MPRA
Munich Personal RePEc Archive

The Effects of Detailing on Prescribing Decisions under Two-Sided Learning

Ching, Andrew and Ishihara, Masakazu

Rotman School of Management, University of Toronto

10 September 2007

Online at https://mpra.ub.uni-muenchen.de/4935/
MPRA Paper No. 4935, posted 16 Sep 2007 UTC

The Effects of Detailing on Prescribing Decisions under Two-Sided Learning*

Andrew Ching†
Masakazu Ishihara
Rotman School of Management
University of Toronto

This draft: September 10, 2007

*An earlier version of this paper was circulated under the title "The Effects of Detailing on Prescribing Decisions under Quality Uncertainty." We thank Dan Ackerberg, Pradeep Chintagunta, Avi Goldfarb, Tom Holmes, Ig Horstmann, Ahmed Khwaja, Nitin Mehta, Sridhar Moorthy, Mengze Shi, Wei Tan, and various conferences and seminars participants for helpful comments. We particularly thank Matt Shum for helping us obtain the data from IMS Canada. The usual disclaimer applies.

†To whom all correspondence should be addressed. Address: Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON, CANADA M5S 3E6. Email: andrew.ching@rotman.utoronto.ca Phone: 416-946-0728.

The Effects of Detailing on Prescribing Decisions under Two-sided Learning

Abstract

A fundamental question in pharmaceutical marketing management is: How does the effectiveness of detailing change when additional information on drugs is revealed via patients' experiences during the product lifecycle? To address this question, we develop a model of detailing and prescribing decisions which incorporates uncertainty about the quality of drugs. Our model assumes that not only physicians/patients, but also drug manufacturers are uncertain about the qualities of drugs, and a representative opinion leader is responsible for updating the prior belief about these qualities. Physicians are heterogeneous in their information sets, and drug manufacturers use detailing as a means to increase/maintain the measure of well-informed physicians. We explicitly model physicians' forgetting by allowing the measure of well-informed physicians to depreciate over time. We estimate our model using product level data of ACE-inhibitor with diuretic in Canada. Our estimation approach allows us to control for the potential endogeneity of detailing. The results show that our model is able to fit the diffusion pattern very well, the effectiveness of detailing depends on the current information set and the measure of well-informed physicians, and the role of detailing-in-utility is minimal. Using our parameter estimates, we examine how a public awareness campaign, which encourages physicians/patients to report their drug experiences, would affect managerial incentives to detail.

Keywords: Detailing, Prescription Drugs, Decisions Under Uncertainty, Two-sided Learning, Representative Opinion Leader, Diffusion

JEL: C15, C35, D83, I11, I18, M31, M37, M38

1 Introduction

> Many serious Adverse Drug Reactions (ADRs) are discovered only after a drug has been on the market for years. Only half of newly discovered serious ADRs are detected and documented in the Physicians' Desk Reference within 7 years after drug approval.
>
> *Lasser et al. (2002), Journal of American Medical Association*

A major tool of marketing communication in the prescription drug market is detailing, in which drug manufacturers send sales representatives to visit physicians. This type of personal selling activities allows sales representatives to directly discuss compliance information, side-effects, and clinical studies of the drugs. One challenge in managing detailing activities throughout a drug's product lifecycle is that even manufacturers may be uncertain about the product attributes of their own drugs. Although some information on product attributes is established from clinical trials when a drug gains approval from the public health agency, many side-effects are not revealed until a large number of patients have tried the drug (Lasser et al. 2002).

In light of an environment filled with uncertainty about efficacies and side-effects of drugs, a fundamental question in pharmaceutical marketing management is: How does the effectiveness of detailing change when additional information on drugs is revealed via patients' experiences during the product lifecycle? The goal of this paper is to provide a framework that allows researchers to address this question. We develop a model of detailing and pharmaceutical demand, which can be estimated using standard *product level* panel data that contains sales volume, prices, and detailing efforts on drugs. To demonstrate the usefulness of our model, we apply it to the ACE-inhibitor with diuretic market in Canada.

There has been a growing literature in economics and marketing that studies the demand for pharmaceuticals using product level data.[1] Most of these studies (e.g., Leffler 1981, Hurwitz

[1]The majority of the studies in this industry use product level data because they are the least expensive data that could be purchased from IMS. Recently, there are a few studies which use proprietary individual level data to study the demand for prescription drugs (e.g., Gonul et al. 2001, Crawford and Shum 2005, Wosinska 2002,

and Caves 1988, Berndt et al. 1997, Rizzo 1999, Narayanan et al. 2004, Osinga et al. 2007) use a reduced-form approach to provide evidence that cumulative detailing can influence the demand for drugs. Another set of studies takes a structural modeling approach to study how uncertainty about drug qualities affects demand (e.g., Ching 2000; 2004; 2005, Narayanan et al. 2005, Mukherji 2002). In particular, Narayanan et al. (2005) and Mukherji (2002) use the framework of Erdem and Keane (1996) to investigate the effects of detailing on demand, in which they assume manufacturers use detailing to convey noisy signals about the true quality of their products to physicians. These studies provide a useful framework for quantifying the impact of aggregate learning on demand and how detailing affects the rate of learning when manufacturers have complete information about the quality of their drugs. However, they have several limitations, which we seek to address:

First, they do not take into account situations where drug manufacturers may need to learn the side-effects and efficacy profiles of their own drugs over the product lifecycle. For drugs with side-effects and efficacy profiles that change over time, their assumption that detailing always provides noisy signals about the true quality would not be appropriate. Such a misspecification could lead to biased estimates of the learning parameters. In particular, the precision of the consumption experience signals could be underestimated.

A second limitation is that they either ignore forgetting or use a reduced-form approach to model physicians' forgetting via the depreciation of the detailing stock in the utility function. By not modeling forgetting explicitly, previous studies may have underestimated the effect of informative detailing and overestimated the effect of the detailing stock in the utility function.

A third limitation is that they either restrict physicians to possess the same information set over time, or assume that every physician uses the same number of information signals to update his/her prior beliefs about drug qualities in each period. As a result, their models have restrictive substitution patterns such as IIA with respect to the effect of detailing.

Narayanan and Manchanda 2006, Dong et al. 2006). In particular, Crawford and Shum (2005) and Narayanan and Manchanda (2006) model how an individual physician/patient learns his/her own match with different drugs. Unfortunately, individual level data in this market is very hard to obtain.

Finally, they do not deal with the endogeneity problem of detailing. Conceivably, when a manufacturer updates his belief about the quality of his own drug favorably, he may react to it by increasing his detailing efforts so as to bring this information to physicians.[2] Ignoring this endogeneity problem would potentially result in biased estimates of the parameters associated with detailing. Nonetheless, the structural modeling literature in pharmaceutical demand that uses product level data has so far neglected to take this endogeneity problem into account.[3]

To address these limitations, our model differs from the previous ones in the following ways: (i) We assume that not only physicians/patients, but also manufacturers are uncertain about drugs' qualities, and they rely on a representative opinion leader to collect past consumption experiences from patients and update the current public information sets for drugs. (ii) We allow physicians to be heterogeneous in their information sets.[4] For each drug, physicians are either informed of its current public information set or uninformed. We allow the measure of physicians who are informed about a particular drug to depend on its cumulative detailing efforts. So, unlike the previous models, detailing does not provide noisy signals about the true qualities in our model. (iii) We explicitly model physicians' forgetting by allowing the measure of well-informed physicians to depreciate over time.

Our focus is to model the effects of detailing on demand instead of equilibrium strategies by manufacturers. Therefore, to take the potential endogeneity problem of detailing into account, we extend the estimation method proposed by Ching (2000; 2005), which does not require solving manufacturers' optimization problem. We use a reduced form approach to model detailing as a

[2]Azoulay (2002) finds evidence that drug companies change their detailing efforts when new information about their drugs becomes available in the U.S. anti-ulcer drugs market.

[3]As far as we know, there is only one recent structural modeling paper by Dong et al. (2006), which endogenizes detailing at the individual level. The endogeneity problem that they focus on is different from ours. In their case, the endogeneity problem is due to the unobserved physician level heterogeneity. In our case, it is due to the unobserved product characteristics because we use product level data. Another difference is that Dong et al. (2006) do not model consumer/physician learning.

[4]It should be noted that Narayanan and Manchanda (2006) model heterogenous physicians' learning. However, they use individual level data instead of product level data. The sources of identification in our model are also different from theirs.

function of observed and unobserved variables that determine demand, and then jointly estimate this pseudo-detailing policy function with the demand side model.

In addition to the economics and marketing literature on pharmaceutical industry, our paper is related to the literature on two-sided learning. Some theoretical papers in this literature study the equilibrium diffusion pattern (e.g., Bergemann and Valimaki 1997) and pricing strategies (e.g., Villas-Boas 2006) in an environment where both buyers and sellers are uncertain about the qualities of the products and consumption signals are observed by all parties (so that the information set is shared by all parties). Bergemann and Valimaki (2006) allow for heterogeneous learning on the buyer's side. Ching (2000; 2004; 2005) estimates a structural model with two-sided learning to examine the equilibrium pricing strategies and diffusion pattern empirically in the U.S. prescription drug market. However, he does not model detailing. To our knowledge, none of the papers in this literature explicitly model buyers' forgetting.

Our paper is also related to the one-sided learning literature. In addition to Erdem and Keane (1996), four other papers are particularly relevant. Mullainathan (2002) studies learning and forgetting in a theoretical model. Mehta et al. (2004) develop and estimate a structural model of learning with forgetting using individual level scanner data instead of product level data. Both Mullainathan (2002) and Mehta et al. (2004) do not model the effect of advertising. Ackerberg (2003) estimates a model in which a consumer infers the value of the product to him/her from the advertising intensity (implicitly through the signaling equilibrium). He does not allow for consumer forgetting. Moreover, similar to Erdem and Keane (1996), he assumes manufacturers know the true mean quality of their products. Hitsch (2006) estimates a structural model in which manufacturers are uncertain about how consumers evaluate their products. However, unlike our model, consumers have complete information in his model.

Although we do not model the manufacturers' problem explicitly, as far as we know, this is the first paper that develops an empirical structural model to study the effects of detailing on demand, allowing for two-sided learning, physician heterogeneity in their information sets, and physicians' forgetting. Our main findings can be summarized as follows: First, our model is able to generate a flexible diffusion pattern – it fits the diffusion pattern of demand very well; Second, we quantify the marginal impact of detailing on current demand at different

points in time and show how it depends on the measure of well-informed physicians and the information sets; Third, we find evidence that the endogeneity problem biases the estimates of the coefficients associated with detailing; Fourth, we find evidence that the role of detailing-in-utility could be much smaller in our framework; Lastly, using our parameter estimates, we conduct a policy experiment to evaluate how a public awareness campaign, which encourages physicians/patients to report their drug experiences, would affect managerial incentives to detail. Given our parameter estimates, we find that the marginal return of detailing has increased under this campaign, suggesting that managers should increase their detailing efforts.

The rest of the paper is organized as follows. Section 2 provides some background of the prescription drug market. Section 3 describes the demand model. Section 4 describes data and the estimation strategy. Section 5 discusses the results. Section 6 is the conclusion.

2 Background

Why would drug manufacturers be uncertain about the quality of their products during the product lifecycle given that they developed these drugs? To understand this, it is important for us to give some background information about the approval process of new drugs. Most countries, including the U.S. and Canada, have a similar approval process. Drug manufacturers are required to prove that a new drug is safe and effective before marketing it. The proof involves a series of clinical trials, which are divided into three phases. Phase I and II studies provide basic evidence that the drug works in a small sample of patients. Phase III studies require a relatively larger sample of patients, which ranges from hundreds to several thousands. These studies are designed to evaluate the safety and effectiveness of the drug, wherein manufacturers need to demonstrate that the drug works better than a placebo. Nevertheless, manufacturers are not required to show that the new drug performs better than existing drugs that treat the same problem. Moreover, although most public health agencies set high standards for phase III clinical studies, it is not uncommon that they do not reveal all the side-effects, as documented by Lasser et al. (2002). These suggest that even manufacturers may not have complete information about the quality of their own drugs at the beginning of the product lifecycle.

Many public health agencies such as the U.S. FDA and Health Canada recognize this fact. This is why they establish computerized information databases for storing and analyzing safety reports submitted by physicians, patients, and drug manufacturers.[5] Public health agencies use reports from these databases to keep track of the side-effects profile over time. If there is sufficient evidence that a previously unknown serious side-effect is associated with a drug, they will require the manufacturer to add more warning statements on the drug label. In rare occasions, they may require the manufacturer stop the sale of the drug if the risks clearly outweigh the benefits. Other channels, such as educational meetings and conferences, also provide opportunities for health care professionals to share their patients' experiences.

Physicians are supposed to keep themselves informed of the most updated information for drugs. However, with many new drugs entering the market each year, it is difficult for physicians to keep up with the enormous amount of information that changes regularly.[6] Most primary care physicians are occupied with seeing patients. It is costly for them to contact public health agencies to learn the updated side-effects, read academic journals on recent clinical trials, or contact opinion leaders to obtain the latest information on the benefits and risks of drugs. They therefore rely on sales representatives as a source of information (Coleman et al. 2004, p.179, Greider 2003, p.67). Without detailing, it is plausible that a primary care physician may forget the information about a drug's side-effects and effectiveness over time, and as a result, become reluctant to prescribe the drug. There is indirect evidence that supports this hypothesis: Caves et al. (1991) find that most drug manufacturers during the 80s dramatically reduces their detailing efforts for drugs whose patents are about to expire, and the total demand for those drugs typically declines over time after patent expiration.

It is possible that the presentations given by sales representatives are biased towards the drugs they promote. This possibility appears to be well-recognized by health care professionals (e.g., Cooper et al. 2003, Ziegler et al. 1995), and physicians are usually cautious when listening

[5]For example, in the U.S., drug manufacturers are required to report each adverse drug experience within 15 days of the initial receipt of the information. Physicians are encouraged to file reports to the FDA on a voluntary basis.

