratio, signals for drug 1 are used to update the posterior mean on drug 2 as well. The posterior mean for drug 2 is given by

$$\beta_{2t+1} = \beta_{2t} + \frac{\pi_t}{\sigma_{\beta 2t}^2 + \sigma_{\nu 1l}^2} \left(\tilde{\beta}_{1l} - \beta_{1t}\right), \tag{14}$$

where π_t denotes the off-diagonal element in the variance–covariance matrix of the perceived quality on the efficiency ratio at time t.

The posterior variance on the efficiency ratio of drug 2 at time t becomes

$$\sigma_{\beta 2t+1}^2 = \sigma_{\beta 2t}^2 - \frac{\pi_t^2}{\sigma_{\beta 2t}^2 + \sigma_{\nu 1l}^2}. \tag{15}$$

The off-diagonal element of the variance–covariance matrix for posterior beliefs becomes

$$\pi_{t+1} = \frac{\pi_t \sigma_{\nu 1l}^2}{\sigma_{\beta 1t}^2 + \sigma_{\nu 1l}^2}. \tag{16}$$

As a result, the variance–covariance matrix for posterior beliefs becomes

$$V[\beta_j|I(k,t+1)] = \begin{pmatrix} \frac{\sigma^2_{\beta 1t}\sigma^2_{v1l}}{\sigma^2_{\beta 1t}+\sigma^2_{v1l}} & \frac{\pi_t\sigma^2_{v1l}}{\sigma^2_{\beta 1t}+\sigma^2_{v1l}} \\ \frac{\pi_t\sigma^2_{v1l}}{\sigma^2_{\beta 1t}+\sigma^2_{v1l}} & \sigma^2_{\beta 2t}-\frac{\pi_t^2}{\sigma^2_{\beta 2t}+\sigma^2_{v1l}} \end{pmatrix}. \quad (17)$$

It is straightforward to extend this example to a model with N drugs.

5.1.1. Remarks. Before leaving this subsection, we make two remarks here. First, one may argue that q_j^c could also have a direct effect in the utility function. We believe this is plausible. However, given our assumption that physicians have complete information about q_j^c, its effect are absorbed by the brand dummy, λ_j.[20] Nevertheless, it is also possible that some physicians may not have complete information about q_j^c. In Section 6.2, we estimate a version of the model that allows q_j^c to enter the utility function directly, and physicians have incomplete information about it. The results remains largely unchanged, and in particular, q_j^c is insignificant. Second, one might think that physicians could be forward-looking and experiment with different drugs to learn about q_j^h. However, as we pointed out earlier, because heart attack and stroke are very rare events, it is unlikely that physicians use patients' experiences to update their belief about q_j^h. Therefore, we assume that physicians do not experiment with different drugs on their patients.

5.2. Roles of Detailing and Publicity

The economics and marketing literature studying the pharmaceutical industry find evidence that detailing can play both informative and persuasive roles (Leffler 1981, Narayanan et al. 2005, Ching and Ishihara 2012). The official role of detailing is to inform physicians about drugs' efficacies and side effects; this is referred to as the informative role. However, it has been argued that other noninformative activities (such as offering physicians free meals and gifts) could also cause physicians to favor prescribing the drug being detailed; this is referred to as the persuasive role. We model both roles and discuss how to separately identify them.

We first describe how we model the persuasive role. Here, we adopt the standard approach by modeling a detailing goodwill stock entering physician k's utility function directly. Therefore, we modify Equation (5) as follows:

$$E\left[U_{ijt}|I(k,t)\right] = \omega \cdot E\left[q_j^h|I(k,t)\right] + \kappa_d \cdot P\_STK\_detail_{jt} + \lambda_j + \epsilon_{ijt}, \quad (18)$$

where $P\_STK\_detail_{jt}$ is a persuasive detailing goodwill stock for drug j at time t. The persuasive detailing stock is defined as

$$P\_STK\_detail_{jt} = \delta_{d\_per} \cdot P\_STK\_detail_{jt-1} + detail_{jt}, \quad (19)$$

where δ_{d_per} is the quarterly carryover rate for persuasive detailing and $detail_{jt}$ denotes the flow of detailing spending for drug j at time t.

We now explain how to model the informative role of detailing and publicity. We modify the model proposed by Ching and Ishihara (2010). They model informative detailing as a means to build and maintain the measure of physicians who know the most updated information about drugs (hereafter, we refer to them as *well-informed physicians*). The basic setup of the model is as follows. There is a continuum of physicians with measure one. They are heterogeneous in their information sets. A physician is either well informed or uninformed about drug j. A well-informed physician knows the current information set maintained by the representative opinion leader, $I_j(t)$. An uninformed physician only knows the initial prior, $\underline{I}_j$.

Let $I\_STK\_detail_{jt}$ and $\boldsymbol{STK\_PUB}_{jt}$ denote the informative detailing stock and publicity stocks, respectively, for drug j at time t. Let $I\_STK\_detail_{-jt}$ and $\boldsymbol{STK\_PUB}_{-jt}$ denote the competitors' informative detailing stock and publicity stocks, respectively, for drug j at time t. Let $\boldsymbol{STK\_PUB}_t^g$ denote general publicity stocks for the class of statin. We model the measure of well-informed physicians of drug j at time t (M_{jt}) as follows:

$$M_{jt} = \frac{\exp(L_{jt})}{1+\exp(L_{jt})}, \quad (20)$$

where $L_{jt} = \alpha_0 + \alpha_d \cdot I\_STK\_detail_{jt} + \alpha_{d\_c} \cdot I\_STK\_detail_{-jt} + \boldsymbol{\alpha}_p \cdot \boldsymbol{STK\_PUB}_{jt} + \boldsymbol{\alpha}_{p\_c} \cdot \boldsymbol{STK\_PUB}_{-jt} + \boldsymbol{\alpha}_{p\_g} \cdot \boldsymbol{STK\_PUB}_t^g$.

Note that we model M_{jt} as a function of $I\_STK\_detail_{-jt}$ to capture the possibility that sales reps of drug j might have discussed other statin's clinical trials to free ride on competitors' clinical trial results and take advantage of correlated learning. We include $\boldsymbol{STK\_PUB}_{-jt}$ because news coverage about one statin could trigger patients/physicians to do research about statins in general and learn about the latest clinical trial results for other statins as well.

The informative detailing stock is defined as

$$I\_STK\_detail_{jt} = \delta_{d\_inf} \cdot I\_STK\_detail_{jt-1} + detail_{jt}, \quad (21)$$

where δ_{d_inf} is the quarterly carryover rate for informative detailing and $detail_{jt}$ denotes the flow of detailing spending for drug j at time t. Similarly, the informative stock of competitors' detailing is defined as

$$I\_STK\_detail_{-jt} = \delta_{d\_inf} \cdot I\_STK\_detail_{-jt-1} + detail_{-jt}, \quad (22)$$

where $detail_{-jt}$ denotes the flow of the sum of detailing spending for all statins except for drug j at time t. We define $\boldsymbol{STK\_PUB}_{jt}$, $\boldsymbol{STK\_PUB}_{-jt}$, and $\boldsymbol{STK\_PUB}_t^g$ in a similar way.

5.3. Prescribing Decisions

Based on patients' choices in the previous period $(t-1)$, we classify patients at time t into two groups: "potential patients" and "existing patients." First, we explain the decision-making process of potential patients. As Figure 6 depicts, our model assumes that their decision-making process consists of two stages. The first stage (adoption decision stage) determines whether a potential patient will use statins. The decision in this stage could be jointly made by the patient and the patient's physician. For example, news articles reporting the problem of high cholesterol levels or the benefits of taking statins could entice the patient to see a physician. Alternatively, a physician detailed by pharmaceutical representatives might recommend the patient get a blood test. Therefore, we model how the general publicity and "aggregate" detailing spending affect the decision-making process in this stage. The probability that physician k prescribes one of the statins to potential patients at time t, $P_t(statin|k_{type})$, is expressed as follows:

$$P_t(statin|k_{type}) = \frac{\exp\left(\gamma_0 + \gamma_i \cdot Inclusive_t(k_{type}) + \boldsymbol{\gamma}_p \cdot \boldsymbol{STK\_PUB}_t^g\right)}{1 + \exp\left(\gamma_0 + \gamma_i \cdot Inclusive_t(k_{type}) + \boldsymbol{\gamma}_p \cdot \boldsymbol{STK\_PUB}_t^g\right)}, \tag{23}$$

where $Inclusive_t(k_{type}) = \ln(\sum_{j=1}^{J} \exp(U_{kjt}(k_{type})))$ is the inclusive value term derived from the brand choice stage, k_{type} denotes the type of physician k, and $\boldsymbol{STK\_PUB}_t^g$ denotes a vector of three types of general publicity (rh_t^g, lc_t^g, and se_t^g) stocks for the class of statin.[21] Note that we model drug-specific detailing and publicity in the brand choice stage. Hence, they can also influence the adoption decision via the inclusive value term.

If a potential patient decides to use statins, the patient moves to the second stage (statin choice stage), in which the physician determines which statin to prescribe. The physician evaluates all the statins available given the physician's information set. The probability that physician k chooses drug j for a new patient, conditional on prescribing, is expressed as follows:

$$P_t(j|statin, k_{type}) = \frac{\exp(U_{kjt}(k_{type}))}{\sum_{r=1}^{J} \exp(U_{krt}(k_{type}))}. \tag{24}$$

The information set of physician k ($I(k,t)$) is a function of physician k's type at time t. Because we assume that, for each drug j, a physician is either well informed about the most updated landmark clinical trials for this drug or is uninformed about them at all, the total number of physician types is 2^H. Note that $H = 5$ in our application because only five statins have landmark clinical trials during our sample period.

Let $P_t(k_{type})$ denote the probability of physician k being a particular type:

$$P_t(k_{type}) = \prod_{j=1}^{H} \left\{ I_{j,k_{type}} \cdot M_{jt} + \left(1 - I_{j,k_{type}}\right) \cdot \left(1 - M_{jt}\right) \right\}, \tag{25}$$

where $I_{j,k_{type}}$ is an indicator for whether physician k is well informed about drug j or not. Let d_{jt} be the demand for drug j in time t. The expected "new patient demand" (group 1) for drug j at time t, $\hat{d}_{jt}^1$, can be expressed as

$$\hat{d}_{jt}^1 = \left(m_t - \sum_{r=1}^{J} d_{rt-1}\right) \cdot \sum_{k_{type}=1}^{2^H} P_t(k_{type}) \cdot P_t(statin|k_{type}) \cdot P_t(j|statin, k_{type}), \tag{26}$$

where m_t is the potential market size for statins at time t and $(m_t - \sum_{r=1}^{J} d_{rt-1})$ is the potential patient pool that has not yet adopted statins at time t.

For existing patients, their decisions are more complicated than those of potential patients. Figure 7 depicts the decision tree of existing patients. In the first stage, they decide to either quit or keep taking statins. Once they decide to keep taking statins, they need to decide whether to stay with the same statin or switch to a different statin. If they decide to switch, then they choose a statin. Note that, because of switching costs, some patients might keep taking the same statin even though there are alternatives that give them higher expected utility.[22]

If an existing patient decides to stay with the current statin, we classify the patient as a "stayer." If the patient decides to switch to another statin, we classify the patient as a "switcher."

The expected demand for stayers (group 2) can be expressed as

$$\hat{d}_{jt}^2 = d_{jt-1} \cdot \left(1 - S_{jt} - D_{jt}\right), \tag{27}$$

where S_{jt} and D_{jt} denote the observed switching and discontinuing rates of drug j at time t. It should be highlighted that the whole sequence of $\{\hat{d}_{jt}^2\}$ is determined by Equation (27) because we observe S_{jt}, D_{jt}, and d_{jt}.

Figure 6. (Color online) Decision Process of Potential Patient

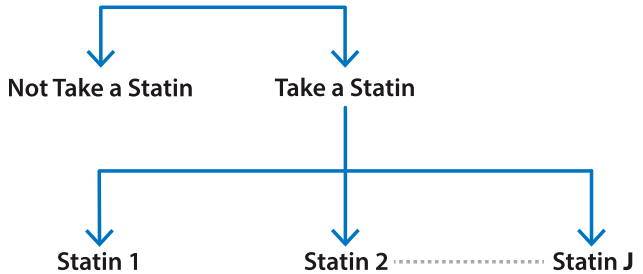
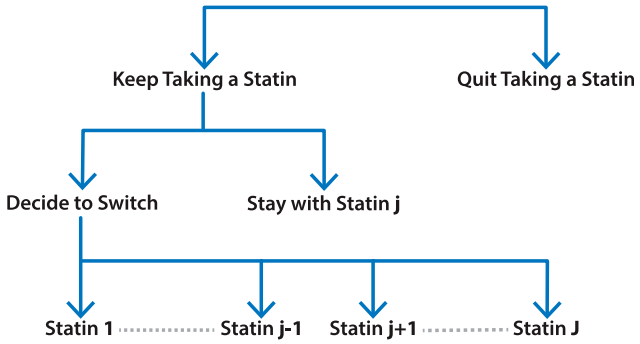


Figure 7. (Color online) Decision Process of Existing Patient



The expected demand for switchers (group 3) can be expressed as

$$\hat{d}_{jt}^{3} = \sum_{m=1,\neq j}^{J} \left\{ d_{mt-1} \cdot S_{mt} \cdot \sum_{k_{type}=1}^{2^H} \left[P_t(k_{type}) \cdot \frac{\exp(U_{kjt}(k_{type}))}{\sum_{r=1,\neq m}^{J} \exp(U_{krt}(k_{type}))} \right] \right\}. \tag{28}$$

where the first component, $d_{mt-1} \cdot S_{mt}$, represents the number of patients who did not choose drug j in $t-1$ and decided to switch to a different drug (they are switchers)[23] and the second component is the choice probability of choosing drug j given that these switchers choose drug m in $t-1$. We should note that, once a patient quits using the statin treatment, we assume the patient will be back to the potential patient pool in the next period.

We emphasize that we do not estimate the switching and refilling cost parameters and that we treat S_{jt} and D_{jt} as exogeneous in this research. We leave it for future research to address these limitations.

5.4. Estimation

5.4.1. Likelihood. We focus on modeling the demand for new and switching patients, d_{jt}^{ns}:

$$d_{jt}^{ns} = d_{jt} - \hat{d}_{jt}^{2} + e_{jt} = \hat{d}_{jt}^{1} + \hat{d}_{jt}^{3} + e_{jt}, \tag{29}$$

where e_{jt} represents a measurement error. Note that we use $\hat{d}_{jt}^{2}$ instead of d_{jt}^{2} in Equation (29), and this is one main source of measurement error because we use the switching and discontinuing rates in Ontario as a proxy for those in Canada.

Assuming that the measurement error e_{jt} in Equation (29) is normally distributed, we can obtain the likelihood function

$$l\left(\left\{d_{jt}^{ns}\right\}_{j=1}^{J} \middle| \left\{\left\{detail_{j\tau}\right\}_{j=1}^{J}\right\}_{\tau=1}^{t}, \left\{\tilde{\beta}_l\right\}_{l=1}^{C_t}, \{N_l\}_{l=1}^{C_t}, \left\{\boldsymbol{PUB}_{\tau}^{g}\right\}_{\tau=1}^{t}, \left\{\left\{\boldsymbol{PUB}_{j\tau}\right\}_{j=1}^{J}\right\}_{\tau=1}^{t}; \theta_d\right), \tag{30}$$

where θ_d is the vector of parameters, $detail_{jt}$ is detailing spending for drug j at time t, C_t denotes the number of landmark clinical trials up to time t, $\tilde{\beta}_l$ is the efficiency ratio signal from landmark clinical trial l, N_l is the number of participants in clinical trial l, and $\boldsymbol{PUB}_t^g$ and $\boldsymbol{PUB}_{jt}$ are vectors of general and drug-specific publicity, respectively. The likelihood of observing $d^{ns} = \{\{d_{jt}^{ns}\}_{j=1}^{J}\}_{t=1}^{T}$ is

$$\begin{aligned} & L\left(d^{ns} \middle| \left\{\left\{detail_{jt}\right\}_{j=1}^{J}\right\}_{t=1}^{T}, \left\{\tilde{\beta}_l\right\}_{l=1}^{C_T}, \{N_l\}_{l=1}^{C_T}, \left\{\boldsymbol{PUB}_t^g\right\}_{t=1}^{T}, \left\{\left\{\boldsymbol{PUB}_{jt}\right\}_{j=1}^{J}\right\}_{t=1}^{T}; \theta_d\right) \\ &= \prod_{t=1}^{T} l\left(\left\{d_{jt}^{ns}\right\}_{j=1}^{J} \middle| \left\{\left\{detail_{j\tau}\right\}_{j=1}^{J}\right\}_{\tau=1}^{t}, \left\{\tilde{\beta}_l\right\}_{l=1}^{C_t}, \{N_l\}_{l=1}^{C_t}, \left\{\boldsymbol{PUB}_{\tau}^{g}\right\}_{\tau=1}^{t}, \left\{\left\{\boldsymbol{PUB}_{j\tau}\right\}_{j=1}^{J}\right\}_{\tau=1}^{t}; \theta_d\right). \end{aligned} \tag{31}$$

We estimate parameters by maximum likelihood. Unlike the previous literature on learning models, all the quality signals are observable in our model. Therefore, the evolution of $E[q_j^h|I(t)]$ is deterministic, and we can simply construct the likelihood function without using any simulation method. Otherwise, one would need to use Monte Carlo methods to integrate out the unobserved quality signals associated with each clinical trial and $E[q_j^h|I(t)]$. As explained in Ching et al. (2013), such a computationally intensive approach is usually needed in estimating structural learning models because experience signals are unobserved in most cases.

5.4.2. Initial Condition Problem. Our data set for prescription volume starts only in Q2 1993. Mevacor, Zocor, and Pravachol were introduced earlier. Hence, by Q2 1993, these three drugs should have accumulated some detailing stocks. If we do not have detailing data prior to Q2 1993, the detailing stocks are subject to the classic initial condition problem (Heckman 1981). To address this, we have collected quarterly detailing data going back to Q3 1988, when the first statin (Mevacor) was introduced. We use these data to construct the initial values of detailing stock in Q2 1993. Similarly, for the publicity variables, we use the presample period data from Q1 1986 to Q1 1993 to construct the initial values of publicity stocks in Q2 1993.[24]

5.4.3. Endogeneity Problem of Detailing. When estimating a demand model using product-level data, one potential concern is that detailing is endogeneous. From the econometric viewpoint, this would be an issue if there were unobserved demand shocks that firms observe before setting their detailing efforts (e.g., Ching and Ishihara (2010)). In most existing works of estimating the demand model for prescription drugs, the unobserved demand shocks are usually a result of omitted information, such as the release of new clinical trials results and news shocks (e.g., Azoulay 2002,

Ching and Ishihara 2010, and Ching et al. 2016). Here, we have collected data on clinical trials and news coverage to control for these two sources of demand shocks. As a result, the endogeneity problem should be alleviated.[25]

We should note that we do not include data on journal advertising. Assuming these marketing activities are positively correlated with detailing,[26] our estimated effect of detailing would include some of the journal advertising effect as well. Hence, one should be cautious when interpreting our results. Another potential piece of missing information is direct-to-consumer advertising (DTCA). We did collect data on DTCA in Canada. But, for most of the periods, pharmaceutical firms did not do DTCA in the Canadian statin market. This is probably because Canada's regulations on DTCA are much stricter than those in the United States. Therefore, we decided not to include DTCA in this study.

5.5. Identification

In this subsection, we provide some intuitions about how the parameters of our model can be identified.

We can identify the parameters in the adoption decision stage (γ_0, γ_i, γ_{lc}, γ_{rh}, γ_{se}) because we observe the total demand for the whole statin category and all the explanatory variables that enter the adoption decision over time. Note that the adoption decision is modeled as a logit model, and hence, the identification of this stage of the model is standard.

To understand how to identify the correlation term in the initial prior, it is important to stress that the quality signals from clinical trials are observable to us. The initial prior in our model captures the physicians' belief prior to the release of any landmark clinical trials. Because drugs entered the market at different points in time, the existing stocks of landmark clinical trials faced by them also differ when they entered the market. As long as $\rho > 0$, the entry date prior beliefs will differ across drugs. Hence, after controlling for the observed quality signals from clinical trial outcomes, q_j^c's, and detailing stocks, the differences of initial sales across drugs help us identify ρ (because it is a function of the entry date prior beliefs).

Correlations in the initial prior beliefs (ρ) can be identified from the observed (to researcher) quality signals on efficiency ratios from clinical trial outcomes and the timing of each clinical trial release as well as the changes in relative market shares of statins before and after the release of each clinical trial. In identifying the correlation parameter, the observed quality signals play a pivotal role. This is because the change in market shares before and after the release of a clinical trial can be influenced by both the realized quality signal from the clinical trial and the extent of correlated learning. For example, if a drug does not gain relative market share after the release of its own clinical trial, there are two possible explanations: (i) the realized quality signal from the clinical trial is the same as the physicians' current perceived quality for the drug, and there is no correlated learning (i.e., $\rho = 0$), or (ii) the realized quality signal is higher than the physicians' current perceived quality, but the extent of correlated learning is extremely high (i.e., $\rho \simeq 1$); consequently, physicians update their prior beliefs about the qualities of both drugs by the same amount. By explicitly using the information reported in a clinical trial, we can observe the realized quality signals. This is how we can tell which explanation plays a bigger role and, hence, identify the correlation parameter. Note that, in principle, one can allow ρ to be drug specific. However, in the current application, we have decided to restrict ρ to be the same across drugs mainly because there are only 14 landmark clinical trials for statins in our sample period.