[6]For example, the number of active drugs in the cardiovascular drug category increased from 215 in March 1993 to 294 in February 1999 in Canada.

to the sales representatives' claims. It is common that during their visits, sales representatives hand out printed documents related to efficacies and side-effects of the drugs being promoted (e.g., published academic articles about clinical trials). Although the printed documents may not be complete, more likely than not it saves physicians' time in gathering the related literature. Moreover, the printed documents will likely contain accurate information. If the documents are wrong or deliberately misleading, public health agencies could use them as concrete evidence to file criminal and civil charges against the manufacturers.[7] Although physicians may be skeptical about the information delivered by the sales representatives, the favorable picture of the drug presented by them may trigger physicians' interests to learn the latest information of the drug being promoted. They may then be more likely to read the related medical literature, or contact peers who are opinion leaders in the related field for more information.[8] One implication of this hypothesis is that the impact of detailing on demand would depend on the actual effectiveness and side-effects of the drug. Venkataraman and Stremersch (2006) test this hypothesis and find that the effect of detailing is indeed higher for drugs that are more effective and have less side-effects in three therapeutic classes: anti-cholesterol drugs (statins), gastrointestinal drugs and erectile dysfunctions drugs. In this paper, our way of modeling detailing will be consistent with this hypothesis.

3 Model

We now turn to discuss our model of detailing and prescribing decisions. Our framework here extends Ching (2000; 2004; 2005), who presents the first empirical structural model that allows both the demand side and the supply side to be uncertain about the qualities of drugs. He studies the competition between brand-name drugs and their generic counterparts. However, as mentioned before, he does not model detailing, which is the main focus of our paper.

[7]The penalty of carrying out misleading promotion is usually very high. For example, Purdue Pharma was fined $600 million for misleading promotion of OxyContin in May 2007.

[8]In addition to detailing, the medical continuing education literature finds that medical journals and opinion leaders are also the main sources of information for primary care physicians (e.g., Haug 1997, Thompson 1997).

In our model, there are three types of agents: physicians, manufacturers, and a representative opinion leader. There are two types of products: inside goods which represent the products that use similar chemical compounds (so-called "me-too" drugs), and an outside good that represents their substitutes (0). Product characteristics can be distinguished as p_j and q_j, $j = 1, ..., J$, where p_j is the price of product j, and q_j is the mean quality level of product j. All agents in the model are perfectly informed about p_j, but are imperfectly informed about the drug's mean quality level, q_j.

To capture the idea that there are opinion leaders who gather the most recent information about drug qualities, we introduce a *representative opinion leader* in our model. The representative opinion leader maintains a vector of public information sets, $I(t) = (I_1(t), ..., I_J(t))$, which describes the most updated belief about $q = (q_1, \ldots, q_J)$ at time t based on past patients' experiences available to the public. For each drug j, a physician either knows $I_j(t)$, or $\underline{I}_j^p$, which is the initial prior that physicians have when drug j is first introduced. Let M_{jt} be the measure of physicians who know $I_j(t)$. We assume that M_{jt} depends on the cumulative detailing efforts at time t. There are two stages in each period. In the first stage, manufacturers choose detailings. Given the amount of detailings, M_{jt} is determined for each j. Each physician makes his/her prescribing decision based on his/her information about the drugs. In the second stage, patients consume the prescribed drugs and some of their experience signals are revealed to the public. The representative opinion leader then uses these signals to update $I(t+1)$ in a Bayesian fashion. We will describe these two stages backward.

3.1 Updating of the Information Set

A drug is an experienced good. Consumption of a drug provides information about its quality. It is assumed that physicians and patients in the model can measure drug qualities according to a fixed scale. For example, a patient can measure quality in terms of how long he/she needs to

wait before the drug becomes effective to relieve his/her symptoms, how long his/her symptoms would be suppressed after taking the drug, or how long the side-effects would last.[9]

Each patient i's experience with the quality of drug j at time t ($\tilde{q}_{ijt}$) may differ from its mean quality level q_j. As argued in Ching (2000), the difference between $\tilde{q}_{ijt}$ and q_j could be due to the idiosyncratic differences of human bodies in reacting to drugs. An experience signal may be expressed as,

$$\tilde{q}_{ijt} = q_j + \delta_{ijt}, \tag{1}$$

where δ_{ijt} is the signal noise. We assume that δ_{ijt} is an $i.i.d.$ normally distributed random variable with zero mean:

$$\delta_{ijt} \quad \sim \quad N(0, \sigma_\delta^2), \tag{2}$$

and the representative opinion leader's initial prior on q_j ($\underline{I}_j^o$) is also normally distributed:

$$q_j \quad \sim \quad N(\underline{q}_j^o, \underline{\sigma}_j^{o2}). \tag{3}$$

The representative opinion leader updates the public information set at the end of each period using the experience signals that are revealed to the public. The updating is done in a Bayesian fashion. In each period, we assume that the number of experience signals revealed is a random subsample of the entire set of experience signals. This captures the idea that not every patient revisits and discusses his/her experiences with physicians, and not every physician shares his/her patients' experiences with others.

According to the Bayesian rule (DeGroot 1970), the expected quality is updated as follows:

$$E[q_j|I(t+1)] = E[q_j|I(t)] + \iota_j(t)(\bar{q}_{jt} - E[q_j|I(t)]), \tag{4}$$

where $\bar{q}_{jt}$ is the sample mean of all the experience signals that are revealed in period t.[10] $\iota_j(t)$ is a Kalman gain coefficient, which is a function of the variance of the signal noise (σ_δ^2), perceived

[9] Obviously, drug qualities are multi-dimensional. Implicitly, we assume patients are able to use a scoring rule to map all measurable qualities to a one-dimensional index. It is the value of this one-dimensional index that enters the utility function.

[10] Let q_j be the true mean quality level of drug j. Then, $\bar{q}_{jt}|(\kappa n_{jt}, I(t)) \sim N(q_j, \frac{\sigma_\delta^2}{\kappa n_{jt}})$.

variance ($\sigma_j^2(t)$), the quantity sold at time t (n_{jt}), and the proportion of experience signals revealed to the public (κ), and it can be expressed as:

$$\iota_j(t) = \frac{\sigma_j^2(t)}{\sigma_j^2(t) + \frac{\sigma_\delta^2}{\kappa n_{jt}}}. \tag{5}$$

ι_j can be interpreted as the weights that the representative opinion leader attaches to the information source in updating its expectation about the level of q_j. In particular, $\iota_j(t)$ increases with $\sigma_j^2(t)$.

The perception variance at the beginning of time $t+1$ is given by (DeGroot 1970):

$$\sigma_j^2(t+1) = \frac{1}{\frac{1}{\sigma_j^2(0)} + \frac{\kappa N_{jt}}{\sigma_\delta^2}}, \tag{6}$$

where $N_{jt}(= \sum_{\tau=1}^{t} n_{j\tau})$ is the cumulative consumption of drug j, or,

$$\sigma_j^2(t+1) = \frac{1}{\frac{1}{\sigma_j^2(t)} + \frac{\kappa n_{jt}}{\sigma_\delta^2}}. \tag{7}$$

Equation (6) implies that, after observing a sufficiently large number of experience signals for a product, the representative opinion leader will learn about q_j, at any arbitrarily precise way (i.e., $\sigma_j(t) \to 0$ and $E[q_j|I(t)] \to q_j$ as the number of signals received grows large). We will next turn to discuss the physicians' choice problem and how detailing influences their choices.

3.2 Detailing and Measure of Well-Informed Physicians

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.e., $I_j(t)$. An uninformed physician only knows the initial prior, i.e., $\underline{I}_j^p = N(\underline{q}_j^p, \underline{\sigma}_j^{p2})$. This implies that the number of physician types is 2^J. Note that physicians' initial prior $\underline{I}_j^p$ could differ from the initial prior of the representative opinion leader, $\underline{I}_j^o$.

We assume that manufacturers observe $I(t)$ when they decide the amount of detailing, $D_{1t}, ..., D_{Jt}$. In general, the measure of well-informed physicians for drug j at time t, M_{jt}, is a function of M_{jt-1} and $D_{1t}, ..., D_{Jt}$. For simplicity, we assume that this function only depends on

M_{jt-1} and D_{jt}, i.e., $M_{jt} = f(M_{jt-1}, D_{jt})$. We assume that $f(M_{jt-1}, .)$ is monotonically increasing in D_{jt}. To capture the idea that physicians may forget, we assume that $f(M, 0) \leq M, \forall M$.

Two remarks should be made regarding the way we model the relationship between detailing and the measure of well-informed physicians. First, similar to Mullainathan (2002), we do not allow uninformed physicians for drug j at time t to possess any $I_j(t')$ for $t' < t$, but $\underline{I}_j^p$. As we mentioned above, even with our current setup, the number of types increases exponentially in J. Although allowing physicians who "partially" forget may seem more appealing, it will dramatically increase the size of the state space – we would need to keep track of the measure of physicians who know $I_j(t')$, for all j and $t' < t$. The number of types will increase to t^J in time t. Such a modification will make the model computationally infeasible to estimate.[11] On the other hand, our assumption is not as restrictive as it may seem. One interpretation is that we approximate the aggregate demand from t^J types of physicians by randomizing the demand of 2^J types.

Second, we assume that M_{jt} depends on D_{jt} partly because the main job of sales representatives is to give physicians documented information about side-effects and efficacies of the drug that they are promoting. We do not mean that physicians simply believe what sales representatives claim during their conversations. Rather, we try to capture the intuition that detailing would increase the chances that physicians obtain the most recent information about the drug (by consulting their peers, reading the medical literature, etc.). This could be because the visits stimulate their interests, increase their awareness of existing or new clinical studies, and make it easier for them to access the relevant journal articles.

Define the average rate of forgetting, $\phi_M \equiv (M - f(M, 0))/M$. We assume further that there exists an $\underline{M}$ and an $\bar{M}$, with $0 < \underline{M} < \bar{M}$, such that $\phi_M(M) = 0$ for $M \leq \underline{M}$, $\phi'_M(\underline{M}) > 0$, $\phi'_M(\bar{M}) = 0$, and $\phi''_M(M) < 0, \forall M$. The assumptions allow us to capture the following intuition. It is likely that physicians are heterogeneous in terms of their rate of forgetting. Some physicians who are more willing to spend time to keep up with the most recent medical literature themselves are likely to have a lower rate of forgetting. Other physicians who prefer to spend most of their

[11] However, with individual level data, it is feasible to estimate a model of learning with partial forgetting (Mehta et al. 2004).

time seeing patients, are likely to have a higher rate of forgetting – they probably will rely more on sales representatives to help them get the most updated information. When M is small, we expect that most of the well-informed physicians would be those who have a lower rate of forgetting. As M increases, we expect that the proportion of well-informed physicians who have a higher forgetting rate would increase. On the other hand, we expect that the number of interactions among well-informed physicians would also increase with M. They might remind each other about how this drug works, which helps reduce the average rate of forgetting (i.e., the network effect). These two forces work against each other. In particular, it is likely that the latter dominates the former when M is large, and vice versa. We therefore expect that when M is small, ϕ_M will first increase with M at a diminishing rate. After M has passed a certain threshold, ϕ_M will eventually decrease with M.

One might argue that the network effect could allow M to expand even without any detailing efforts. We rule this possibility out in our model because it is inconsistent with the stylized fact that we mentioned before: The demand for a brand-name drug typically declines over time after a manufacturer has dramatically reduced its detailing efforts (Caves et al. 1991). This suggests that without detailing, the measure of well-informed physicians would depreciate.

In our econometric model, we capture the relationship between M_t and (M_{t-1}, D_t) by introducing a detailing goodwill stock, G^I_{jt}, which accumulates as follows:

$$G^I_{jt} = (1 - \phi_I)G^I_{jt-1} + D_{jt}, \tag{8}$$

where D_{jt} is manufacturer j's detailing efforts in time t, and $\phi_I \in [0, 1]$ is the corresponding depreciation rate. We specify the relationship between M_{jt} and G^I_{jt} as:

$$M_j = \frac{exp(\beta_0 + \beta_1 G^I_j)}{1 + exp(\beta_0 + \beta_1 G^I_j)}. \tag{9}$$

Although ϕ_I is a constant, G^I_j affects M_j nonlinearly. In particular, the implied average forgetting rate, ϕ_M, will exhibit an inverted-U shape described above.

3.3 Prescribing Decisions

Now we turn to discuss how physicians make their prescribing decisions. Each physician takes the current expected utility of his/her patients into account when making prescribing decisions. Physician h's objective is to choose $d_{hij}(t)$ to maximize the current period expected utility for his/her patients:

$$E[\sum_{j\in\{0,1,\ldots,J\}} u_{ijt} \cdot d_{hij}(t)|I^h(t)], \tag{10}$$

where $d_{hij}(t) = 1$ indicates that alternative j is chosen by physician h for patient i at time t, and $d_{hij}(t) = 0$ indicates otherwise. We assume that $\sum_j d_{hij}(t) = 1$. The demand system is obtained by aggregating this discrete choice model of an individual physician's behavior.

We assume that a patient's utility of consuming a drug can be adequately approximated by a quasilinear utility specification, additively separable in a concave subutility function of drug return, and a linear term in price. The utility of patient i who consumes drug j at time t is given by the following expression:

$$u_{ijt} = \alpha - \exp(-r\tilde{q}_{ijt}) - \pi_p p_{jt} + \zeta_{ikt} + e_{ijt}, \tag{11}$$

where p_{jt} is the price for product j at time t; r is the risk aversion parameter; α is the common intercept across drugs; π_p is the utility weight for price; $(\zeta_{ikt} + e_{ijt})$ represents the distribution of patient heterogeneity; k indexes nest (i.e., inside good or outside good).[12] ζ_{ikt} and e_{ijt} are unobserved to the econometrician but observed to the physicians when they make their prescribing decisions. We assume that ζ_{ikt} and e_{ijt} are *i.i.d.* extreme value distributed. The exponential specification of the subutility function of drug return is known as the Constant Absolute Risk Aversion (CARA) utility. In this specification, r represents the coefficient of absolute risk aversion.