The parameters that determine the persuasive (κ_d, δ_{d_per}) and informative detailing (α_d, α_{d_c}, δ_{d_inf}) can be separately identified because we assume that clinical trial outcomes only affect the informative detailing, and we explicitly use the information from clinical trials. As a result, clinical trials provide exclusion restrictions needed to disentangle the persuasive and the informative effects of detailing. It is worth emphasizing that clinical trials differ in terms of (a) which drugs they study, (b) number of subjects (patients), (c) reported mean efficiency ratio, and (d) release time. All of these would change the well-informed physicians' expected utility for the drug studied in the clinical trial, and this would, in turn, change the impact of informative detailing on market shares. The identification of publicity variables in the brand choice stage is more straightforward because they only affect the measure of well-informed physicians, who determine the market shares.

The persuasive detailing provides us with an additional channel to explain the rise and decline of a drug. It should be highlighted that this alternative channel does not rely on correlated learning. Hence, the impact of persuasive detailing on physicians' expected utility does not depend on landmark clinical trials and publicity. These are the exclusion restrictions that help us separately identify the effect of persuasive and informative detailing.

6. Results

6.1. Parameter Estimates

Table 6 shows the parameter estimates. Let us first look at the parameters in the utility function. We find the coefficient for the perceived q^h (ω) to be positive and significant. We also find that the persuasive detailing parameter (κ_d) is positive and significant. This indicates that, other than its informative role and correlated learning, persuasive detailing is also responsible for explaining each statin's demand. In fact, as we show in our counterfactual experiments 3 and 4, persuasive detailing plays a strong role in the success of Lipitor and Crestor.

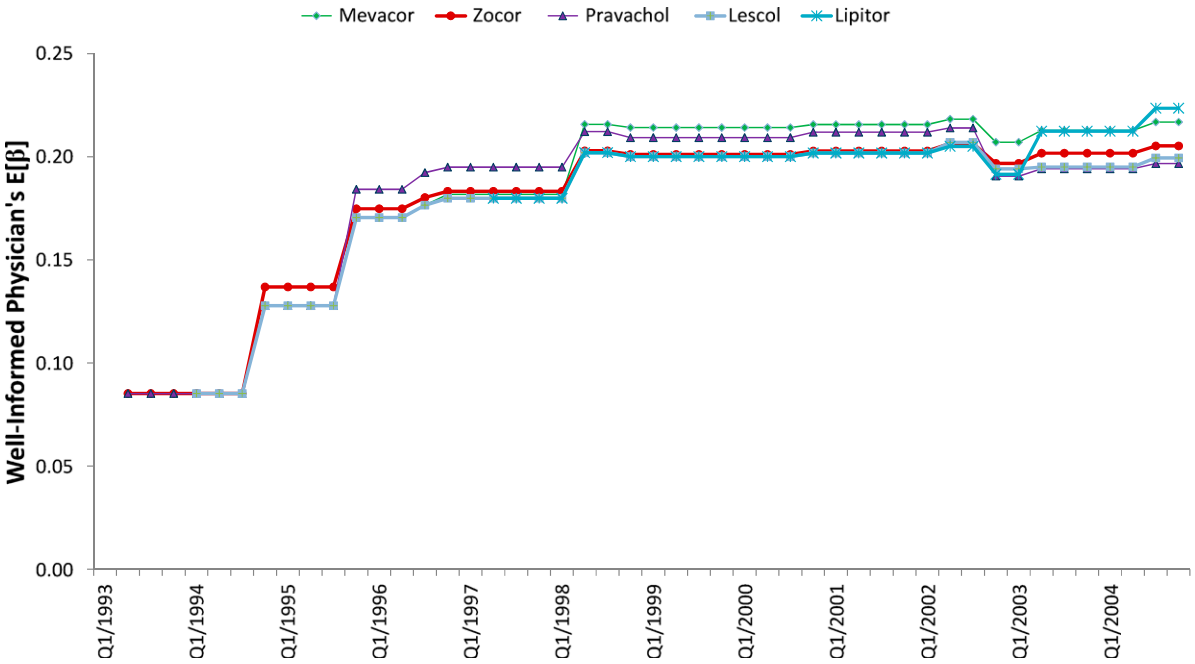
Table 6. Parameter Estimates for the Benchmark Model

Variable descriptions	Estimates	Standard error
Statin choice stage		
Learning parameters		
$\underline{\beta}$ (initial prior belief on efficiency ratio)	0.0851	0.0210
$\underline{\sigma_\beta}^2$ (initial prior variance on efficiency ratio)	1.2435	0.0962
$\sigma_\varepsilon{}^2$ (signal variance from different design)	0.5227	0.2069
$\sigma_\zeta{}^2$ (signal variance from 1,000 patients)	5.4968	2.2340
ρ (correlated learning parameter in initial prior)	0.8247	0.0607
Parameters determining measure of informed physicians		
α_0 (constant)	–5.7250	0.6151
α_d (informative detailing)	2.5943	0.2180
α_{d_c} (informative detailing of competitors)	1.3840	0.0562
α_{lc} (informative publicity in lowering cholesterol levels)	0.0216	0.3156
α_{lc_c} (informative publicity of competitors in lowering cholesterol levels)	–0.0094	0.0434
α_{lc_g} (informative general publicity in lowering cholesterol levels)	0.3865	0.1662
α_{rh} (informative publicity in reducing heart disease risks)	1.2032	0.1640
α_{rh_c} (Informative publicity of competitors in reducing heart disease risks)	0.2504	0.3271
δ_{rh_g} (Informative general publicity in reducing heart disease risks)	0.1176	0.1986
δ_{d_inf} (carryover rate of informative detailing in statin choice)	0.8999	0.0143
δ_{rh_inf} (carryover rate of informative publicity in statin choice)	0.2142	0.0290
Utility parameters		
ω (coefficient of perceived quality in reducing heart disease)	2.0144	0.3112
κ_d (persuasive detailing)	1.0735	0.0828
δ_{d_per} (carryover rate of persuasive detailing in statin choice)	0.9272	0.0077
Brand dummies		
Zocor	1.2683	0.0352
Pravachol	1.0809	0.0560
Lescol	–0.1991	0.2524
Lipitor	1.7644	0.0175
Baycol	0.3259	0.1135
Crestor	1.0314	0.0610
Adoption decision stage		
γ_0 (constant)	–6.9597	0.1059
γ_1 (inclusive value)	1.0395	0.0385
γ_{lc} (general publicity stock in lowering cholesterol levels)	0.0434	0.0979
γ_{th} (general publicity stock in reducing heart disease risks)	0.3272	0.1095
γ_{se} (general publicity stock in side effects)	–0.0297	0.0152
δ_p (carryover rate of publicity in adoption decision)	0.9262	0.0053
Additional parameter		
Standard deviation of e_{jt} (in hundred thousand)	0.2319	0.0111
Log likelihood	–2,695.46	

The first section in the table describes learning parameters. Physicians' initial prior mean on efficiency ratio is 0.085.[27] As shown in Table 2, most signals on the efficiency ratios from landmark clinical trials are between 0.1 and 0.3. Therefore, it appears that physicians' initial prior belief on statins' efficiency ratios is quite low compared with the true efficiency ratios. The initial prior variance (σ_β^2), signal variance from different clinical trial designs (σ_ϵ^2), and the signal variance per 1,000 patients (σ_ζ^2) are all statistically significant. The initial prior correlation on efficiency ratio (ρ) is 0.825. This implies that if one statin receives a new clinical trial result, physicians will also use its information to update their beliefs about other statins.

To demonstrate the rate of learning, in Figure 8, we graph the posterior belief of a well-informed physician (i.e., $E[\beta_{jt}]$) over time based on our estimated model. Recall that a well-informed physician has learned about all the clinical trial results available up to time t. The figure shows that the physician updates the physician's beliefs about all the statins whenever a landmark clinical trial is released. Before Q4 1994, there were no landmark clinical trials to support statins' efficacy in reducing heart disease risks. Hence, the physician has exactly the same prior belief about Mevacor, Zocor, and Pravahol up to Q4 1994. In Q4 1994, Zocor received the first landmark clinical trial (4S study) supporting its efficacy in reducing heart disease risks. Then, the well-informed physician updates the physician's beliefs about all statins (not just Zocor). However, because the information spillover is not 100%, the physician's belief on Zocor is slightly higher than those on other statins. We can see similar imperfect information spillover patterns happened when other clinical trials are released.

Figure 8. (Color online) The Posterior Beliefs of Well-Informed Physicians



Before Lipitor releases its own landmark clinical trials (i.e., before Q2 2003), $E[\beta_{Lipitor,t}]$ is the lowest among all existing statins. However, after the second landmark clinical trial of Lipitor is released in Q3 2004, $E[\beta_{Lipitor,t}]$ became the highest among all statins. Figure 8 suggests that Lipitor benefits much from other statins' investment in landmark clinical trials. But it appears that there is still room for Lipitor to improve from its own investment in landmark trials. In one of our counterfactual experiments, we investigate this further by examining how the evolution of $E[\beta_{j,t}]$ translates to sales.

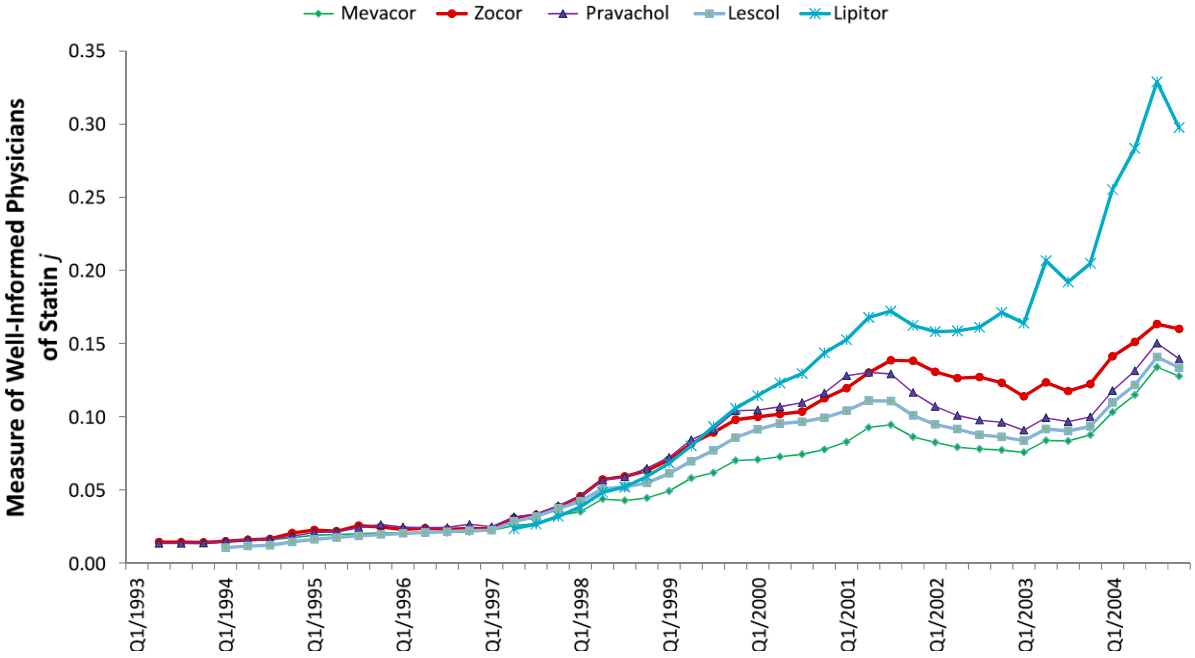
As expected, we find that firms' own detailing goodwill stock (α_d) has a positive and significant effect on sales. The results indicate that detailing plays an informative role in physicians' prescription choices. Brand-specific publicity in reducing heart disease risks (α_{rh}) is also positive and significant, indicating that it can help inform physicians (perhaps the pressure of patients prompts physicians to look into the latest clinical evidence of a drug) about the drug's clinical trials. This could happen if patients who are exposed to publicity in heart disease risks encourage their physicians to read clinical trial results. The results here are consistent with the findings from the reduced-form analysis done in Ching et al. (2016).

We also find evidence that rivals' detailing could help inform physicians about a drug's clinical trials (α_{d_c} is positive and significant). This is consistent with correlated learning (or information spillover). For instance, when its own landmark clinical trials are still underway, Lipitor may want to take advantage of correlated learning and use its detailing to inform physicians about the landmark trials of other drugs. This is one way to enhance its late-mover advantages. Note that this result is not reported in Ching et al. (2016) because their reduced-form model does not consider the possibility that a drug's own clinical trial stock can interact with rivals' detailing.

Figure 9 shows how the measure of well-informed physicians changes over time by drug. For all statins, the evolutions appear to be in an S-shape. For the earlier half of the sample period, the measure of physicians well informed about Lipitor is lower than those who are well informed about Zocor and Pravachol. This is because Lipitor is a late entrant, and it takes time for Lipitor to build up a stock of physicians who are well informed about Lipitor. But in Q4 2001, Lipitor has more well-informed physicians than other statins, and its measure rises above 0.3 by Q4 2004. As a comparison, its next competitors, Zocor and Pravachol, reach around 0.16 and 0.14, respectively, by Q4 2004.

Next, we discuss parameters in the adoption decision stage. The estimate of the inclusive value term (γ_i) is positive (1.04) and significant. This indicates that detailing and publicity stocks in the brand choice stage have positive influence on the adoption decision in the first place. The stock of general publicity in reducing heart disease risks (γ_{rh}) is also estimated to be positive and significant, but the stock of lowering cholesterol (γ_{lc}) is insignificant. However, we do not want to read too much into the estimates of these two variables because these two stock variables are highly correlated. Interestingly, the coefficient on the stock of general publicity in side effects (γ_{se}) is negative and significant, indicating that consumers are generally worried about side effects. Figure 10 shows the goodness-of-fit of our

Figure 9. (Color online) Measure of Well-Informed Physicians of Statin j



model. In general, our estimated model is able to fit the data quite well.

6.2. Extension and Remarks

In this subsection, we consider a version of the model in which patients (and, hence, physicians) directly care about q_j^c and not all physicians have complete information about q_j^c. The utility of patient i who consumes drug j at time t is given by

$$U_{ijt} = \omega_h \cdot q_j^h + \omega_c \cdot q_j^c + \lambda_j + \epsilon_{ijt}, \tag{32}$$

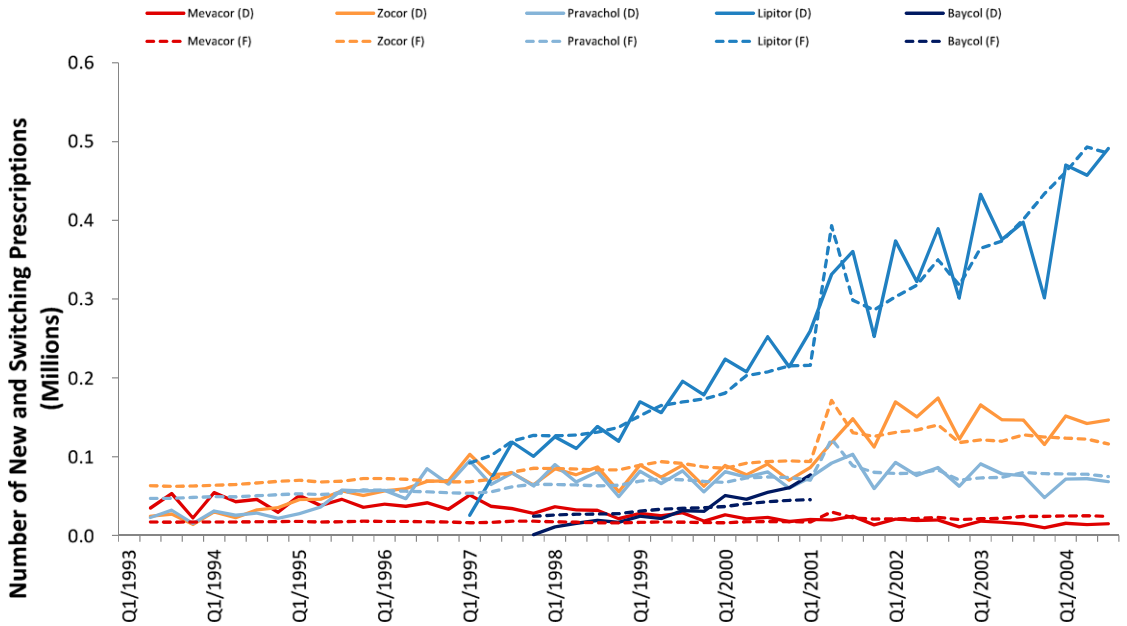
where q_j^h and q_j^c denote drug j's efficacy in reducing heart disease risks and lowering LDL, respectively. Hence, physician k's expected utility of prescribing drug j to patient i at time t is

$$E\left[U_{ijt}|I(k,t)\right] = \omega_h \cdot E\left[q_j^h|I(k,t)\right] + \omega_c \cdot E\left[q_j^c|I(k,t)\right] + \kappa_d \cdot P\_STK\_detail_{jt} + \lambda_j + \epsilon_{ijt}, \tag{33}$$

where

$$E\left[q_j^h|I(k,t)\right] = E\left[q_j^c|I(k,t)\right] \cdot E\left[\beta_j|I(k,t)\right]. \tag{34}$$

Figure 10. (Color online) Fit: Actual and Simulated Prescription Volume



Like the original model, physicians are either well informed or uninformed about the most updated information about a drug. But in this version, it applies to both q_j^c and q_j^h. The measure of well-informed physicians is still determined in the same way explained in the original model. If a physician is well informed about drug j, the physician knows the true q_j^c and the most updated information about q_j^h. If a physician is uninformed about drug j, the physician only knows the initial priors, $\underline{q}_j^c$ and $\underline{q}_j^h$. This introduces another set of initial prior parameters, $\underline{q}_j^c$, to be estimated.

We estimate this extension, and the results are reported in Table 7. In general, the results are similar to what we find in the original model. In particular, we find that the correlated learning parameter is still significant and at the magnitude similar to our benchmark model. However, we also see that the utility weight for q_j^c is not significant. This provides a justification for our original model, in which we assume that patients and physicians primarily care about q_j^h.

We should note that an alternative way to model learning about q_j^c is to follow Ching (2010b) and assume physicians slowly learn about it from their patients' experience signals. Because the signals are unobserved to researchers, we need to use Monte Carlo methods to integrate them when constructing the likelihood. This significantly increases the computational burden. Moreover, with product-level data and unobserved information signals, allowing heterogenous information sets across physicians is computationally infeasible because it is very difficult to track how they evolve over time. This problem requires us to impose the assumption that physicians share a homogenous information set and use the same set of signals for updating as in Ching (2010b) or assume that every physician receives the same number of signals in each period as in Narayanan et al. (2005). On the contrary, the approach we adopt here is more tractable and allows us to introduce physician heterogeneity in their information sets with respect to q_j^c. These are the reasons why we choose this approach here.

6.2.1. Remarks. When Lipitor just entered the market, it received a very high amount of detailing. Such a pattern is very common when a drug just enters the market. The high initial detailing has three effects: (i) it quickly builds up Lipitor's persuasive detailing stock (which partly captures the awareness effect), (ii) it quickly builds up the measure of well-informed physicians for Lipitor, and (iii) it also helps build up the measure of well-informed physicians for other statins. The first channel directly helps the demand for Lipitor. The second channel gets Lipitor prepared for the future release of its own landmark clinical trials. The last channel indirectly helps Lipitor because of correlated learning.

As pointed out in Simon and Kotler (2003), Pfizer/Warner–Lambert launched an intensive education campaign in the United States when Lipitor entered. This campaign could have generated more media coverage about the risk of high LDL and statins' ability to reduce LDL. Such media coverage could have been viewed in Canada. We are able to capture this channel because our publicity variables include not only Canadian sources, but also large U.S. newspapers and internet news sources.

We do observe that the introduction of Lipitor followed a steady increase in publicity related to statins (measured by our general publicity variable) and a steady flow in publicity related to Lipitor regarding its effectiveness in reducing LDL. The introduction of Lipitor may have led to an increase in general publicity for statins, which leads to category expansion. At the same time, it also leads to an increase in brand-specific publicity (especially for Lipitor), and that has led to further market expansion (captured via the inclusive value term in the first stage of our model). When it gets to the brand choice stage, the vast majority of the market expansion goes to Lipitor.