Note that $\tilde{q}_{ijt}$ is observed neither by physicians nor patients when prescribing decisions are made. It is observed by physicians/patients only after patients have consumed the drug,

[12]This is equivalent to modeling physicians' choice as a two-stage nested process, where they choose between the inside goods and the outside good in the first stage, and then choose an alternative among the inside goods in the second stage.

but it remains unobserved by the econometrician. Physicians make their decisions based on the expected utility of their patients. Let $I(t)$ and $I^h(t)$ denote the representative opinion leader's information set and physician h's information set at time t, respectively. If physician h is well-informed about drug j at time t, his/her expected utility will be:

$$\begin{aligned} E[u_{ijt}|I^h(t)] &= E[u_{ijt}|I_j(t)] \\ &= \alpha - exp(-rE[q_j|I(t)] + \frac{1}{2}r^2(\sigma_j^2(t) + \sigma_\delta^2)) - \pi_p p_{jt} \\ &\quad +\zeta_{ikt} + e_{ijt}. \end{aligned} \tag{12}$$

If physician h is uninformed about drug j at time t, his/her expected utility of choosing drug j becomes:

$$\begin{aligned} E[u_{ijt}|I^h(t)] &= E[u_{ijt}|\underline{I}_j^p] \\ &= \alpha - exp(-r\underline{q}_j^p + \frac{1}{2}r^2(\underline{\sigma}_j^{p2} + \sigma_\delta^2)) - \pi_p p_{jt} + \zeta_{ikt} + e_{ijt}. \end{aligned} \tag{13}$$

It should be noted that patient heterogeneity components of the utility function (ζ_{ikt}, e_{ijt}) reappear in the expected utility equation because they are stochastic only from the econometrician's point of view.

Equations (11)-(13) apply only to the inside alternatives. In each period, physicians may also choose an outside alternative that is not included in our analysis (i.e., other non-bioequivalent drugs). We assume the expected utility associated with the outside alternative takes the following functional form:

$$E[u_{i0t}|I^h(t)] = \alpha_0 + \pi_t t + \zeta_{i0t} + e_{i0t}. \tag{14}$$

The time trend of the outside alternative allows the model to explain why the total demand for inside goods may increase or decrease over time.

The quantity demand, n_{jt}, can be expressed as,

$$n_{jt} = Size_t \cdot S(j|D_t, (E[q_j|I(t)], \sigma_j(t), M_{jt-1})_{j=1}^2; \theta_d) + \epsilon_{jt}, \tag{15}$$

where $Size_t$ is the size of the market, $S(j|\cdot)$ is the market share of drug j, ϵ_{jt} represents a measurement error, and θ_d is a set of demand side parameters.

3.4 Empirical Implications

We now discuss some empirical implications of our model. We consider the case of two products. In this case, there are four types of physicians (2^2) who differ in their information sets. Let $s_{jt}(I_j, I_k)$ be the probability of choosing drug j at time t by physicians who have the information sets I_j and I_k for drugs j and k, respectively ($j \neq k$). Then the market share for drug j at time t is given by,

$$\begin{aligned} S_{jt} &= M_{jt}M_{kt}s_{jt}(I_j(t), I_k(t)) + M_{jt}(1 - M_{kt})s_{jt}(I_j(t), \underline{I}_k^p) \\ &\quad +(1 - M_{jt})M_{kt}s_{jt}(\underline{I}_j^p, I_k(t)) + (1 - M_{jt})(1 - M_{kt})s_{jt}(\underline{I}_j^p, \underline{I}_k^p), \end{aligned} \tag{16}$$

where $s_{jt}(I_j, I_k)$ has a closed form expression given that we use the nested logit framework. It follows that the marginal return of detailing on current market share for drug j is,

$$\frac{\partial S_{jt}}{\partial D_{jt}} = \frac{\partial M_{jt}}{\partial D_{jt}} \times \{M_{kt}\Delta s_{jt}(I_k(t)) + (1 - M_{kt})\Delta s_{jt}(\underline{I}_k^p)\}, \tag{17}$$

where $\Delta s_{jt}(I_k) \equiv s_{jt}(I_j(t), I_k) - s_{jt}(\underline{I}_j^p, I_k)$. Intuitively, $\Delta s_{jt}(I_k)$ is the change in the probability of choosing j when a physician switches his/her information set for drug j from $\underline{I}_j^p$ to $I_j(t)$, conditional on him/her knowing I_k ($= I_k(t)$ or $\underline{I}_k^p$). Equation (17) shows that the marginal return of detailing depends on $\Delta s_{jt}(I_k(t))$ and $\Delta s_{jt}(\underline{I}_k^p)$, which are weighted by M_{kt} and $1 - M_{kt}$, respectively. This weighted average is further adjusted by $\partial M_{jt}/\partial D_{jt}$. It is worth noting that $\partial S_{jt}/\partial D_{jt}$ increases (decreases) with M_{kt} if $(\Delta s_{jt}(I_k(t)) - \Delta s_{jt}(\underline{I}_k^p))$ is positive (negative).

Consider a situation where a new drug enters a market with a matured incumbent (in the sense that the representative opinion leader has learnt the true quality of the incumbent, i.e., $I_k(t) \to I_k(\infty)$). Conditional on M, equations (16) and (17) imply that the entrant's marginal return of detailing will increase with its market share. Moreover, the detailing elasticity of demand in our model could **increase** or **decrease** over time partly depending on how $I(t)$ evolves. In particular, even after the uncertainty about the drug quality is completely resolved, detailing still affects demand as long as $\phi_I > 0$, and its effect depends on $I(t), \underline{I}^p$ and M_{jt-1} (i.e., G_{jt-1}^I). On the contrary, previous models of learning and detailing/advertising, which follow the framework of Erdem and Keane (1996), imply that the detailing/advertising elasticity of demand **diminishes** over time as uncertainty about product quality is slowly resolved. This

demonstrates that the empirical implications from our model are quite different from those from the previous models.

3.5 Identification

As shown in equation (16), S_{jt} is a non-linear function of the structural parameters of the model, history of sales, and cumulative detailing. This indicates that the parameters of the model can be identified from the diffusion paths of inside goods. The new element of our learning model is the measure of well-informed physicians. It is worth discussing how this feature of the model helps us fit the data. Note that a traditional Bayesian learning model, which assumes the entire population is well-informed, typically implies an S-shaped diffusion path (Bergemann and Valimaki 1997). However, the diffusion paths observed in the data may not have such a pattern. For example, in the market of ACE-inhibitor with diuretic which will be analyzed by our model in the next section, the diffusion paths appear to be linear. When fitting this diffusion path, a traditional learning model will typically underpredict and overpredict the demand in the earlier part and the later part of the product lifecycle, respectively. Adding M_{jt} to a learning model allows it to generate a more flexible diffusion path. In order to fit a linear diffusion path, we will need to adjust the parameters of $f(M_{jt-1}, D_{jt}; \beta)$ so that it takes time to build M_{jt} and removes the steep increasing feature of the predicted diffusion path implied by a traditional learning model.

4 Estimation

4.1 Overview of the Data

Having described our model, we now turn to an application. We estimate our model using Canadian data for ACE-inhibitor with diuretic, which treats hypertension. ACE-inhibitor (Angiotensin Converting Enzyme Inhibitor) works by limiting the production of a substance that promotes salt and water retention in the body. Diuretic induces the production and elimination

of urine, which helps in lowering blood pressure. This class of combination drugs is usually not prescribed until therapy is under way.

We choose Canada and ACE-inhibitor with diuretic for three reasons. First, most of the patients who have high blood pressure are elderly, and their prescription drugs are covered by the Canadian government. Moreover, Canada has price regulations on brand-name drugs. The Patented Medicine Price Review Board restricts Canadian prices of patented drugs to be below the median prices of G7 countries. There is evidence which suggests that this constraint is binding on average (Elgie 2001). These institutional details, which suggest that price does not play an important role in determining demand, allow us to treat prices as exogenous and focus on modeling the effects of detailing. Second, the market of ACE-inhibitor with diuretic does not have direct-to-consumer (DTC) advertising. DTC advertising has increased dramatically in the U.S. since 1997. It is believed that it plays an important role in the demand for prescription drugs. However, the way that DTC advertising influences physicians' choice is likely to be different from detailing. Modeling the effects of DTC advertising is beyond the scope of this paper. Third, the market of ACE-inhibitor with diuretic only has two dominant drugs. We feel that it is sensible to first apply our framework to this simple market before tackling markets with more competitors.

Data sources for this study come from IMS Canada, a firm that specializes in collecting sales and advertising data for the Canadian pharmaceutical industry. The revenue data is drawn from their Canadian Drugstore and Hospital Audit (D&H); the number of prescriptions is drawn from their Canadian Compuscript Audit (CCA); the number of detailing minutes is drawn from their Canadian Promotion Audit (CPA). Although D&H does not include purchases made by the government, mail order pharmacies, and nursing homes or clinics, IMS believes that it covers about 90% of total sales. The price is obtained by dividing the revenue by the number of prescriptions. We deflated the prices using the consumer price index in the Canadian pharmaceutical industry. We note that on average less than one percent of sales is from hospital purchases. Due to its dominance, we only model the segment of the drugstore market and ignore how hospitals reach their purchase decisions.

The data set contains monthly data from March 1993 to February 1999. There are two main brand-name drugs in the market – Vaseretic and Zestoretic. Vaseretic is marketed by Merck; its generic ingredients are enalapril and hydrochlorothiazide. It was approved by Health Canada in September 1990. Zestoretic is marketed by AstraZeneca; its generic ingredients are lisinopril and hydrochlorothiazide. It was approved in October 1992. Both of them are present throughout the sample period, and they capture more than 80% of sales of the ACE-inhibitor with diuretic category. We therefore focus our analysis on these two drugs. Treating product/month as one observation, the total sample size is 144. We report the summary statistics in Table 1.

For an overview of the data, we plot the number of prescriptions filled for Vaseretic and Zestoretic in Figure 1. The sales of both drugs increase over time. The monthly sales of Vaseretic grow slowly and steadily from 2,500 prescriptions to 4,500 prescriptions, while Zestoretic's monthly sales grow at a much faster rate from around 300 prescriptions to more than 14,000 prescriptions. Being the incumbent of the ACE-inhibitor with diuretic, the sales of Vaseretic is about eight times that of Zestoretic at the beginning of the sample period (March 1993). It took Zestoretic more than two years to overtake Vaseretic's sales. By the end of the sample period (February 1999), the sales of Zestoretic is more than three times that of Vaseretic. The sales trend of Zestoretic is remarkable, and illustrates the slow diffusion of new drugs well documented in this industry.

The potential size of the market is defined as the total number of prescriptions for drugs that belong to ACE-inhibitor, Thiazide Diuretic, and ACE-inhibitor with diuretic. In Figure 2, we plot the size of the market over time. It increases from 655,000 to 860,000 during the sample period.

4.2 Simultaneity Problem

As we argued above, we assume price to be exogenous, and focus on the potential endogeneity problem of detailing when estimating our physicians choice model. It is plausible that manufacturers observe $I(t)$ before detailing takes place in each period. Therefore, it is likely that detailing is a function of $I(t)$. In particular, we expect that D_{jt} may be correlated with $E[q_j|I(t)]$

and $\sigma_j(t)$. For instance, if $E[q_j|I(t)]$ is higher than $E[q_k|I(t)]$, manufacturer j may have an incentive to increase D_{jt} so as to disseminate the information. If we ignore this correlation, the parameters for building up the measure of well-informed physicians will likely be biased upward.

A popular method to estimate this class of model using product level data is developed by Berry et al. (1995) (BLP). They show that there is a one-to-one mapping between the mean utility levels and the observed market shares, conditional on a parameter vector. As a result, it is possible to construct a GMM objective function based on the mean utility function without explicitly solving the supply side model. However, as pointed out by Chernozhukov and Hong (2003), BLP's GMM objective function is highly nonconvex with many local optima. This poses a formidable challenge when minimizing it in practice. Another way to handle this endogeneity problem is to explicitly model manufacturers' decision on detailing, and incorporate their detailing policy functions in a full-information maximum likelihood procedure. Since detailing has a long-lived effect, this would involve developing a forward-looking dynamic oligopoly structural model. Unfortunately, estimating this type of dynamic oligopoly model using a full-solution method has proved to be infeasible given today's computational power.

In this paper, we estimate our model using the approach developed by Ching (2000; 2005). Similar to BLP, this method does not require solving the dynamic oligopolistic supply side model. To take the endogeneity of detailing into account, he proposes to approximate manufacturers' policy functions by expressing it as a polynomial of the state variables (both observed and unobserved), and then jointly estimate this pseudo-policy function and the demand model.[13]

There are two drawbacks in this approach: (i) It increases the number of parameters to estimate due to the pseudo-detailing policy functions; (ii) The estimates are not efficient because the supply side model is not used in the estimation. However, this approach does not require us to make any strong assumptions about the equilibrium solution, and whether drug manufacturers maximize their total discounted profits or current profits. So we avoid some risks of misspecifying the supply side, which may result in biased estimates. More importantly, it allows us to avoid the computational burden of solving a dynamic oligopoly model when estimating the demand model.

[13]This method can also be applied to address price endogeneity. See Ching (2005) for further details.

The state variables of our model consist of $(E[q_j|I(t)], \sigma_j^2(t), M_{jt-1})_{j=1}^2$. We therefore assume that the detailing policy function depends on these variables. The detailing policy function may also depend on variables that we do not explicitly model. For instance, the total detailing minutes by manufacturer j in the cardiovascular drug category could affect D_j. It is possible that a manufacturer sets its detailing budget for the entire cardiovascular drug category first, and then determines the detailing for individual drugs in the category. We therefore include the total detailing minutes by manufacturer j in the cardiovascular drug category net D_j in the pseudo-detailing policy function.[14] This variable is useful in identifying the parameters associated with detailing in the demand model because it plays the role of exclusion restriction, and essentially serves as an instrumental variable for D_{jt}.

When specifying the pseudo-detailing policy function, one should use a flexible high order polynomial to do the approximation if the sample is large. In practice, however, one may need to make some trade-offs between flexibility and the number of parameters by choosing a functional form carefully. After experimenting with a number of functional forms, we specify the detailing policy function as follows: For $j, k = 1, 2$, and $j \neq k$,

$$\begin{aligned} log(D_{jt}) &= \lambda_{j0} + (\lambda_{j1} + \lambda_{j2} * M_{kt-1}) * (1 - M_{jt-1}) * |\Delta u_{jkt}^q| * \mathbb{I}(\Delta u_{jkt}^q > 0) \\ &\quad +(\lambda_{j3} + \lambda_{j4} * M_{kt-1}) * M_{jt-1} * |\Delta u_{jkt}^q| * \mathbb{I}(\Delta u_{jkt}^q < 0) \\ &\quad +\lambda_{j5} * IV_{jt} + \nu_{jt}, \end{aligned} \tag{18}$$

where

$$\Delta u_{jkt}^q = E[u_{jt}^q|I(t)] - E[u_{kt}^q|I(t)], \tag{19}$$

$$E[u_{jt}^q|I(t)] = -exp(-rE[q_j|I(t)] + \frac{1}{2}r^2(\sigma_j^2(t) + \sigma_\delta^2)), \tag{20}$$

ν_{jt} is the prediction error, and $\mathbb{I}(\cdot)$ is an indicator function. Note that $E[u_{jt}^q|I(t)]$ is part of the expected utility that depends on $E[q_j|I(t)]$ and $\sigma_j^2(t)$.

Our model suggests that manufacturer j has an incentive to increase detailing if $\Delta u_{jkt}^q > 0$. Such an incentive is stronger if M_{jt-1} is small because of the diminishing return of $\partial M_j / \partial D_j$. We therefore interact $(1 - M_{jt-1})$ with $|\Delta u_{jkt}^q|$ when $\Delta u_{jkt}^q > 0$. We expect the coefficient associated with the interaction term to be positive (i.e., $\lambda_{j1} > 0$).

[14] Berndt et al. (2003), use this variable as the instrument for detailing in their reduced form model.

Similarly, when $\Delta u^q_{jkt} < 0$, we interact M_{jt-1} with $|\Delta u^q_{jkt}|$. We expect that manufacturer j would have less incentives to detail when M_{jt-1} is large. However, when M_{jt-1} is small, manufacturer j, if forward-looking, may detail more in order to build up M_j even though $\Delta u^q_{jkt} < 0$. The sign of the coefficient for the interaction term (i.e., λ_{j3}) is therefore ambiguous.

As shown in equation (17), the marginal return of detailing depends on the measure of well-informed physicians for a competing drug as well. Therefore, we also allow M_{kt-1} to interact with M_{jt-1} and Δu^q_{jkt}. According to equation (15), we expect the sign of λ_{j2} and λ_{j4} to be positive if $\Delta s_{jt}(I_k(t)) > \Delta s_{jt}(\underline{I}^p_k)$, and vice versa.

The following two subsections describe the likelihood function and the initial conditions problem. Readers who are not interested in details may skip to Section 5 directly.

4.3 The Likelihood Function

Assuming that the prediction error, ν_{jt}, in equation (18) is normally distributed, we obtain the conditional likelihood of observing D_t,

$$f_d(D_t|(E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}; \theta_s), \tag{21}$$

where θ_s is the vector of parameters.

Assuming that the measurement error, ϵ_{jt}, in equation (18) is normally distributed, and denote $f_n(n_t|D_t, (E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1},\ Size_t; \theta_d)$ as the likelihood of observing n_t conditional on $(D_t, (E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}, Size_t)$. The joint likelihood of observing (n_t, D_t) is simply the product of $f_n(n_t|D_t, .)$ and $f_d(D_t|.)$:

$$\begin{aligned} &l(n_t, D_t|(E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}, Size_t; \theta_d, \theta_s) = \\ &\quad f_n(n_t|D_t, (E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}, Size_t; \theta_d) f_d(D_t|(E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}; \theta_s). \end{aligned} \tag{22}$$

Now note that $\sigma_j(t)$ is a function of $\{n_{j\tau}\}^{t-1}_{\tau=1}$ (see (7)). Therefore, one can rewrite (22) as,

$$\begin{aligned} &l(n_t, D_t|(E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}, Size_t; \theta_d, \theta_s) = \\ &\quad l(n_t, D_t|(E[q_j|I(t)], \{n_{j\tau}\}^{t-1}_{\tau=1}, M_{jt-1})^2_{j=1}, Size_t; \theta_d, \theta_s). \end{aligned} \tag{23}$$

The likelihood of observing $n = \{n_t\}_{t=1}^T$ and $D = \{D_t\}_{t=1}^T$ is,

$$L(n, D|\{E[q|I(\tau)], M_{\tau-1}, Size_\tau\}_{\tau=1}^T; \theta_d, \theta_s) = \prod_{t=1}^T l(n_t, D_t|E[q|I(t)], \{n_\tau\}_{\tau=1}^{t-1}, M_{t-1}, Size_t; \theta_d, \theta_s). \tag{24}$$

But $E[q|I(t)]$ is unobserved to the econometrician and therefore must be integrated over to form the unconditional sample likelihood for (n, D). Evaluating such an integral numerically is very difficult. It involves high order integrals because $E[q|I(t)]$ is autocorrelated. We resolve this problem by using the method of simulated maximum likelihood. The details of the simulation procedures are similar to Ching (2005).

4.4 Initial Conditions Problem

Notice that both Vaseretic and Zestoretic were introduced before March 1993, the first period of our data set. Therefore, we do not observe the initial values of the state variables at $t = 1$: G_{j0}^I, $E[q_j|I(1)]$ and $\sigma_j(1)$. Given this initial conditions problem, consistent estimation for fixed T requires integration over the joint unconditional distribution of the state variables at $t = 1$. As discussed in Heckman (1981), this integration is extremely difficult. It requires us to explicitly incorporate complete dynamic equilibrium since the inception of both drugs into the estimation procedure. As discussed above, this approach is not computationally feasible at this point.

We therefore adopt a middle-ground approach. We set $(D_{jt_j^I}, ..., D_{j0})$ equal to the average D_{jt} for the first 30 observations, where t_j^I is the period that drug j is introduced. In other words, for $t = t_j^I, ..., 0$, we set $D_{jt} = \bar{D}_j$, where $\bar{D}_j = \frac{\sum_{t=1}^{30} D_{jt}}{30}$. Also, for $t = t_j^I, ..., 0$, we set p_{jt} at the average observed values. For the size of market, we first run a linear regression of the size of market on a constant and time trend and then use the predicted values to fill in $Size_t$, for $t = t_j^I, ..., 0$. Given the imputed values of $(D_{jt_j^I}, ..., D_{j0})$, $(p_{jt_j^I}, ..., p_{j0})$, and $(Size_{t_j^I}, ..., Size_0)$, we use our physician's choice model to simulate the unconditional joint distribution of $(G_{j0}^I, E[q_j|I(1)], \sigma_j(1))$, which is then incorporated in our likelihood function.

5 Results

5.1 Parameter Estimates

We now discuss the parameter estimates. Recall that we treat Vaseretic and Zestoretic as inside goods because they compose more than 80% of the demand for the ACE-inhibitor with diuretic. We combine all other drugs that belong to ACE-inhibitor with diuretic, ACE-inhibitor, and Thiazide Diuretic as the outside good. For identification reasons, we need to normalize the scaling parameter for the number of consumption experience signals, κ, the intercept term for the utility of the outside good, α_0, and the true mean quality of Vaseretic, q_1. We set $\kappa = 1/30000$, and $\alpha_0 = q_1 = 0$. For simplicity, we also restrict $\underline{I}_j^o = \underline{I}_j^p \equiv \underline{I}_j$ and $\underline{\sigma}_j^o = \underline{\sigma}_j^p \equiv \underline{\sigma}, \forall j$ because we do not observe the data during the initial part of the product lifecycle, which is important in identifying their difference. We refer to $\underline{I}$ as the market initial prior.

Table 2 shows the parameter estimates. Model 1 refers to the model presented above. Drug 1 is Vaseretic (incumbent) and drug 2 is Zestoretic (entrant). The time trend of the outside good (π_t) is negative and significant, indicating that the value of the outside good relative to inside goods is declining over time. This is consistent with the continuous expansion of demand for both Vaseretic and Zestoretic, as shown in Figure 1. The parameter estimates for the true mean quality and the initial priors are all statistically significant. The true mean quality of Zestoretic (q_2) is 29.04, which is higher than that of Vaseretic (q_1). The initial prior mean qualities of Vaseretic and Zestoretic are -10.24 and -18.92, respectively, which are lower than their true mean qualities. This indicates that the market has pessimistic priors about both drugs when they are first introduced into the market. It should also be noted that the initial prior mean quality for Vaseretic is better than that for Zestoretic.

All of the preference parameter estimates are statistically significant. The price coefficient is not significant. This is not surprising because, as mentioned before, Canada provides prescription drug coverage to patients who are 60 or older, and most of the patients who have hypertension are elderly. The risk coefficient (r) is positive and significant, indicating risk-averse behavior. In other words, an increase in the perceived variance of a product will lower the ex-

pected utility of choosing it. However, the estimate for r is 0.05, which is quite small. Given the functional form of the utility function, this implies that $E[q_j|I(t)]$ carries significantly more weight than $\sigma_j(t)$ in physicians' choice.

The parameters associated with the measure of well-informed physicians are all statistically significant. The estimate for β_0 is -1.42, which implies that nearly 20 percent of physicians will be well-informed about $I_j(t)$ (i.e., $M_j = 0.2$) when $G_j^I = 0$. This represents the percentage of physicians who keep up with the most updated information about ACE-inhibitor with diuretic themselves even without any help from detailing. Recall that the average rate of forgetting is a non-linear function of M_{jt-1}, which exhibits an inverted-U shape. The estimate of ϕ_I is close to 3%. The implied average rate of forgetting is shown in Figure 3. The average rate of forgetting starts from 0% at around $M_{jt-1} = 0.2$. It increases and reaches the maximum of 2.1% at around $M_{jt-1} = 0.6$, and then declines. The estimate of β_1 is 5.80e-05. In Figure 4 we plot its implied rate of building M_{jt} without forgetting (i.e., $\phi_I = 0$), conditioning on M_{jt-1} and $D_{jt} = 1300$, which is the average per period detailing for both Vaseretic and Zestoretic in our sample. The rate of building M_{jt} starts off at slightly above 6% when M_{jt-1} is around 0.2 (i.e., $G_I = 0$). Then it declines almost linearly at the rate of 0.775% per 0.1 increase in M_{jt-1}.

Measures of well-informed physicians, expected qualities and perceived variances play crucial roles in our model. They are also potentially important for marketing managers, who need to make strategic decisions on how to allocate their sales forces. Although these variables are not directly observed in the data, having explicitly modeled how these elements influence physicians' choice, we are able to recover them from the evolution of market shares and detailing data. Figure 5 shows the evolution of the measures of well-informed physicians during the sample period. For Vaseretic, the measure of well-informed physicians starts off at around 0.57. It increases to 0.7 after 30 months, and then gradually reduces to around 0.55 at the end of the sample period. For Zestoretic, the measure of well-informed physicians increases from 0.3 to around 0.85. Figure 6 shows how $E[q_j|I(t)]$ evolves during the sample period. For Vaseretic, it increases slowly from around -5 to -2. For Zestoretic, it increases at a much faster rate from -18 to 23.[15]

[15]Since our estimate of r implies that $\sigma_j^2(t)$ does not play an important role in physicians' choice, we do not report the evolution of $\sigma_j^2(t)$ in the interest of space. It is available upon request.

As for the pseudo-detailing policy functions, most of the parameters are statistically significant except $\lambda_{13}, \lambda_{14}, \lambda_{15}$, and λ_{22}. The instrumental variable for Zestoretic (λ_{25}) is positive and significant while the instrumental variable for Vaseretic (λ_{15}) is not significant. Both λ_{11} and λ_{21} are positive, suggesting that manufacturers respond to favorable information about their own drugs by increasing the amount of detailing. λ_{23} is positive, which is consistent with Zestoretic being the entrant. This suggests that the incentive to detail in order to build up M is stronger than the disincentive to detail due to $\Delta u^q_{21t} < 0$. This is consistent with our parameter estimates, which imply $|\Delta u^q_{21t}| * \mathbb{I}(\Delta u^q_{21t} < 0)$ is shrinking over time.

Also, both λ_{j2} and λ_{j4} are negative for $j = 1, 2$. This implies that D_{jt} decreases as M_{kt-1} increases. If manufacturers are rational and use the marginal impact of detailing on current demand as a guide for determining their detailing efforts, the estimated signs of λ_{j2} and λ_{j4} suggest that the marginal impact of detailing on demand would decrease as M_{kt} increases. Interestingly, using our parameter estimates, we simulate sequences of $(\Delta s_{jt}(\underline{I}^p_k), \Delta s_{jt}(I_k(t)))$, and find that $\Delta s_{jt}(\underline{I}^p_k) > \Delta s_{jt}(I_k(t))$ for all j, k and t. It follows from equation (17) that the implied marginal return of detailing indeed decreases as M_{kt} increases. Overall, our results suggest that the endogeneity problem of detailing exists in this market.

5.2 Goodness-of-fit

Our estimated model provides a good fit to the data. To illustrate this, we simulate 5000 sequences of quantity demanded (expressed in terms of number of prescriptions) for both Vaseretic and Zestoretic using the demand model and the pseudo-detailing policy functions. We compute the average predicted quantity by averaging simulated quantities. Figures 7 and 8 plot the average predicted demand and the actual demand for Vaseretic and Zestoretic, respectively. In general, the model is able to fit the diffusion pattern of demand very well. This indicates that even though we only have four types of physicians in our model, it is flexible enough to fit the data. Figures 9 and 10 plot the average predicted detailing minutes and the actual ones for Vaseretic and Zestoretic, respectively. As we can see, the average predicted detailing minutes is able to capture the data trend reasonably well. In particular, the average predicted detailing minutes is able to mimic the observed fluctuation for Zestoretic. This is mainly due to the pos-

itive correlation between detailing for Zestoretic and its instrument (total detailing minutes by Zestoretic's manufacturer in the cardiovascular category net the detailing minutes for Zestoretic) used in the pseudo-detailing policy function.