6.3. Counterfactual Experiments

Now we turn to counterfactual experiments. Because our results remain robust under the extension we just considered and q_j^c is not significant in the utility function, we decide to use our original model to conduct the counterfactual experiments. In the first two counterfactual experiments, we investigate the impact of one's own landmark clinical trials for both late entrants and early incumbents by removing them. In the last two counterfactual experiments, we investigate the impact of correlated learning on the statin market by removing it.

6.3.1. Experiment 1: Quantifying the Return of Lipitor's Landmark Clinical Trials. To design a clinical trial to show a drug's efficacy in reducing heart disease risks, researchers needs to follow up on thousands of patients for a few years. As discussed earlier, such a large sample is required because heart attacks, strokes, etc., are rare events. Therefore, sponsoring such a clinical trial is a very large investment for the firm. If physicians can indirectly learn about the ability of a new statin in reducing heart disease risks through incumbents' clinical trials, it might not be worthwhile for a late entrant to sponsor a landmark clinical trial for its own statin. This could be the case for Lipitor. Prior to its entry in 1997, several incumbent firms had already obtained landmark clinical results for reducing heart disease risks. Lipitor obtained its own landmark clinical trial results several years after its introduction in 1997. Were the landmark clinical trials for its own drug, Lipitor, worth the investment of the drug company?

Table 7. Parameter Estimates for the Extended Model

Variable descriptions	Estimates	Standard error
Statin choice stage		
Learning parameters		
$\underline{\beta}$ (initial prior belief on efficiency ratio)	0.0777	0.0193
σ_β^2 (initial prior variance on efficiency ratio)	1.2200	0.0967
σ_ε^2 (signal variance from different design)	0.5349	0.2086
σ_ζ^2 (signal variance from 1,000 patients)	5.5081	2.3210
ρ (correlated learning parameter in initial prior)	0.8258	0.0642
Initial priors on efficacy in lowering cholesterol levels		
Mevacor	1.6662	0.5516
Zocor	1.6360	0.3061
Pravachol	1.2312	0.5158
Lescol	1.1528	0.6955
Lipitor	2.2321	0.4991
Baycol	2.2073	0.9468
Crestor	2.4352	0.5280
Parameters determining measure of well-informed physicians		
α_0 (constant)	–5.7230	0.4238
α_d (informative detailing)	2.5831	0.2284
α_{d_c} (informative detailing of competitors)	1.4021	0.0581
α_{lc} (informative publicity in lowering cholesterol levels)	0.0068	0.3339
α_{lc_c} (informative publicity of competitors in lowering cholesterol levels)	–0.0122	0.0445
α_{lc_g} (informative general publicity in lowering cholesterol levels)	0.3196	0.2134
α_{rh} (informative publicity in reducing heart disease risks)	1.2007	0.1641
α_{rh_c} (informative publicity of competitors in reducing heart disease risks)	0.2504	0.3383
α_{rh_g} (informative general publicity in reducing heart disease risks)	0.0920	0.2132
δ_{d_inf} (carryover rate of informative detailing in statin choice)	0.8999	0.0130
δ_{rh_inf} (carryover rate of informative publicity in statin choice)	0.2139	0.0290
Utility parameters		
ω (coefficient of perceived quality in reducing heart disease risks)	2.0043	0.2598
ω_c (coefficient of perceived quality in lowering cholesterol levels)	–0.0191	0.0590
κ_d (persuasive detailing)	1.0647	0.1089
δ_{d_per} (carryover rate of persuasive detailing in statin choice)	0.9265	0.0078
Brand dummies		
Zocor	1.2152	0.0584
Pravachol	1.0211	0.0668
Lescol	–0.2171	0.2620
Lipitor	1.7417	0.0232
Baycol	0.3292	0.1218
Crestor	1.0531	0.0652
Adoption decision stage		
γ_0 (constant)	–7.0022	0.1031
γ_i (inclusive value)	1.0854	0.0453
γ_{lc} (general publicity stock in lowering cholesterol levels)	0.0392	0.0522
γ_{rh} (general publicity stock in reducing heart disease risks)	0.3546	0.1058
γ_{se} (general publicity stock in side effects)	–0.0289	0.0065
δ_p (carryover rate of publicity in adoption decision)	0.9250	0.0057
Additional parameter		
Standard deviation of e_{jt} (in hundred thousand)	0.2285	0.0113
Log likelihood	–2,692.00	

To shed some light on this question, we use our model to forecast the demand for Lipitor in a counterfactual situation in which we shut down Lipitor's landmark clinical trials in studying its efficacy on reducing heart disease risks. Figure 11 graphs the benchmark and counterfactual total demand for statins. Without its own landmark clinical trials, the counterfactual demand for Lipitor is 2%–5.8% lower than the benchmark demand for most quarters from Q2 2003 to Q4 2004 (note that the first landmark clinical trial for Lipitor was released in May 2003). The counterfactual total demand for Lipitor is about 150,000 prescriptions lower than the benchmark counterpart in Q4 2004. This is roughly about 5.8% of the actual total demand in Canada.

The total change in the Canadian market might appear to be small. However, Lipitor's global annual sales were almost $10.9 billion in 2004.[28] Therefore, even just 5.8% loss in sales would cost about $632 million per year. According to the U.S. Department of Health and Human Services, the average cost of phase 4 clinical trials

(i.e., postmarketing clinical trials[29]) for cardiovascular drugs is $27.8 million.[30] We drop two landmark clinical trials for Lipitor in this counterfactual experiment. Hence, Lipitor's investment in its own postmarketing clinical trials appears to be justifiable from a profits viewpoint.

It is also worth pointing out that the counterfactual demand for other drugs also drops but at a smaller magnitude. For Crestor, the new and switching patients' demand decreases by 4%–6.15% during Q2 2003 to Q4 2004. For Zocor and Pravachol, it also drops but by an even smaller percentage. Interestingly, we see negative effects on other drugs because we take away the positive externality generated by the clinical trials of Lipitor. The reason why the effects are asymmetric is because (i) the information spillover is not perfect and (ii) β works like a multiplier, and how a change in β affects a drug's q_j^h (and ultimately its demand) depends on its q_j^c. Because $q^c_{crestor} > q^c_{zocor} > q^c_{pravachol}$, the magnitudes of the changes in demand across drugs reflects this order. The results here also echo our reduced form results in Section 4.6.

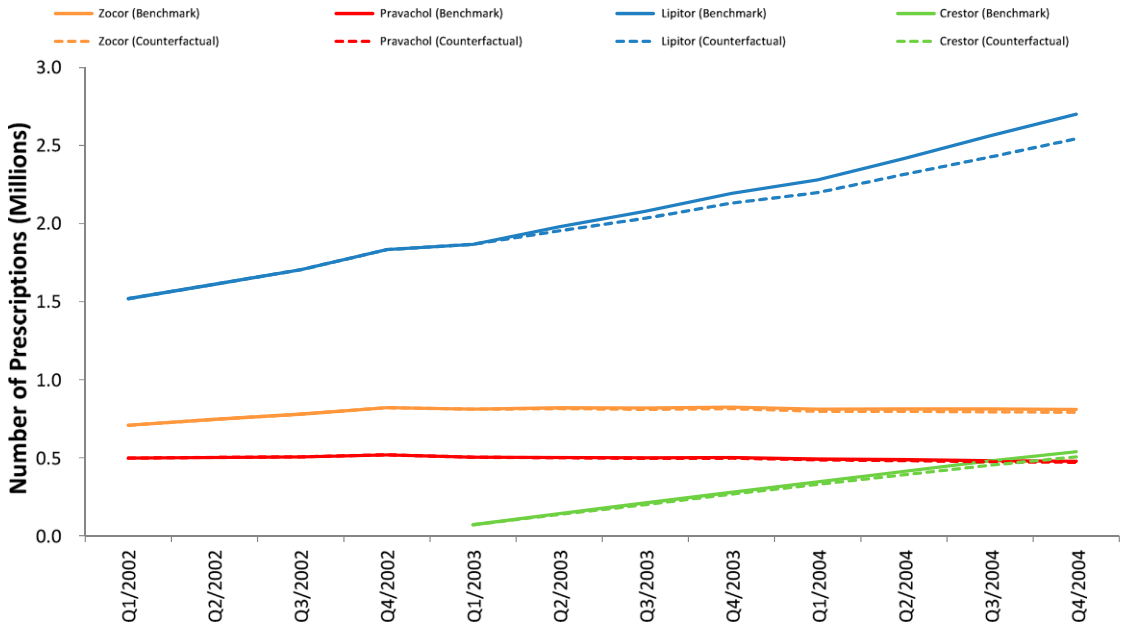
6.3.2. Experiment 2: Removing Incumbents' Landmark Clinical Trials. Our results show that late entrants can free ride on the landmark clinical trials done by incumbent firms. If the late entrants are also more effective in lowering cholesterol levels than incumbent drugs, they may be able to steal a significant portion of demand from incumbents. This leads to the following question: Did incumbents benefit from the results of their landmark clinical trials or were they hurt by the results of landmark clinical trials? To answer this question, we conduct two counterfactual exercises: (i) shutting down the landmark clinical trials for Mevacor and (ii) shutting down the landmark clinical trials for Pravachol. In these exercises, we can see which drugs benefited most or least from incumbents' clinical trials. The results on the total demand are shown in Figure 12.

As this figure shows, all statins would lose some patients without the landmark clinical trials for Mevacor. In other words, the absence of landmark clinical trials done by the early entrants would likely shrink the whole statin market, including the incumbents. The number of prescriptions would decrease by 0.1%–1.8% for Mevacor, 0.1%–1.3% for Pravachol, 0.3%–2.2% for Lipitor, and 1.8%–2.0% for Crestor. It turns out that the incumbents' clinical trials benefit Lipitor and Crestor most because they are the strongest in q_j^c. The counterfactual demand for Mevacor would be about 100 to 2,100 prescriptions per quarter (on average 1,100) lower than the benchmark counterpart. Compared with Lipitor's case in counterfactual experiment 1, the demand loss for Mevacor is very small.

We also conduct another counterfactual exercise by eliminating the landmark clinical trials for Pravachol. The results are qualitatively similar to the Mevacor case. More precisely, the number of prescriptions would decrease by 0.1%–2.1% for Mevacor, 0.1%–2.0% for Pravachol, 1.1%–3.3% for Lipitor, and 1.4%–1.6% for Crestor. The counterfactual demand for Pravachol would be about 300 to 10,100 prescriptions—on average 5,000 prescriptions—lower than the benchmark counterpart. Again, the reduction in demand is much smaller than that for Lipitor.

Based on these results, although Mevacor and Pravachol still benefit from their own landmark clinical trials, it seems unlikely that such small benefits can justify their investments (note that there were two and

Figure 11. (Color online) Counterfactual Experiment 1 (Removing Lipitor's Landmark Trials)



six landmark clinical trials for Mevacor and Pravachol, respectively).