5.3 Effectiveness of Detailing

5.3.1 The effect of a temporary increase in detailing

Measuring the effectiveness of detailing is important for managers because they often need to decide how to allocate their sales forces. In this subsection, we discuss the effectiveness of detailing using our parameter estimates. It is worth reiterating that M_{jt} and $E[q_j|I(t)]$ play important roles in determining the marginal return of detailing in our model. Although these variables are not directly observed in the data, we are able to use the estimates of our structural parameters to generate them. We will first illustrate how the marginal impact of detailing on current demand depends on them.

Notice that the marginal return of detailing for drug j not only depends on $I_j(t)$ and M_{jt}, but also $I_{-j}(t)$ and M_{-jt}. To simplify the illustration, we set $M_{1t} = M_{2t}$ for all t. In the baseline case, we simulate 5000 histories of demand and $I(t)$ by setting $D_{1t} = D_{2t} = 1300$ (the average observed amount of detailing across both drugs) for $t \geq 1$, and p_{jt} at the average observed values for all t. Recall that Vaseretic and Zestoretic enter the market before $t = 1$ (when our sample begins). To ensure $M_{1t} = M_{2t}$ and obtain the initial value of the information sets at $t = 1$, we also set $M_{1t} = M_{2t} = 0.5$ for $t < 1$ in our baseline simulation. For $t \geq 1$, M_{jt} is determined by D_{jt}. We evaluate the effects of a one-time increase in detailing at three different points in time, based on the average expected qualities in the baseline simulation: (i) $t = 1$ when the average expected quality for Vaseretic is higher; (ii) $t = 23$ when the average expected qualities are about the same for both drugs; (iii) $t = 60$ when the average expected quality for Zestoretic is higher. In each case, we increase the detailing amount by 50% for one of the drugs, holding the other one fixed, and examine its effect on current demand.

Panel 1 of Table 3 shows the results. For Vaseretic, the percentage changes in current demand are 0.348%, 0.417%, and 0.414% at $t = 1$, 23, and 60, respectively. The effect at $t = 23$ is higher than that at $t = 1$, mainly because $E[q_1|I(t)]$ increases from -5.52 to -3.68 during that period. However, the effect at $t = 60$ is about the same as that at $t = 23$ despite the fact that $E[q_1|I(t)]$ improves from -3.68 to -2.06. One reason is that the average expected quality of Zestoretic improves from -3.11 to 19.79 during that period, which is much more than that of Vaseretic. This makes Vaseretic less attractive to physicians. Another reason is that there is diminishing return in building up M_1. During that period, M_1 increases from 0.64 to 0.73. According to equation (17), a lower return in building up M results in a smaller effect of detailing on current demand. We find a similar pattern for Zestoretic. The percentage changes in current demand are 0.283%, 0.996%, and 0.903% at $t = 1$, 23, and 60, respectively.

It should be noted that at $t = 23$, the percentage change in current demand is much larger for Zestoretic (0.996%) than for Vaseretic (0.417%) although the average expected qualities of Vaseretic and Zestoretic are about the same. This is because the initial prior for Zestoretic's quality is lower than that for Vaseretic's. Consequently, it follows from equation (17) that the marginal impact of detailing is higher for Zestoretic.

The magnitudes of our detailing elasticities are consistent with Berndt et al. (1997). According to their estimates, the upper bound of the elasticity of demand with respect to cumulative detailing minutes ranges from 0.67 to 0.92.[16] In our simulation above, a 50% increase in detailing corresponds to increases of 2.6%, 1.9%, and 1.6% in cumulative detailing minutes

[16]Berndt et al. (1997) estimates the following equation using the data on anti-ulcer drugs in the U.S.:

$$log\left(\frac{n_{jt}}{n_{1t}}\right) = \beta \cdot log\left(\frac{G^I_{jt}}{G^I_{1t}}\right) + \cdots, \tag{25}$$

where n_{jt} is the sales of drug j at time t, G^I_{jt} is the cumulative detailing minutes of drug j at time t, and drug 1 is the first entrant in this market. This equation implies that

$$\varepsilon_{jj} = \beta + \varepsilon_{1j}, \tag{26}$$

where ε_{jk} is the elasticity of demand for drug j with respect to cumulative detailing minutes of drug k. If $\varepsilon_{jk} < 0$ for $j \neq k$, β is the upper bound of ε_{jj}.

at $t = 1, 23$, and 60, respectively. Thus our elasticity of demand with respect to cumulative detailing minutes falls in a range between 0.1 and 0.6.[17]

5.3.2 Policy Experiment: A campaign that encourages sharing drug experiences

We now turn to discuss a policy experiment. In order to enhance the speediness of updating the safety profile of drugs, public health agencies have been considering various measures to encourage health care professionals and patients to share their drug experiences with them. For example, Health Canada set up a program called MEDEffect to promote awareness about the importance of filing reports using their on-line report system for the general public. It is likely that such a program would increase the portion of experience signals revealed to the public (correspond to an increase in κ in our model). How should marketing managers respond to this kind of campaign? Given that we have obtained estimates for the structural parameters, we are able to examine how the effectiveness of detailing would change if the government implements such a campaign. To illustrate this, we re-simulate the effects of detailing in our model using the procedure above by doubling the value of κ. Panel 2 of Table 3 shows the results. Compared with the baseline case in Panel 1 of Table 3, the information set, $I(t)$, has improved much quicker, and the marginal returns of detailing from the current demand are also higher at $t = 1, 23$, and 60. In particular, the increases in the effectiveness of detailing are higher in the earlier part of the product lifecycle. Given these results, marketing managers should consider increasing the amount of detailing in this market if this campaign is carried out.

It is important to understand the intuition behind these results. They are mainly driven by the pessimistic initial prior in this market. As more experience signals are revealed in each period under this campaign, the expected qualities are revised upward more quickly over time. Consequently, this shifts up the effectiveness of detailing. Following this argument, it should be emphasized that the effectiveness of detailing could shift down under this campaign if a market has optimistic initial prior about drug qualities. In that case, the expected qualities will be revised downward more quickly over time.

[17] We do not compare our detailing elasticity with those implied by Narayanan et al. (2005) and Mukherji (2002) because they use detailing expenditures instead of detailing minutes, which is used in our paper.

The discussion above again highlights the difference between our model and the traditional learning models, which assume that advertising/detailing signals and consumption experience signals are substitutes for each other in updating the prior belief about product qualities. In those models, increasing the value of κ will cause the marginal return of advertising/detailing to decrease, which suggests that managers should reduce their advertising/detailing efforts. This is just the opposite of what our model suggests, given our parameter estimates.

5.4 Robustness Checks

5.4.1 The Importance of Endogeneity of Detailing

Our estimates in the pseudo-detailing policy function suggest that detailing is endogenous. However, it is hard to assess the economic significance of the endogeneity problem from the estimates. To investigate the extent of the parameter bias if one fails to take the endogeneity problem of detailing into account, we re-estimate the demand model without using the pseudo-detailing policy functions. The parameter estimates are reported in Table 2, under Model 2 (demand only model). The estimate for β_1 is 6.74e-05. This is higher than the estimate from the base model (i.e., Model 1), which is 5.80e-05. The depreciation rate of the detailing stock, ϕ_I, is 0.022. This is lower than the estimate 0.029 in the base model. A likelihood ratio test rejects the hypothesis that the estimates of $(\beta_0, \beta_1, \phi_I)$ in the base model are the same as those in Model 2. This suggests that the estimated marginal return of detailing is biased upward if we do not take the endogeneity problem into account. To show the extent of the bias, we plot the implied average rate of forgetting from the demand only model in Figure 3, and the implied rate of building M in Figure 4. The average rate of forgetting is biased downward, with its peak at 1.5% instead of 2.1%; the rate of building M is biased upward, starting at around 7% instead of 6%.

To understand how the bias would affect the estimates of the effectiveness of detailing, we repeat the exercise in Section 5.3 by using the parameter estimates from Model 2. We use the same simulated values of $I(t)$ and M_{jt-1} at $t = 1$, 23, and 60 from the baseline simulation in Panel 1 of Table 3. Conditional on these simulated $I(t)$ and M_{jt-1}, we use the parameter

estimates from Model 2 to simulate the effect of the one-time temporary increase in detailing. The results, reported in Panel 3 of Table 3, confirm that the effectiveness of detailing would be biased upward if we do not take the endogeneity into account. The extent of bias also appears to be quite significant.

5.4.2 The Role of Persuasive Detailing

Sales representatives often give away gifts during their visits. Critics argue that these gifts may affect physicians' prescribing behavior although many physicians disagree. Our model so far does not take this kind of "bribery" effect into account. One difficulty is that there is no data on the amount of gifts given by sales representatives. Nevertheless, it is plausible that detailing and the amount of gifts are positively correlated. As a robustness check, we follow Nerlove and Arrow (1962) by incorporating a detailing goodwill stock, G^P, directly into the utility function. Following the economics literature (e.g., Leffler 1981, Hurwitz and Caves 1988), we refer to G^P as the *persuasive detailing goodwill stock*.[18] We modify the physicians' utility function as follows:

$$E[U_{hijt}|I^h(t)] = E[u_{ijt}|I^h(t)] + \gamma G^P_{jt}, \tag{27}$$

where γ captures the effect of G^P_{jt}. We construct G^P_{jt} in the same way as G^I_{jt}, but with a different depreciation rate ϕ_P. Following Narayanan et al. (2005), we set ϕ_P to be 30% when estimating other parameters. The estimation results are reported in Table 2, under Model 3. The coefficient for the persuasive detailing stock, γ, is negative, but small and insignificant. This suggests that persuasive detailing does not play an important role in influencing physicians' choice in our framework. The log-likelihood of Model 3 turns out to be very close to that of Model 1 – the difference is less than 1. A likelihood ratio test rejects the hypothesis that γ is different from zero. We interpret this as evidence that our results about how detailing depends on $I(t)$ and M_t are robust, and the persuasive component of detailing is negligible in the Canadian ACE-inhibitor with diuretic market.

[18]Some researchers (e.g., Anand and Shachar 2005, Narayanan et al. 2005) refer to G^P as the "direct" effect of advertising/detailing. Some (Becker and Murphy 1993, Ackerberg 2001) refer to it as the prestige effect.

Although our results may seem to contrast with those of Narayanan et al. (2005), who find that the persuasive detailing stock in the utility function is positive and significant, our results do not necessarily contradict their findings. Unlike our model, which separates the reminding role of detailing from the "bribery" role, they combine both of them via G^P in the utility function. Thus, the significant and positive effect of G^P found in their estimation could be mainly due to its reminding role. This demonstrates an advantage of modeling the reminding role of detailing more structurally: We are able to give more precise interpretations to our parameters. Moreover, our results suggest that the modeling approach adopted here could avoid some counter-intuitive implications by the traditional approach of incorporating advertising/detailing in the utility function: Manufacturers can always raise their prices and maintain the demand unchanged as long as they spend enough on advertising/detailing.

6 Conclusion

In this paper, we develop a new structural model of physicians' prescribing decisions and detailing under quality uncertainty, which can be estimated using product level sales and detailing data. We introduce a representative opinion leader, whose role is to update the most current information about drug qualities based on past consumption experiences. Unlike the previous literature which assumes detailing is a way to convey noisy signals about the true quality of the drug to physicians, we assume that detailing changes the measure of physicians who are informed of the current public information sets maintained by the representative opinion leader. This allows our model to directly link the marginal return of detailing to the measure of well-informed physicians and current public information sets. We also explicitly model physician forgetting by allowing the measure of well-informed physicians to decrease if current detailing is not sufficiently large.

We estimate our model using product level data on the ACE-inhibitor with diuretic market in Canada. Our estimation approach, which makes use of a pseudo-detailing policy function, allows us to control for the potential endogeneity of detailing. The results show that our model is able to fit the diffusion pattern very well, the effectiveness of detailing depends on the current

information set and the measure of well-informed physicians, and the role of detailing-in-utility is minimal. We examine how a public awareness campaign, which encourages physicians/patients to report their drug experiences, would affect managerial incentives to detail. Given our parameter estimates, our model suggests that managers should increase the detailing efforts, whereas the previous learning models suggest that managers should reduce the detailing efforts.

One limitation of this paper is that we do not explicitly incorporate data from clinical trials outcomes and side-effect information. Conceivably, such data will be very valuable for analyzing the effects of detailing. Also, we do not model how direct-to-consumer advertising, journal advertising, free samples, and educational meetings or conferences sponsored by drug companies may affect pharmaceutical demand. We leave modeling the role of these marketing communication methods in the environment with two-sided learning for future research.

Another limitation is that we do not allow for heterogeneous opinion leaders in our model. Some opinion leaders may obtain more past patients' experiences than others, (perhaps some work for larger hospitals and therefore are able to collect more patients' experiences) and as a result, they may possess different public information sets representing their various levels of learning. Physicians may receive more influence from opinion leaders who are located in their neighborhoods. Although these are attractive features, unfortunately, incorporating them will dramatically complicate the model. One would also need a richer data set to estimate such a model. Instead, our approach of using a representative opinion leader leads to a tractable model which can be estimated simply using product level data. We hope future research will extend our framework to allow for multiple representative opinion leaders. Another interesting research direction is to use individual level data to examine the role of opinion leaders. A recent study by Bhatia, Manchanda and Nair (2006) is taking this important step to examine the effects of heterogeneous opinion leaders on physician decisions.

Our model can potentially help a marketing manager evaluate the future return of alternative long-term detailing strategies. Conditional on his/her own future detailing strategies and his/her rivals' future detailing strategies, we can take the uncertainty about true quality into account by integrating out the prior distributions of q. However, when the marketing manager changes his/her own detailing strategies, it is likely that his/her rivals will react and change

theirs as well. Although our pseudo-detailing policy function approach allows us to correct the endogeneity problem, it does not allow us to predict how rivals react when one changes his/her own detailing strategy due to its reduced form nature. In order to utilize our demand model to evaluate alternative future detailing strategies, we would need to combine it with a supply side model explicitly. By developing a tractable demand side model, we hope that our framework has laid some groundwork for this challenging research direction.

Finally, although we present our model in the context of pharmaceutical demand, it could also be applied to other markets such as movies, video games, softwares, restaurants, etc., where both sides of the market are uncertain about how new products will perform, and opinion leaders (e.g., professional critics) may play an important role in influencing consumer purchase decisions. Given that data on reviews and critics are typically available in the public domain, it is surprising that structural modeling of opinion leaders is relatively scarce. Our model could be used as a starting point to analyze their roles and potentially improve our understanding about how information is transmitted in markets other than prescription drugs.

References

Ackerberg, Daniel A. 2001. Empirically distinguishing informative and prestige effects of advertising. *RAND Journal of Economics* **32** 100–118.

Ackerberg, Daniel A. 2003. Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination. *International Economic Review* **44**(3) 1007–1040.

Anand, Bharat N., Ron Shachar. 2005. Advertising, the Matchmaker. Working paper, Faculty of Management, Tel Aviv University.

Azoulay, Pierre. 2002. Do Pharmaceutical Sales Respond to Scientific Evidence? *Journal of Economics and Management Strategy* **11**(4) 551–594.

Becker, Gary S., Kevin M. Murphy. 1993. A Simple Theory of Advertising as a Good or Bad. *Quarterly Journal of Economics* 942–964.

Bergemann, Dirk, Juuso Valimaki. 1997. Market Diffusion with Two-Sided Learning. *RAND Journal of Economics* **28**(4) 773–795.

Bergemann, Dirk, Juuso Valimaki. 2006. Dynamic Pricing of New Experience Goods. *Journal of Political Economy* **114**(4) 713–743.

Berndt, Ernst, Linda T. Bui, David H. Reiley, Glen L. Urban. 1997. The Role of Marketing, Product Quality and Price Competition in the Growth and Composition of the U.S. Anti-Ulcer Drug Industry. *The economics of new goods, NBER Studies in income and wealth*, vol. 58. University of Chicago Press, Chicago and London, 277–322.

Berndt, Ernst R., Robert S. Pindyck, Pierre Azoulay. 2003. Consumption Externalities and Diffusion in Pharmaceutical Markets: Antiulcer drugs. *Journal of Industrial Economics* **51**(2) 243–270.

Berry, Steven T., James Levinshohn, Ariel Pakes. 1995. Automobile Prices in Market Equilibrium. *Econometrica* **63**(4) 841–890.

Bhatia, Tulikaa, Puneet Manchanda, Harikesh Nair. 2006. Asymmetric Peer Effects in Physician Prescription Behavior: The Role of Opinion Leaders. Working paper, Graduate School of Business, Stanford University.

Caves, Richard E., Mark A. Hurwitz. 1988. Persuasion or Information? Promotion and the Shares of Brand Name and Generic Pharmeceuticals. *Journal of Law and Economics* **31**.

Caves, Richard E., Mihchael D. Whinston, Mark A. Hurwitz. 1991. Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry. *Brookings Papers on Economic Activity, Microeconomics* **1** 1–48.

Chernozhukov, Victor, Hang Hong. 2003. An MCMC approach to Classical Estimation. *Journal of Econometrics* **115**(2) 293–346.

Ching, Andrew. 2000. Dynamic Equilibrium in the US Prescription Drug Market After Patent Expiration. Ph.D. dissertation, University of Minnesota.

Ching, Andrew. 2004. A Dynamic Oligopoly Structural Model for the US Prescription Drug Market After Patent Expiration. Working paper, Rotman School of Management, University of Toronto.

Ching, Andrew. 2005. Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs After Patent Expiration. Working paper, Rotman School of Management, University of Toronto.

Coleman, James S., Elihu Katz, Herbert Menzel. 2004. *Medical Innovation: A Diffusion Study*. The Bobbs-Merrill Company, Indianapolis, I.N.

Cooper, Richelle J., David L. Schriger, Roger C. Wallace, Vladislav J. Mikulich, Michael S. Wilkes. 2003. The quantity and quality of scientific graphs in pharmaceutical advertisements. *Journal of General Internal Medicine* **18** 294–297.

Crawford, Gregory S., Matthew Shum. 2005. Uncertainty and Experimentation in Pharmacetical Demand. *Econometrica* **73** 1137–1174.

DeGroot, Morris H. 1970. *Optimal Statistical Decisions*. McGraw-Hill, New York.

Dong, Xiaojing, Puneet Manchanda, Pradeep K. Chintagunta. 2006. Qualifying the Benefits of Individual Level Targeting in the Presence of Firm Strategic Behavior. Working paper, Graduate School of Business, University of Chicago.

Elgie, Robert G. 2001. Enhancing Transparency in Drug Price Regulation. Mimeo, Patented Medicine Prices Review Board.

Erdem, Tulin, Michael P. Keane. 1996. Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent consumer Goods Markets. *Marketing Science* **15**(1) 1–20.

Gonul, Fusun F., Franklin Carter, Elina Petrova, Kannan Srinivasan. 2001. Promotion of Prescription Drugs and Its Impact on Physicians' Choice Behavior. *Journal of Marketing* **65** 79–90.

Greider, Katharine. 2003. *The Big Fix: How the Pharmaceutical Industry Rips off American Consumers*. PublicAffairs, New York, N.Y.

Haug, James D. 1997. Physicians' Preferences for Information Sources: A Meta-Analytic Study. *Bulletin of the Medical Library Association* **85**(3) 223–232.

Heckman, James. 1981. Heterogeneity and State Dependence. *Studies in Labor Markets*. University of Chicago Press, Chicago and London, 91–139.

Hitsch, Gunter J. 2006. An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty. *Marketing Science* **25**(1) 25–50.

Hurwitz, Mark, Richard Caves. 1988. Persuasion or Information? Promotion and the Shares of Brand Name and Generic Pharmaceuticals. *Journal of Law and Economics* **31** 299–320.

Lasser, Karen, Paul Allen, Steffie Woolhandler, David Himmelstein, Sidney Wolfe, David Bor. 2002. Timing of New Black Box Warnings and Withdrawals for Prescription Medications. *Journal of American Medical Association* **287**(17) 2215–2220.

Leffler, Keith. 1981. Persuasion or Information? The Economics of Prescription Drug Advertising. *Journal of Law and Economics* **24** 45–74.

Mehta, Nitin, Surendra Rajiv, Kannan Srinivasan. 2004. Role of Forgetting in Memory-Based Choice Decisions: a structural model. *Quantitative Marketing and Economics* **2** 107–140.

Mukherji, Prokriti. 2002. A Structural Model of Learning For New Products. Mimeo, Marshall School of Business, University of Southern California.

Mullainathan, Sendhil. 2002. A Memory-Based Model of Bounded Rationality. *Quarterly Journal of Economics* **117**(3) 735–774.

Narayanan, Sridhar, Ramarao Desiraju, Pradeep K. Chintagunta. 2004. Return on Investment Implications for Pharmaceutical Promotional Expenditures: The Role of Marketing-Mix Interactions. *Journal of Marketing* **68** 90–105.

Narayanan, Sridhar, Puneet Manchanda. 2006. Heterogeneous Learning and the Targeting of Marketing Communication for New Products. Working paper, Graduate School of Business, Stanford University.

Narayanan, Sridhar, Puneet Manchanda, Pradeep K. Chintagunta. 2005. Temporal Differences in the Role of Marketing Communication in New Product Categories. *Journal of Marketing Research* **42**(3) 278–290.

Nerlove, Marc, Kenneth J. Arrow. 1962. Optimal Advertising Policy Under Dynamic Conditions. *Economica* **29**(114) 129–142.

Osinga, Ernst C., Peter S.H. Leeflang, Jaap E. Wieringa. 2007. The Persistence of Marketing Effects over the Brand's Life Cycle. Working paper, University of Groningen.

Rizzo, John. 1999. Advertising and Competition in the Ethical Pharmaceutical Industry. *Journal of Law and Economics* **42** 89–116.

Thompson, Margaret L. 1997. Characteristics of Information Resources Preferred by Primary Care Physicians. *Bulletin of the Medical Library Association* **85**(2) 187–192.

Venkataraman, Sriram, Stefan Stremersch. 2006. The Debate on Influencing Doctors' Decisions: Are Drug Characteristics the Missing Link? Working paper, Department of Business Economics, School of Economics, Erasmus University Rotterdam.

Villas-Boas, J. Miguel. 2006. Dynamic Competition with Experience Goods. *Journal of Economics and Management Strategy* **15**(1) 37–66.

Wosinska, Marta. 2002. Just What the Patient Ordered? Direct-to-Consumer Advertising and the Demand for Pharmaceutical Products. HBS Marketing Research Paper No. 02-04.

Ziegler, M. G., P. Lew, B. C. Singer. 1995. The Accuracy of Drug Information from Pharmaceutical Sales Representatives. *The Journal of the American Medical Association* **273** 1296–1298.

Table 1: Summary statistics

	Brand	Mean	Standard deviation	Max	Min
Number of prescriptions	Vaseretic	4,007.63	676.80	5,446	2,429
	Zestoretic	6,388.75	4,900.28	16,330	322
Detailing Minutes	Vaseretic	1,032.63	689.11	3,240	97
	Zestoretic	1,627.08	828.67	4,203	93
Price	Vaseretic	40.54	8.76	69.21	24.45
	Zestoretic	34.29	8.65	61.48	15.74

Table 2: Parameter estimates

1 - Vaseretic (incumbent)
2 - Zestoretic (entrant)

	Model 1 (base)		Model 2 (demand only)		Model 3 (w' persuasive effect)	
	estimates	s.e.	estimates	s.e.	estimates	s.e.
Learning parameters						
σ_δ^2	**8.410**	0.566	**7.560**	0.122	**8.416**	0.174
$\underline{q}_1$	**-10.237**	0.888	**-12.153**	2.302	**-10.055**	1.034
$\underline{q}_2$	**-18.916**	0.418	**-19.542**	2.534	**-18.892**	0.345
$\underline{\sigma}^2$	**3.479**	0.262	**3.654**	0.093	**3.487**	0.213
q_1	0		0		0	
q_2	**29.038**	1.529	**20.106**	3.618	**29.721**	3.807
κ	1/30000		1/30000		1/30000	
Preference parameters						
α	**-3.893**	0.063	**-3.864**	0.040	**-3.892**	0.038
r	**0.049**	0.000	**0.046**	0.006	**0.049**	0.001
π_p	4.98E-04	5.12E-04	5.72E-04	4.04E-04	5.22E-04	5.48E-04
π_t	**-0.005**	6.51E-04	**-0.006**	3.31E-04	**-0.006**	4.73E-04
γ	0		0		-1.09E-06	5.78E-06
Detailing stock parameters						
Φ_P	0		0		0.3	
Φ_I	**0.029**	0.003	**0.022**	0.006	**0.030**	0.006
β_0	**-1.420**	0.183	**-1.360**	0.095	**-1.483**	0.336
β_1	**5.80E-05**	9.79E-06	**6.74E-05**	1.06E-05	**5.98E-05**	9.50E-06
Other parameters for error terms						
s.d.(ε)	**180.673**	11.272	**171.320**	7.452	**180.083**	11.562
s.d.(ζ)	1		1		1	
s.d.(e)	**0.729**	0.021	**0.621**	0.036	**0.732**	0.023
Pseudo-detailing policy functions						
λ_{10}	**7.777**	1.324			**7.785**	1.509
λ_{11}	**2.811**	0.995			**2.645**	1.327
λ_{12}	**-6.166**	1.127			**-5.823**	2.194
λ_{13}	0.708	1.512			0.936	1.923
λ_{14}	-2.831	1.701			-3.158	2.455
λ_{15}	-0.085	0.145			-0.087	0.166
λ_{20}	**0.173**	1.669			**0.179**	0.999
λ_{21}	**19.195**	7.463			17.473	11.284
λ_{22}	-18.698	10.946			-16.876	13.791
λ_{23}	**16.370**	4.031			**16.953**	4.583
λ_{24}	**-22.477**	6.040			**-23.570**	7.351
λ_{25}	**0.690**	0.186			**0.688**	0.119
s.d.(ν)	**0.643**	0.029			**0.644**	0.033
log likelihood	-2110.439		-962.275		-2109.932	

* Estimates shown in bold are significant at 5% level .