6.3.3. Experiment 3: Removing Correlated Learning. Our estimation results suggest that (i) there is information spillover of landmark clinical trial results across drugs, and (ii) Lipitor (and Crestor) can gain a late-mover advantage by free riding on the information provided by its rivals' clinical trials. Therefore, we are interested in quantifying the importance of correlated learning. What would happen if Lipitor were not able to free ride on the information from the clinical trials conducted by other drug companies? To answer this question, we generate the demand for each statin in a counterfactual situation in which there is no correlated learning. Under this counterfactual experiment, we set the correlated learning parameter ρ to be zero. Here, we only study the impact on new and switching patients because our model assumes the switching rate is exogenous (and, hence, the nonswitchers will not be affected by this experiment). Figure 13 presents the benchmark and counterfactual demand from new and switching patients. Lipitor's counterfactual demand is about 2%–10% lower than its benchmark demand from Q2 1997 to Q1 2003. After Lipitor's first landmark clinical trial is launched in Q2 2003, the difference between the counterfactual and benchmark demand sharply drops and then quickly converges to about 4% in Q4 2004 (which amounts to about 19,400 prescriptions). For Crestor, the counterfactual demand is about 8%–25% lower than the benchmark demand during Q1 2003 to Q4 2004. The differences are substantial and indicate that correlated learning plays a role in the early success of both Lipitor and Crestor.

However, note that, even without correlated learning, the demand for Lipitor from new and switching patients keeps increasing over time. This suggests that correlated learning is not the only driving force for the early success of Lipitor. So what else can contribute to its success? Because Lipitor had the highest q^c before Crestor was launched, Lipitor could still have a high $E[q^h]$ even with a relatively low $E[\beta]$ (see Equations (6) and (7)). To investigate this possibility, in Figure 14, we graph the well-informed physicians' $E[q^h]$ over time under the condition that there is no correlated learning. It shows that, in the absence of correlated learning, the $E[q^h]$ for Lipitor is lower than those for Mevacor, Zocor, and Pravachol up until Q2 2003. Hence, we can rule out this explanation.

6.3.4. Experiment 4: Removing Correlated Learning and Persuasive Detailing. We then investigate if persuasive detailing could be another reason for Lipitor's success. We take it a step further by removing both correlated learning and persuasive detailing. We show the results in Figure 15. The difference between this figure and Figure 13 demonstrates that persuasive detailing plays an important role in the success of Lipitor and Crestor. It is interesting to note that, even after we remove both correlated learning and persuasive detailing, we still see that Lipitor gets more new patients and switchers than other drugs. This is due to Lipitor's brand dummy, which has the highest value among all brand dummies. The larger brand dummy of Lipitor may reflect the values of other unobservables (e.g., Lipitor's price is slightly lower than other statins'). Recall that the switching rate is very low. Lipitor's high brand dummy helps get more new patients in each

Figure 12. (Color online) Counterfactual Experiment 2 (Removing Mevacor's Landmark Trials)

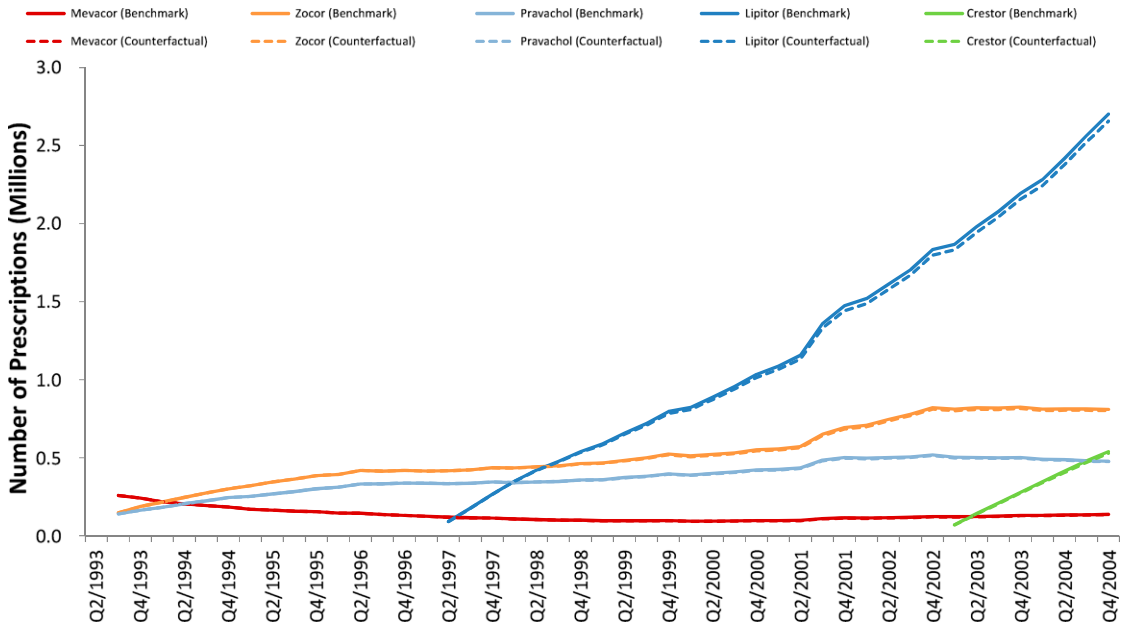
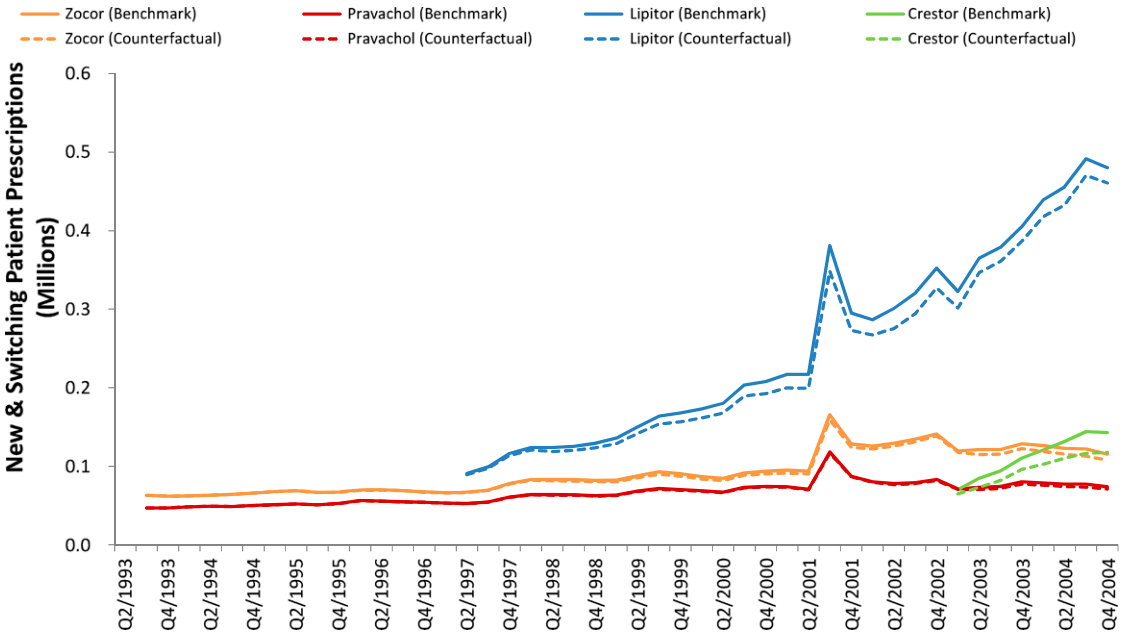


Figure 13. (Color online) Counterfactual Experiment 3 (Removing Correlated Learning)



period, and with the low switching rate, it builds up its patient base by accumulating the new demand over time.

7. Conclusion and Future Research

We develop a new demand model with correlated learning in which consumers can observe some attributes of a product, but they are uncertain about how efficiently the product can translate them to completing tasks (or generating output) that consumers care about. For products that rely on a similar production technology, we argue that correlated learning and information spillover can happen at the production-function level, and this can lead to late-mover advantages.

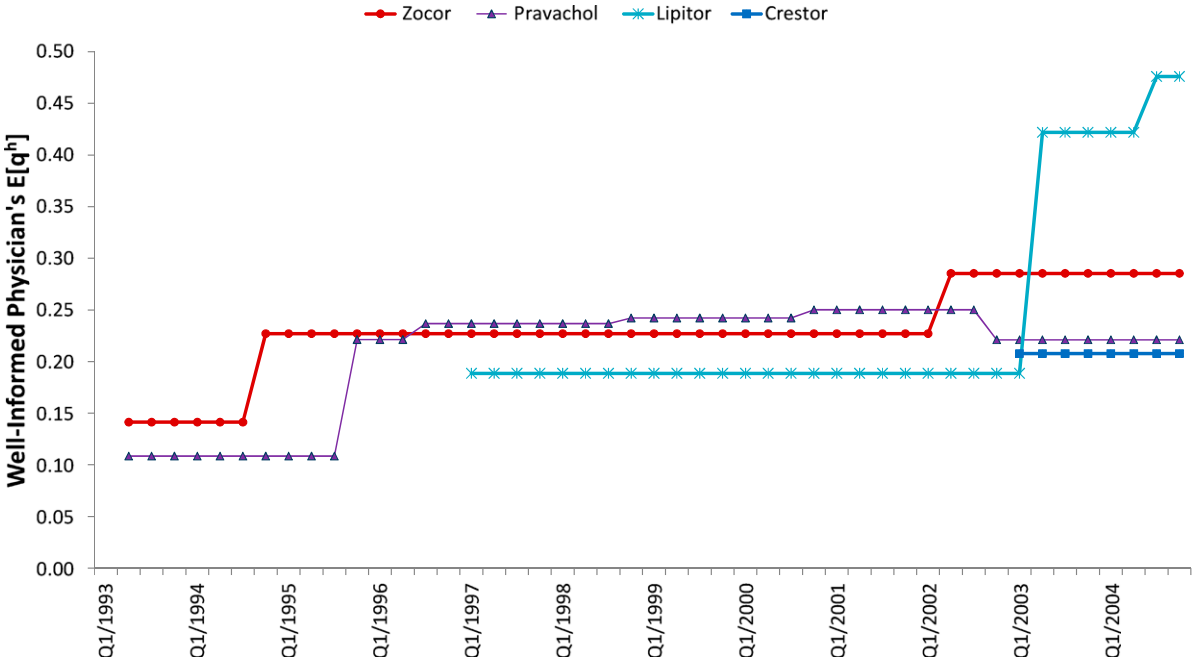
We apply our model to the statin market, in which physicians/patients need to infer a statin's efficacy in reducing heart disease risks based on its observed efficacy in lowering LDL. This relationship is captured by the efficiency ratio parameter, about which physicians are uncertain. We allow physicians' initial prior perceptions about the efficiency ratio to be correlated across drugs, and physicians learn about the efficiency ratio for each drug from landmark clinical trials via informative detailing and publicity. Our estimation results show that the initial prior perceptions on the efficiency ratio are positively correlated. Because of the information spillover, late entrants (especially for Lipitor and Crestor) can benefit from incumbents' landmark clinical trials.

Unlike the previous structural learning literature in which researchers do not observe quality signals, we are able to observe them from clinical trial results. Observing the information signals helps us identify the strength of correlated learning and disentangle the informative and persuasive effects of detailing, Moreover, it greatly reduces the burden of computing the likelihood in our estimation procedure.

Although we find that late entrants are able to gain advantages in the statin market, we are not advising managers that they should postpone launching their products. This is because such advice ignores the importance of switching costs, which allow early entrants to gain first-mover advantages by locking in early adopters of statins. Another point to note is that pharmaceutical companies usually file for patent protections for their drugs before gaining approval for selling them. Because patent protection expires after a fixed period of time, delay in launching a drug shortens its effective patent life, which is another important factor to consider. To find out the optimal entry time, it is important to evaluate these trade-offs. A complete analysis is beyond the scope of this paper because it requires us to understand the nature of switching costs and model the future demand for the drugs (taking into consideration the remaining patent life and generic competition afterward) under different entry times. We leave this for future research.

We should also note that our model does not allow for intrinsic sales force heterogeneity, which could be partly responsible for the success of Lipitor. However, as far as we know, the existing structural models of pharmaceutical marketing based on product-level data all make this assumption. In the statin market, all brand-name drugs are marketed by large pharmaceutical companies. It seems likely that the sales reps for these companies all receive high-quality training. Hence, we feel that intrinsic sales force heterogeneity may not be of first order importance. To allow for this

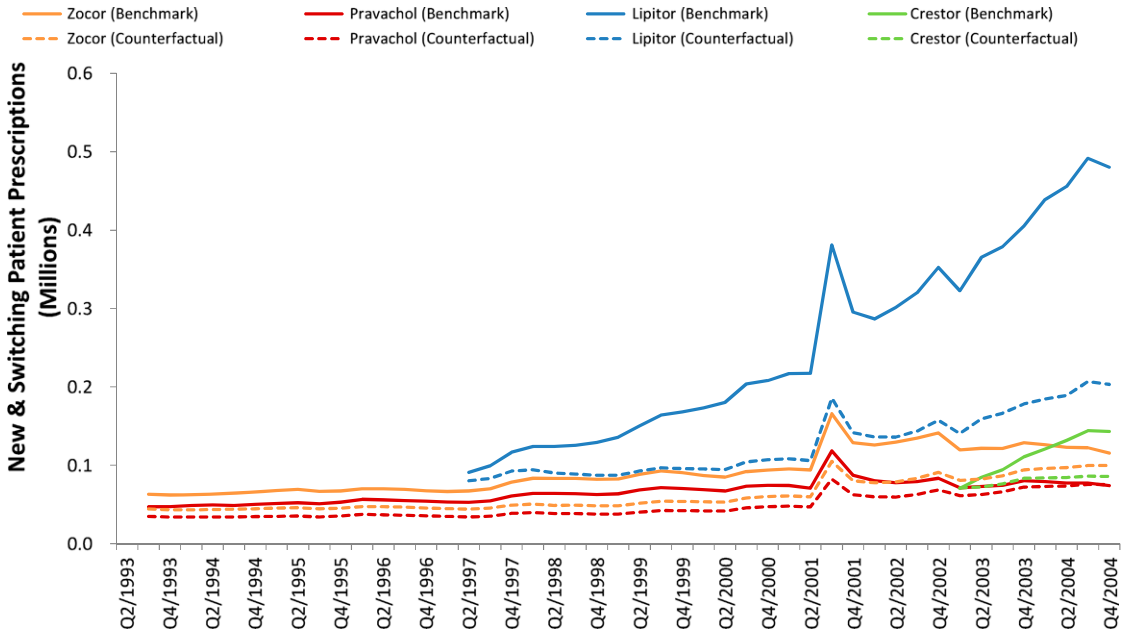
Figure 14. (Color online) Well-Informed Physician's E[q^h] over Time (Counterfactual Experiment 3)



possibility, one would need to use individual-level data. This should be another interesting future research topic.

Last but not least, we should emphasize that the model developed here could be applied to settings other than prescription drugs. For instance, when the iPhone entered the market, it was a very innovative product. The heavy promotion done by Apple has informed a large population about what a touchscreen phone can accomplish (like a mini-computer). When Samsung entered the market later (as a late entrant), it leveraged what consumers already knew about basic ideas of this product and launched its own android-based phones with larger screen sizes, more memory, and faster processors. Some consumers may then have inferred that a Samsung phone is better than an iPhone. This can explain why Samsung has quickly become the largest smartphone player even though it entered the smartphone market after Apple. Other examples may include electric cars, droids, and so forth.

Figure 15. (Color online) Counterfactual Experiment 4 (Removing Correlated Learning and Persuasive Detailing)



Acknowledgments

The authors thank the department editor, Matthew Shum, and the review team for constructive comments, which led to a significant improvement of the paper. They thank Qiang Liu, Ig Horstmann, Susan Mendez, Minjae Song, Demetrios Vakratsas, and Kevin Williams; as well as seminar participants at Stanford University, Emory University, Penn State University, Fudan University, Western University, Peking University, Erasmus University, CKGSB, Chinese University of Hong Kong, University of Hong Kong, HKUST Marketing Research Camp, Annual Health Econometric Workshop, FTC Microeconomics Conference, Annual Marketing Science Conference, Workshop on Physicians Prescription Behavior at Toulouse School of Economics, University of Calgary's Empirical Microeconomics Workshop, Empirical and Theoretical Symposium, Annual Workshop on Industrial Organization and Management Strategy, and the Annual Health Economics Conference, for helpful comments. All remaining errors are ours. Both authors have made equal contribution to this research. This paper was previously circulated under the title "A Structural Analysis of Detailing, Publicity and Correlated Learning: The Case of Statins."

Endnotes

[1] For instance, a recent clinical trial shows that a new anticholesterol combination drug, Vytorin, does not reduce heart disease risks even though it is very effective in lowering bad cholesterol (LDL) and raising the good cholesterol (HDL) (Park 2008).

[2] Landmark clinical trials we include in this research are phase 4 trials recruiting more than 1,000 participants with treatment duration of at least two years.

[3] Pharmaceutical detailing is a direct-to-physician marketing activity wherein sales representatives visit physicians and explain the efficacies and side effects of a drug.

[4] It is also possible that patients who have been exposed to publicity about a drug ask their physicians about the drug. Such inquiries could motivate physicians to look up clinical evidence for that drug.

[5] Liu et al. (2015) also study the category expansion (spillover) effect of DTCA and meetings and events for statins. However, they do not explicitly model learning, and hence, their modeling framework is different from ours. Ching et al. (2009, 2014, 2019) also provide evidence on category expansion resulting from advertising in consumer package goods.

[6] Baycol was only on the market briefly. When the news about Baycol's potentially strong negative side effects came out (more serious muscle pain and potentially permanent damages to muscle), its manufacturer quickly pulled Baycol off the market.

[7] See https://www.nhs.uk/news/heart-and-lungs/statins-side-effects-are-minimal-study-argues, accessed March 18, 2018.

[8] To convert from nominal to real dollars for detailing, we use the consumer price index from Statistics Canada.

[9] A discontinuing user is one who filled a statin prescription in the previous quarter but did not fill a prescription for any of the statins in the current quarter. Similarly, a switcher is one who filled a statin prescription in the previous quarter and switched to fill a prescription of a different statin in the current quarter.

[10] In a famous nonlandmark clinical trial, CURVES, which studies the efficacy of four statins with different doses, the largest treatment group only had 63 patients and the follow-up period was only eight weeks.

[11] According to our calculation, the ratio between doctor and high-cholesterol population about is 1:165 in Canada. Moreover, not everyone with high cholesterol takes a statin or uses medication to control cholesterol levels. Furthermore, without observing a control group, it is difficult for doctors to draw inference about the efficacy of lowering heart disease risks simply based on patients who are taking statins.

[12] Although Table 3 uses mmol/L as unit of LDL reduction, the unit in Table 2 is mg/dL. Because molar mass of cholesterol is 386.65 g, 1 mmol/L of LDL can be converted to 38.6 mg/dL.

[13] Law et al. (2003) include all double-blind clinical trials reporting mean absolute LDL reductions (mmol/L) in the statin-treated group and in the placebo group from Medline, Cochrane Collaboration, and Web of Science databases. We have also collected data from the CURVES study, which Pfizer used to gain the approval to sell Lipitor. The results on the LDL reduction abilities are consistent with those of Law et al. (2003). However, because CURVES does not report the efficacy of Crestor, we decided not to use the results from CURVES. We should also note that Baycol is not included in the meta-analysis done in Law et al. (2003). Baycol's potential serious side effects led to its early withdrawal. Because we cannot find any meta-analysis that includes Baycol, we set Baycol's q_j^c to be the same as Lipitor's in our analysis. In a robustness check, we set Baycol's q_j^c to be the same as Crestor's (which is higher than Lipitor's). In that case, we found that Baycol's demand increased by slightly more than 4%. But for all other statins, the changes in demand are all less than 0.2%.

[14] For drug-specific publicity, we sometimes encounter articles that compare drugs. Therefore, we further classify drug-specific publicity into *comparison* (*c*) or *noncomparison* (*nc*). But the results of Ching et al. (2016) indicate that comparison publicity does not have much variation. Therefore, we only use noncomparison drug-specific publicity in this paper.

[15] The carryover rate used here is similar to what Berndt et al. (1996) found. We also tried $\delta_d = 0.90$. But the model with $\delta_d = 0.95$ produces significantly better fit, and hence, we decided to set $\delta_d = 0.95$.

[16] Because the switching rate is very low, the vast majority of d_{jt}^{ns} comes from new patients' demand.