Table 3: Effect of a one-time increase in detailing by 50% on current demand

Panel 1

	Increase in detailing for Vaseretic		Increase in detailing for Zestoretic		Average I(t) ($E[q_j \mid I(t)]$, $\sigma_j^2(t)$)		Measure of well-informed physicians in the last period
	% change in demand		% change in demand				
time	Vaseretic	Zestoretic	Vaseretic	Zestoretic	Vaseretic	Zestoretic	
1 (Mar 93)	0.348	-0.097	-0.017	0.283	(-5.52, 1.87)	(-16.97, 3.34)	0.50
23 (Jan 95)	0.417	-0.070	-0.154	0.996	(-3.68, 1.24)	(-3.11, 2.33)	0.64
60 (Feb 98)	0.414	-0.035	-0.282	0.903	(-2.06, 0.70)	(19.79, 0.67)	0.73

Panel 2

	Increase in detailing for Vaseretic		Increase in detailing for Zestoretic		Average I(t) ($E[q_j \mid I(t)]$, $\sigma_j^2(t)$)		Measure of well-informed physicians in the last period
	% change in demand		% change in demand				
time	Vaseretic	Zestoretic	Vaseretic	Zestoretic	Vaseretic	Zestoretic	
1 (Mar 93)	0.464	-0.126	-0.035	0.529	(-3.64, 1.23)	(-14.99, 3.19)	0.50
23 (Jan 95)	0.496	-0.051	-0.274	1.156	(-2.18, 0.73)	(11.39, 1.28)	0.64
60 (Feb 98)	0.445	-0.036	-0.295	0.912	(-1.12, 0.38)	(25.07, 0.29)	0.73

Panel 3

	Increase in detailing for Vaseretic		Increase in detailing for Zestoretic		Average I(t) ($E[q_j \mid I(t)]$, $\sigma_j^2(t)$)		Measure of well-informed physicians in the last period
	% change in demand		% change in demand				
time	Vaseretic	Zestoretic	Vaseretic	Zestoretic	Vaseretic	Zestoretic	
1 (Mar 93)	0.412	-0.194	-0.026	0.381	(-5.52, 1.87)	(-16.97, 3.34)	0.50
23 (Jan 95)	0.510	-0.130	-0.263	1.214	(-3.68, 1.24)	(-3.11, 2.33)	0.64
60 (Feb 98)	0.509	-0.056	-0.519	1.057	(-2.06, 0.70)	(19.79, 0.67)	0.73

Note: Estimates used in Panels 1-3
Panel 1: Model 1
Panel 2: Model 1 except that kappa is doubled
Panel 3: Model 2