[17] We exclude Crestor because it entered the market near the end of our sample period (Q1 2003). There are three landmark clinical trials released after its entry and before our sample ends. Two are ASCOT-LLA and ALERT, which were released in Q2 2003, and the other is CARDS, which was released in Q3 2004. The last observation of our sample period is Q4 2004. Therefore, there is only one observation after CARDS and before ASCOT-LLA and ALERT. Consequently, the number of observations either before or after these clinical trials is not enough to identify the effects of the clinical trials. We also exclude Baycol. Baycol was withdrawn shortly after its introduction because of serious side effects. It was on the market from Q1 1998 to Q2 2001 with 13 quarters of observations. During this period, there were only three landmark clinical trials: AFCAPS/TexCAPS was released in Q2 1998, LIPID was released in Q4 1998, and GISSI was released in Q4 2000. Hence, we do not have enough observations to identify the impacts of these three clinical trials on the demand for Baycol (we have too few observations before LIPID and AFCAPS/TexCAPS and too few observations after GISSI).

[18] Because of price regulations on prescription drug prices in Canada, prices for statins hardly changed during our sample period.

[19] As we mentioned in Section 4.3, the effect on lowering LDL usually shows soon after the statin treatment has started. This is evidenced by the well-known study CURVES, which only follows patients for eight weeks. Hence, given that the length of a period is a quarter in our study, physicians should be able to learn the true q_j^c very quickly in our framework either from nonlandmark clinical trials or their patients.

[20] Lipitor is also able to reduce triglyceride. Unfortunately, explicitly modeling learning about another dimension of product quality would go beyond the scope of this paper. Although we do not explicitly model

triglyceride, the brand dummy of Lipitor should include its effect as well.

[21] We assume that they all share the same carryover rate.

[22] Many statin users are elderly, and they may find it troublesome to switch to a different drug and worry about its potential side effects.

[23] Note that the summation is done with "$\neq j$" because switchers, who choose drug j in $t-1$, choose a drug other than j in t and, hence, do not contribute to $\hat{d}^3_{jt}$.

[24] It is unlikely that there is much news about statins available prior to Q1 1986 because the first statin was launched in Q2 1988.

[25] One might argue that landmark clinical trials are funded by drug companies, and hence, they may also be endogeneous. We note that landmark clinical trials take years to complete. Given that we have also included a drug fixed effect, it seems unlikely that the landmark clinical trial's release dates and outcomes are correlated with the demand shocks. Ching et al. (2016) make use of the same argument to justify treating detailing and publicity variables as exogenous from the estimation viewpoint.

[26] Berndt et al. (1996) find that detailing and journal advertising are highly correlated in the antiulcer drug market.

[27] At first, the initial prior for the efficiency ratio might seem too low as it suggests the initial belief about the link between high cholesterol and heart disease risks is weak. But we should note that because we use product-level data, the initial prior here should be interpreted as the mixed initial belief of doctors and patients. Moreover, it represents the belief prior to the statin market starting in Q3 1988 before any detailing and publicity about statins were available. Even though the public health agency had the professional knowledge about high LDL being a strong risk factor for heart disease (e.g., heart attack, stroke), it is plausible that regular consumers were not widely aware of it in 1988. This is consistent with our data, which shows that the number of prescriptions written in 1993 is only 9.34% of those in 2004.

[28] See https://www.statista.com/statistics/254341/pfizers-worldwide-lipitor-revenues-since-2003/, accessed April 14, 2018.

[29] Phase 4 refers to postmarketing studies to provide additional information, such as a treatment's risks, benefits, and optimal use.

[30] See https://aspe.hhs.gov/sites/default/files/pdf/77166/rpt_erg.pdf, accessed March 18, 2018.

References

Arrow K (1962) Economic welfare and the allocation of resources for invention. Committee on Economic Growth of the Social Science Research Council, eds. *The Rate and Direction of Inventive Activity: Economic and Social Factors* (Princeton University Press, Princeton, NJ), 609–626.

Azoulay P (2002) Do pharmaceutical sales respond to scientific evidence? *J. Econom. Management Strategy* 11(4):551–594.

Berndt ER, Bui LT, Lucking-Reiley DH, Urban GL (1996) The roles of marketing, product quality and price competition in the growth and composition of the U.S. antiulcer drug industry. Bresnahan TF, Gordon RJ, eds. *The Economics of New Goods*, vol. 58 (University of Chicago Press, Chicago), 277–328.

Chan T, Narasimhan C, Xie Y (2013) Treatment effectiveness and side effects: A model of physician learning. *Management Sci.* 59(6): 1309–1325.

Ching AT (2010a) A dynamic oligopoly structural model for the prescription drug market after patent expiration. *Internat. Econom. Rev.* 51(4):1175–1207.

Ching AT (2010b) Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration. *Internat. J. Indust. Organ.* 28(6):619–638.

Ching AT, Ishihara M (2010) The effects of detailing on prescribing decisions under quality uncertainty. *Quant. Marketing Econom.* 8(2):123–165.

Ching AT, Ishihara M (2012) Measuring the informative and persuasive roles of detailing on prescribing decisions. *Management Sci.* 58(7):1374–1387.

Ching AT, Erdem T, Keane M (2009) The price consideration model of brand choice. *J. Appl. Econom.* 24(3):393–420.

Ching AT, Erdem T, Keane MP (2013) Learning models: An assessment of progress, challenges, and new developments. *Marketing Sci.* 32(6):913–938.

Ching AT, Erdem T, Keane MP (2014) A simple method to estimate the roles of learning, inventories and category consideration in consumer choice. *J. Choice Model.* 13:60–72.

Ching AT, Erdem T, Keane MP (2017) Empirical models of learning dynamics: A survey of recent developments. Wierenga B, van der Lans R, eds. *Handbook of Marketing Decision Models*, 2nd ed., International Series in Operations Research & Management Science, vol. 254 (Springer, Cham, Switzerland), 223–257.

Ching AT, Erdem T, Keane MP (2019) How much do consumers know about the quality of products? Evidence from the diaper market. *Japanese Econom. Rev.* Forthcoming.

Ching AT, Clark RC, Horstmann IJ, Lim H (2016) The effects of publicity on demand: The case of anti-cholesterol drugs. *Marketing Sci.* 35(1):158–181.

Chintagunta PK, Renna Jiang R, Jin GZ (2009) Information, learning, and drug diffusion: The case of cox-2 inhibitors. *Quant. Marketing Econom.* 7(4):399–443.

Cholesterol Treatment Trialists' Collaborators (2005) Efficacy and safety of cholesterol-lowering treatment: Prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet* 366(9493):1267–1278.

Cosh C (2010) National pharmacare, as run by omniscient angels. *Maclean's* (September 18), http://www2.macleans.ca/2010/09/18/national-pharmacare-as-run-by-omniscient-angels/.

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

DeGroot M (1970) *Optimal Statistical Decisions* (McGraw-Hill, New York).

Delahoy PJ, Magliano DJ, Webb K, Grobler M, Liew D (2009) The relationship between reduction in low-density lipoprotein cholesterol by statins and reduction in risk of cardiovascular outcomes: An updated meta-analysis. *Clinical Therapeutics* 31(2):236–244.

Erdem T (1998) An empirical analysis of umbrella branding. *J. Marketing Res.* 35(3):339–351.

Erdem T, Keane MP (1996) Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Sci.* 15(1):1–20.

Erdem T, Sun B (2002) An empirical investigation of spillover effects of marketing mix strategy in umbrella branding. *J. Marketing Res.* 39(4):408–420.

Gallant AR, Hong H, Khwaja A (2018) The dynamic spillovers of entry: An application to the generic drug industry. *Management Sci.* 64(3):1189–1211.

Ghemawat P (1986) Sustainable advantage. *Harvard Bus. Rev.* 64(5): 53–58.

Ghemawat P, Spence M (1985) Learning curve spillovers and market performance. *Quart. J. Econom.* 100(Suppl):839–852.

Golder PN, Tellis GJ (1993) Pioneer advantage: Marketing logic or marketing legend? *J. Marketing Res.* 30(2):158–170.

Guasch JL, Weiss A (1980) Wages as sorting mechanisms in competitive markets with asymmetric information: A theory of testing. *Rev. Econom. Stud.* 47(4):653–664.

Heckman JJ (1981) The incidental parameters problem and the problem of initial condition in estimating a discrete time-discrete data stochastic process. Manski CF, McFadden DL, eds. *Structural*

Analysis of Discrete Data and Econometric Applications (MIT Press, Cambridge, MA), 179–195.

Janakiraman R, Sismeiro C, Dutta S (2009) Perception spillovers across competing brands: A disaggregate model of how and when. *J. Marketing Res.* 46(4):467–481.

Law MR, Wald NJ, Rudnicka AR (2003) Quantifying effect of statins on low density lipoprotein cholesterol, ischaemic heart disease, and stroke: Systematic review and meta-analysis. *BMJ* 326(7404):1423.

Leffler KB (1981) Persuasion or information? The economics of prescription drug advertising. *J. Law Econom.* 24(1):45–74.

Lieberman MB (1987) The learning curve, diffusion, and competitive strategy. *Strategic Management J.* 8(5):441–452.

Liu H, Liu Q, Chintagunta PK (2017) Promotion spillovers: Drug detailing in combination therapy. *Marketing Sci.* 36(3):382–401.

Liu Q, Steenburgh TJ, Gupta S (2015) The cross attributes flexible substitution logit: Uncovering category expansion and share impacts of marketing instruments. *Marketing Sci.* 34(1):144–159.

Marcoul P, Weninger Q (2008) Search and active learning with correlated information: Empirical evidence from mid-Atlantic clam fishermen. *J. Econom. Dynam. Control* 32(6):1921–1948.

Narayanan S, Manchanda P, Chintagunta PK (2005) Temporal differences in the role of marketing communication in new product categories. *J. Marketing Res.* 42(3):278–290.

Neslin SA, Rhoads EE, Wolfson P (2009) A model and empirical analysis of patient compliance and persistence in pharmaceuticals. Working paper, Dartmouth College, Hanover, NH.

Park A (2008) Is Vytorin a failure? *Time* (January 15), http://www.time.com/time/health/article/0,8599,1703827,00.html.

Reinganum J (1983) Uncertain innovation and the persistence of monopoly. *Amer. Econom. Rev.* 73(4):741–748.

Schumpeter JA (1961) *The Theory of Economic Development: An Inquiry Into Profits, Capital, Credit, Interest, and the Business Cycle* (Oxford University Press, Oxford, UK).

Shapiro BT (2018) Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *J. Political Econom.* 126(1):381–437.

Simon F, Kotler P (2003) *Building Global Biobrands: Taking Biotechnology to Market* (Free Press, New York).

Spence M (1984) Cost reduction, competition, and industry performance. *Econometrica* 52(1):101–121.

Tang M-J (1988) An economic perspective on escalating commitment. *Strategic Management J.* 9(S1):79–92.

Zhou Z, Rahme E, Pilote L (2006) Are statins created equal? Evidence from randomized trials of pravastatin, simvastatin, and atorvastatin for cardiovascular disease prevention. *Amer. Heart J.* 151(2):273–281.