Figure 1: Total sales vs time

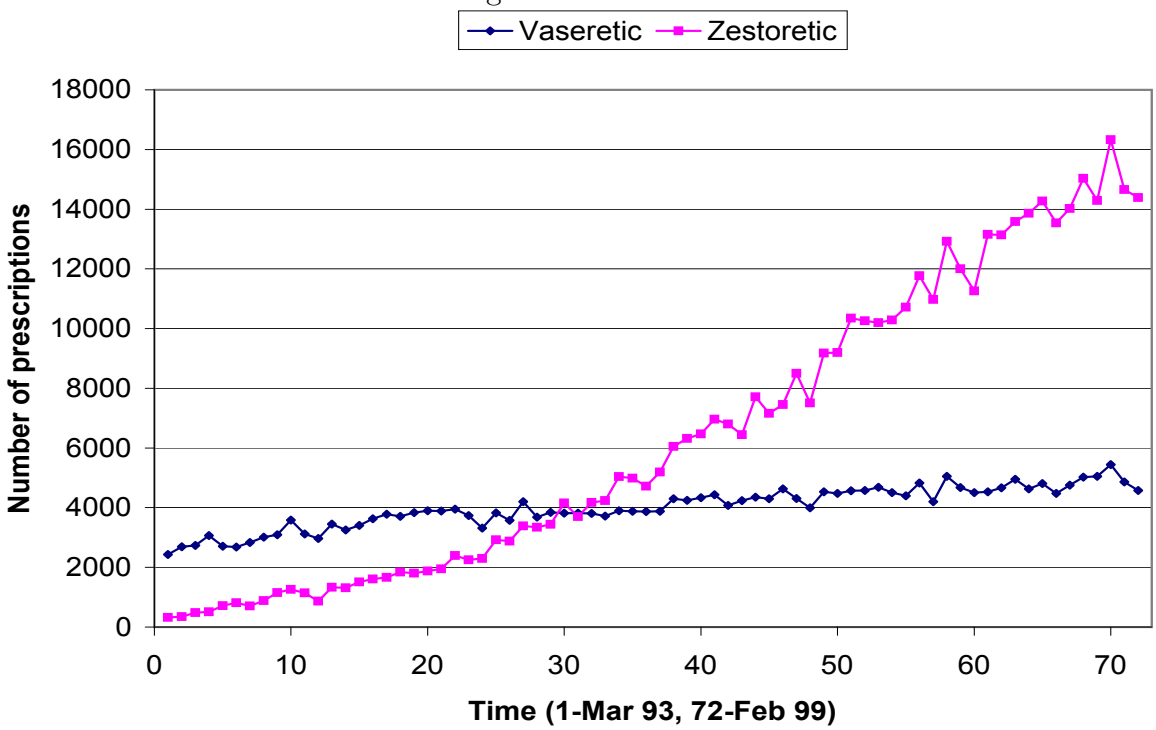

Figure 2: Size of market vs time

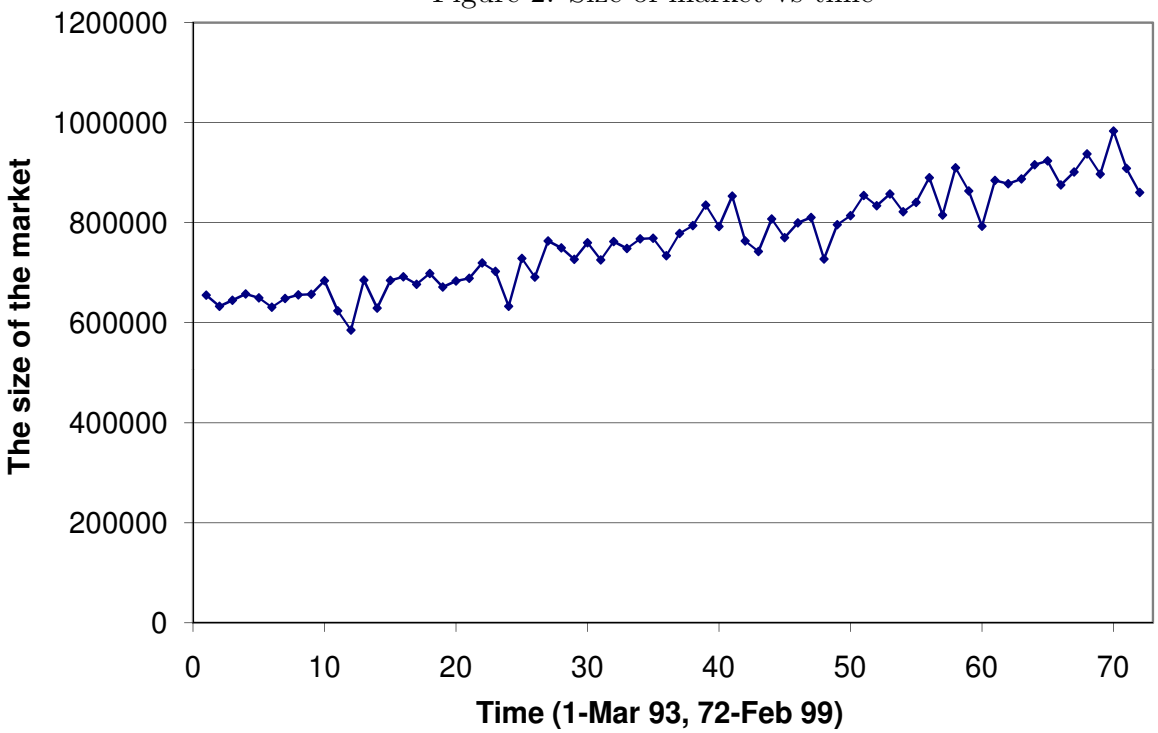

Figure 3: Rate of forgetting

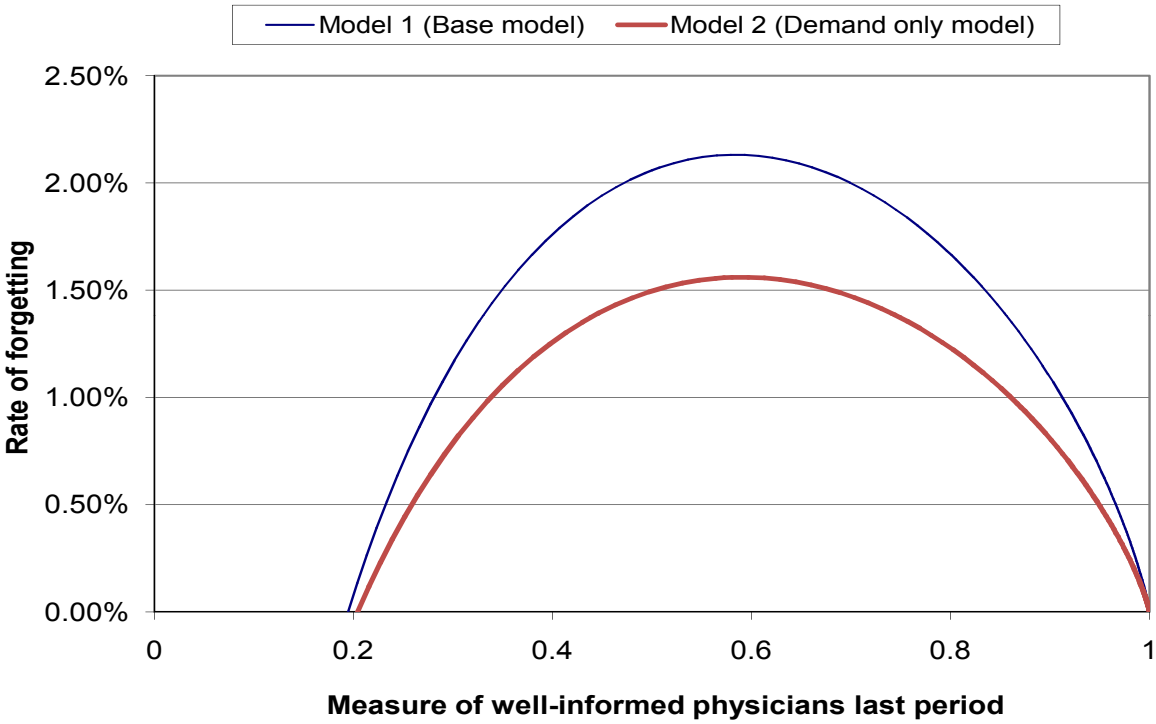

Figure 4: Rate of building the measure of well-informed physicians

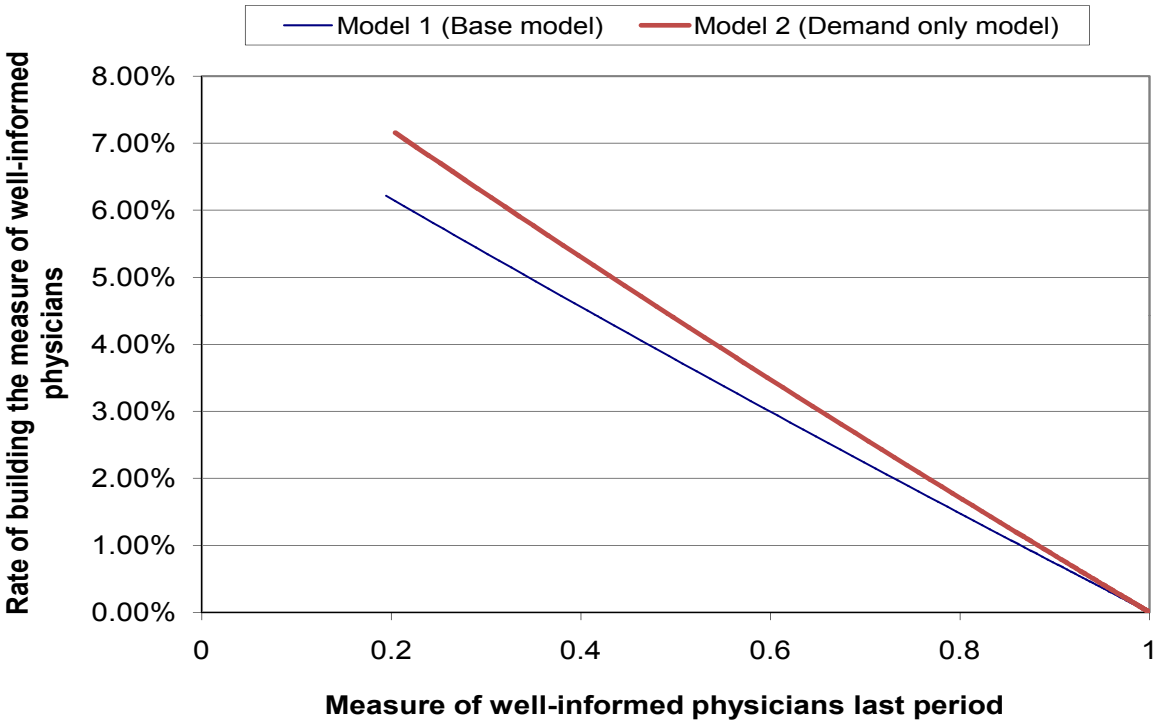

Figure 5: Measure of informed physicians

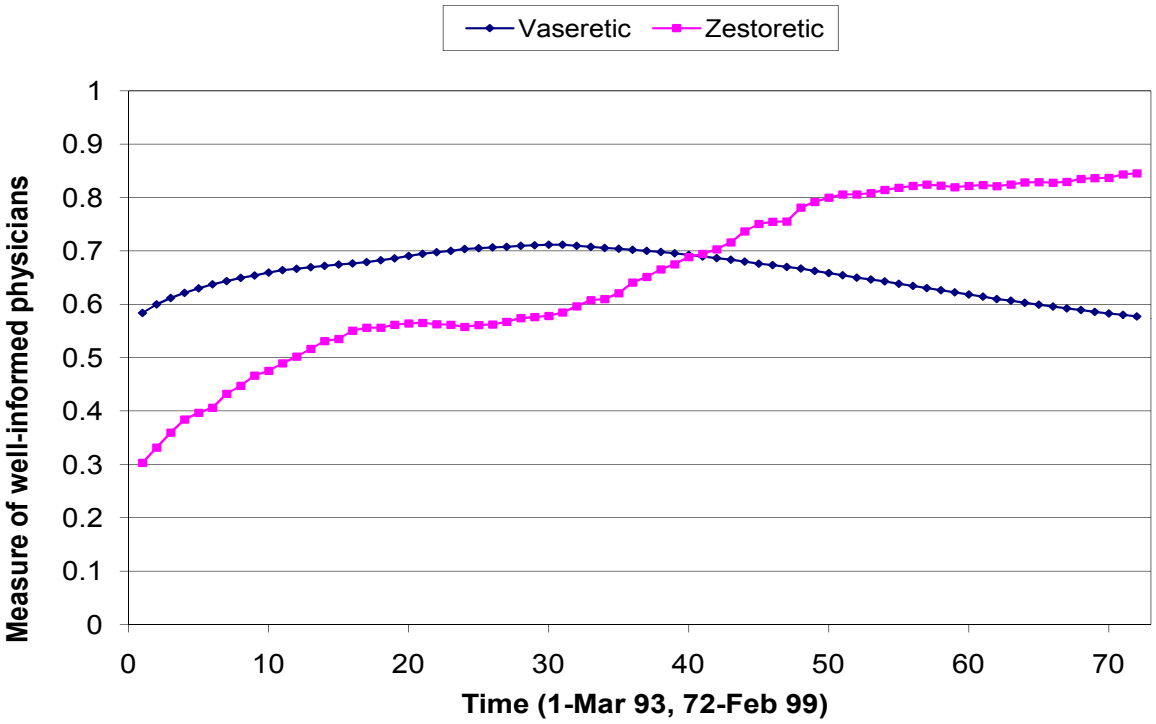

Figure 6: Expected qualities

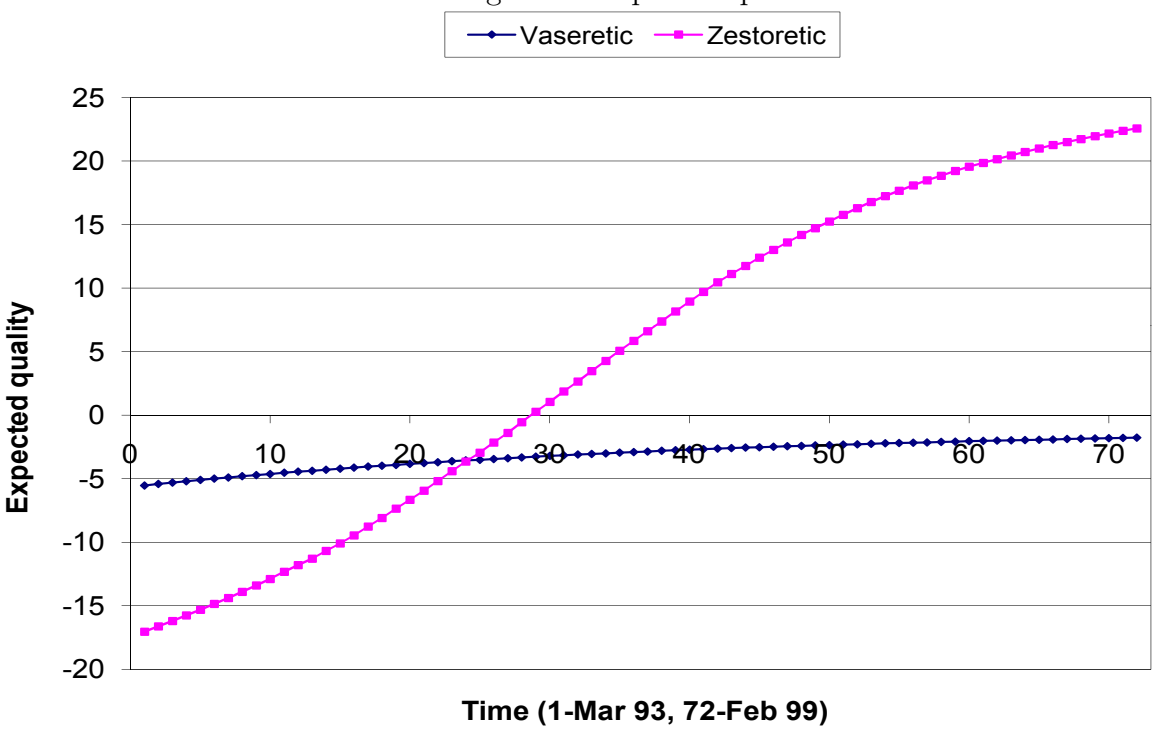

Figure 7: Predicted and Actual Demand for Vaseretic

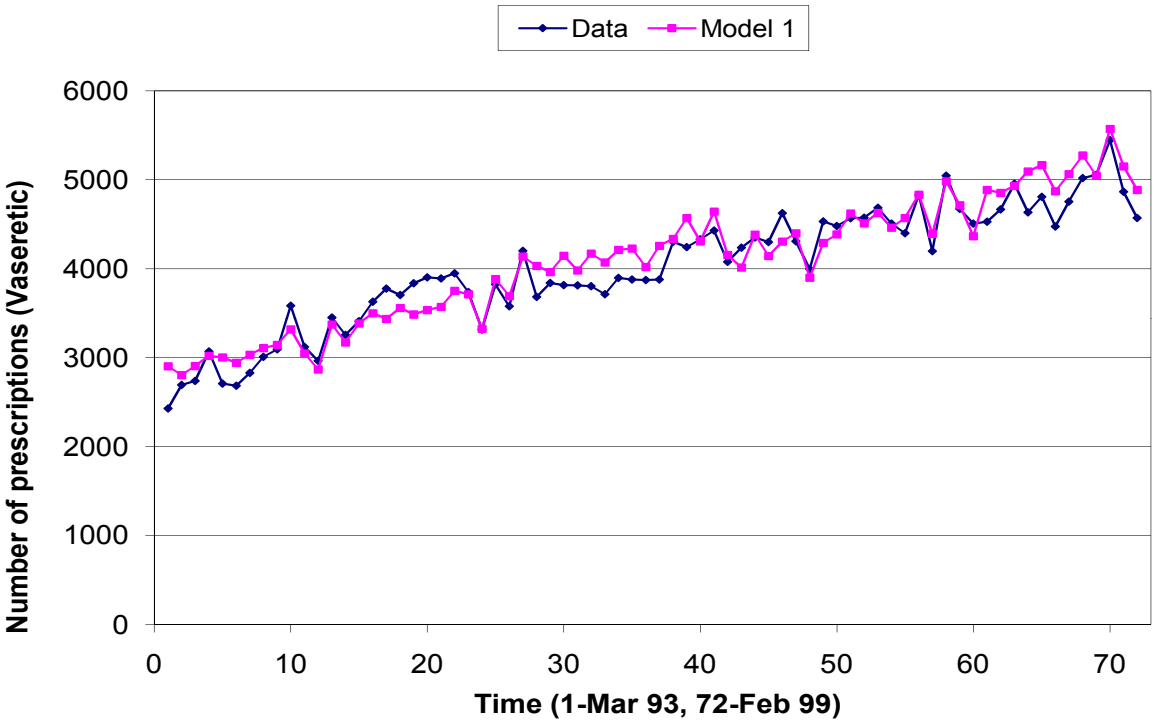

Figure 8: Predicted and Actual Demand for Zestoretic

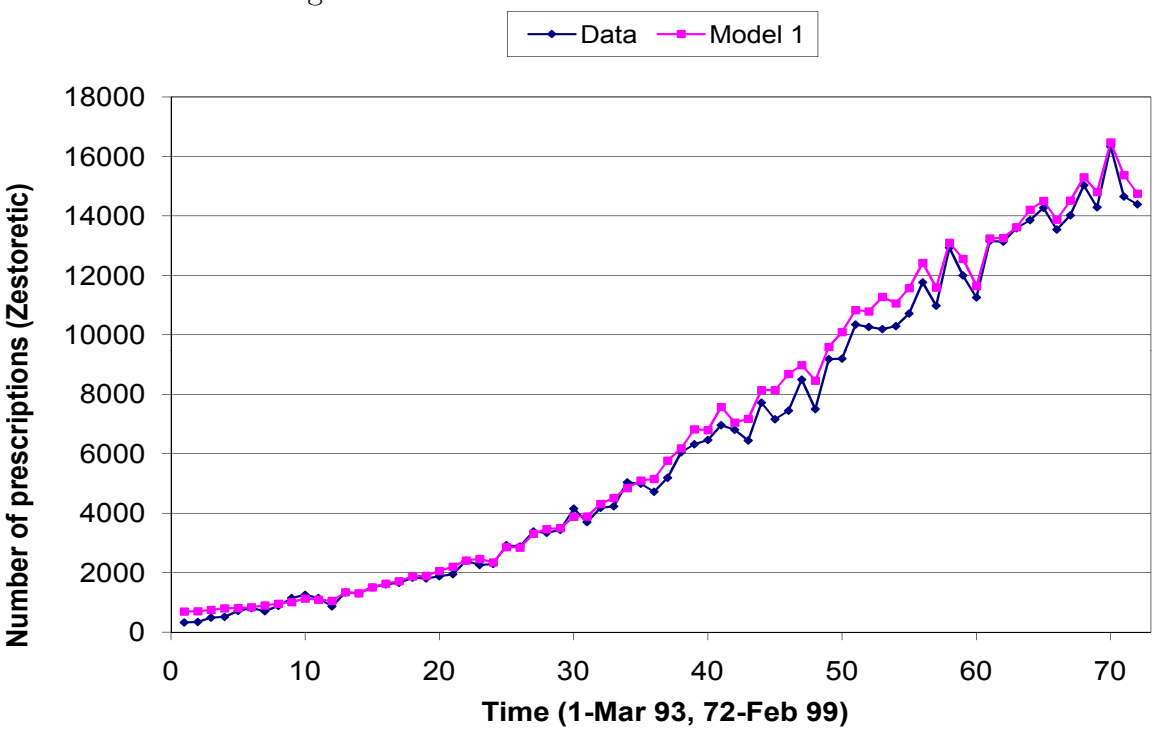

Figure 9: Predicted and Actual Detailing Minutes for Vaseretic

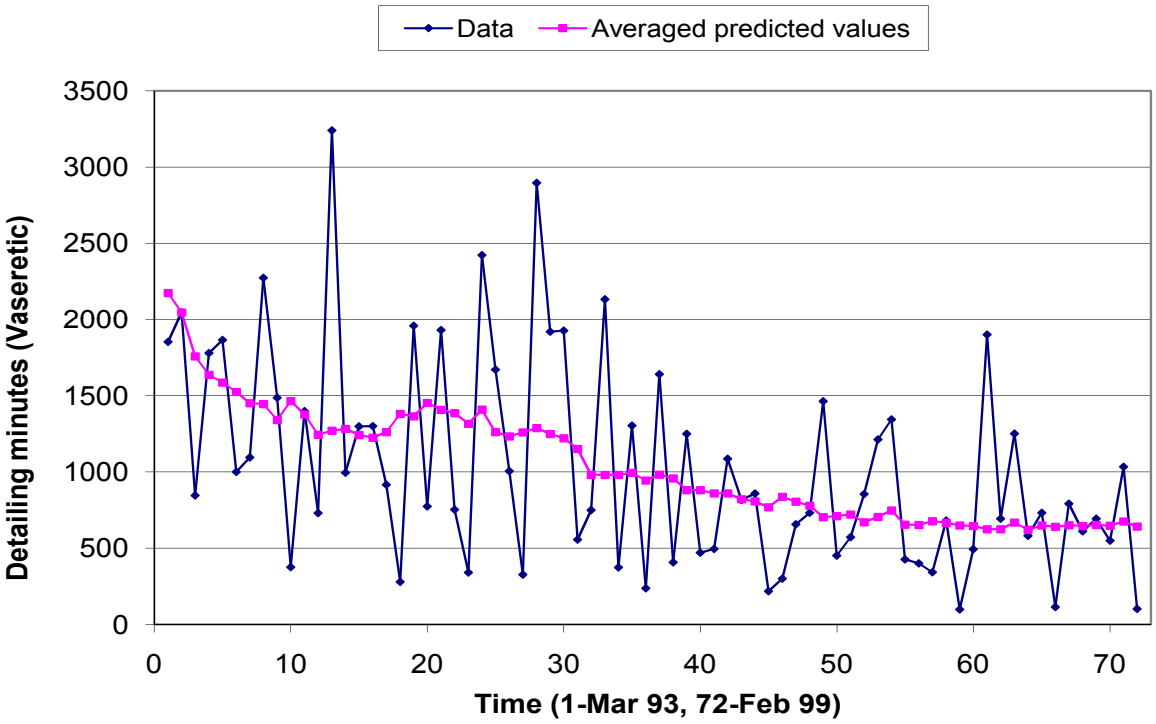
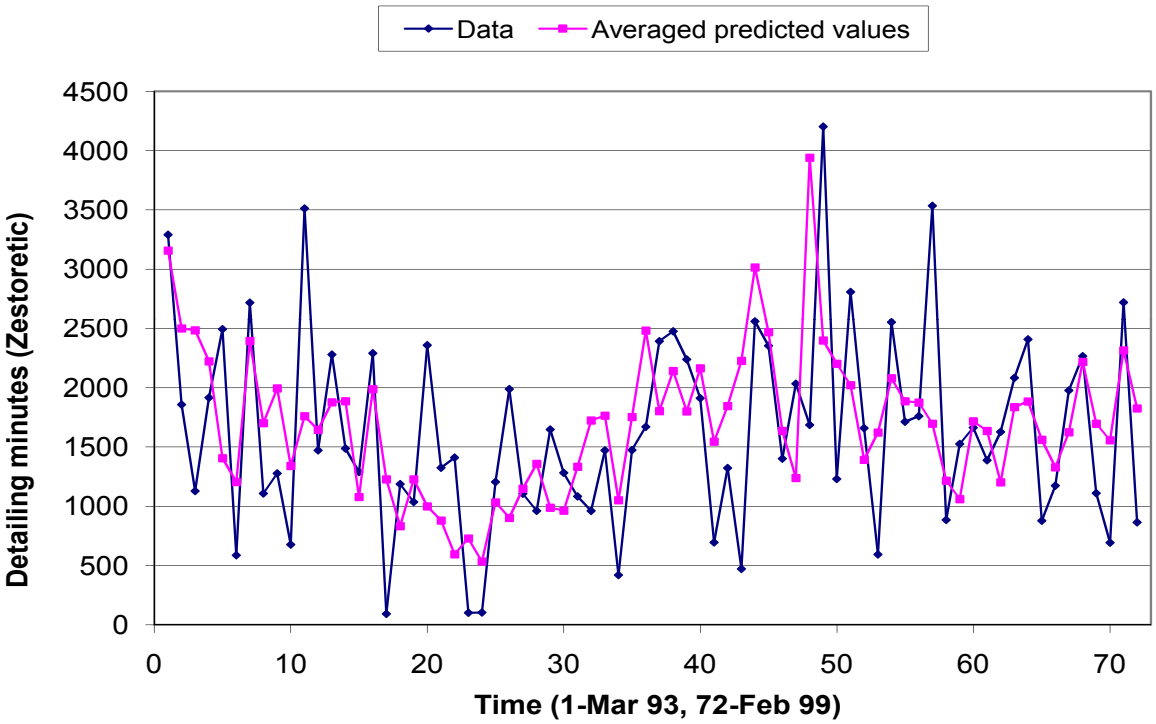
Figure 10: Predicted and Actual Detailing Minutes for Zestoretic