Quant Mark Econ (2010) 8:123–165
DOI 10.1007/s11129-010-9082-z

The effects of detailing on prescribing decisions under quality uncertainty

Andrew Ching · Masakazu Ishihara

Received: 17 April 2008 / Accepted: 21 January 2010 / Published online: 11 February 2010
© Springer Science+Business Media, LLC 2010

Abstract Motivated by recent empirical findings on the relationship between new clinical evidence and the effectiveness of detailing, this paper develops a new structural model of detailing and prescribing decisions under the environment where both manufacturers and physicians are uncertain about drug qualities. Our model assumes (1) a representative opinion leader is responsible for updating the prior belief about the quality of drugs via consumption experiences and clinical trial outcomes, and (2) manufacturers use detailing as a means to build/maintain the measure of physicians who are informed of the current information sets. Unlike previous learning models with informative detailing, our model directly links the effectiveness of detailing to the current information sets and the measures of well-informed physicians. To illustrate the empirical implications of the new model, we estimate our model using a product level panel data on sales volume, prices, detailing minutes, and clinical trial outcomes for ACE-inhibitors with diuretics in Canada. Using our estimates, we demonstrate how the effectiveness of detailing depends on the information sets and the measures of well-informed physicians. Furthermore, we conduct a policy experiment to examine how a public awareness campaign, which encourages physicians/patients to report their drug experiences, would affect managerial incentives to detail. The results demonstrate that the empirical and managerial implications of our model can be very different from those of previous models. We argue that our results point out the

A. Ching (✉) · M. Ishihara
Rotman School of Management, University of Toronto,
105 St. George Street, Toronto, ON, Canada M5S 3E6
e-mail: andrew.ching@rotman.utoronto.ca

M. Ishihara
e-mail: mishihara@rotman.utoronto.ca

TEVA_CAOC_14205460

TE-SF-02127.00001

CCSF v Purdue Pharma, et al.
3:18-CV-7591
TE-SF-02127
Admitted:

importance of developing a structural model that captures the mechanism of how detailing/advertising conveys information in the market under study.

Keywords Detailing · Prescription drugs · Decisions under uncertainty · Representative opinion leader · Diffusion

JEL Classification 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 (Schweitzer 1997, p. 48). This type of personal selling activities allows sales representatives to directly discuss with physicians the 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, in some therapeutic categories, many new indications and side-effects are not revealed until more post-marketing clinical trials are conducted and a large number of patients have tried the drug (Lasser et al. 2002). As a result, the information available today for detailing may be different from the information available tomorrow. This suggests that the current information set, which results from accumulated clinical evidence and past patients' experiences, could directly influence the effectiveness of detailing.

This view of informative detailing is supported by empirical evidence. A recent study by Venkataraman and Stremersch (2007) examines three therapeutic classes: anti-cholesterol drugs (statins), gastrointestinal drugs and erectile dysfunctions drugs, and finds that the impact of detailing on physicians' prescribing behavior depends on the cumulative information on the efficacies and side-effects of drugs. A study by Azoulay (2002) also finds that the detailing efforts for a drug are positively correlated with its cumulative clinical outcomes in the anti-ulcer drug market, suggesting that the marginal return of detailing increases as more favorable information about a drug is accumulated. Although these studies point out this important feature of detailing, none of the existing structural models are able to capture it. Most of the structural modeling papers (e.g., Chan et al. 2007; Narayanan et al. 2005; Mukherji

TEVA_CAOC_14205461

TE-SF-02127.00002

2002) adopt the framework of Erdem and Keane (1996) to investigate the informative effects of detailing on demand. The fundamental assumptions of this framework are: (1) manufacturers have complete information about the quality of their products when they launch them; (2) informative detailing (advertising) simply conveys noisy signals about the true quality of the products to physicians (consumers). It follows that any new information available about the drug (including patients' experiences and clinical trial outcomes) will act as substitutes for the detailing signals, and consequently, will only reduce the effectiveness of informative detailing under their framework. This would be the case even if a clinical trial reveals a very good news about a drug in terms of efficacy.

In light of the shortcomings of the previous literature, the goal of this research is to develop a structural model that is consistent with the empirical findings stated above. In our model, the information sets are updated based on patients consumption experiences and clinical trial outcomes over time, and detailing serves as a means to build/maintain the measure of physicians who are informed of the most updated information. For each drug, physicians are either informed of the most updated information or uninformed. We assume that the measure of physicians who are informed about a particular drug depends on its cumulative detailing efforts. We also assume that the most updated information is maintained by a representative opinion leader. This is to capture the idea that opinion leaders play an important role in disseminating new information about drugs, and are often considered as an important source of the most up-to-date information about the drug categories in which they specialize (e.g., Haug 1997; Thompson 1997). Furthermore, we model physicians' forgetting by allowing the measure of well-informed physicians to depreciate over time.[1] One important implication of our framework is that informative detailing will continue to affect physicians' prescribing decisions even after the uncertainty about drugs' efficacies and side-effects is completely resolved, as long as the depreciation rate for the measure of well-informed physicians is strictly positive. In other words, our way of modeling informative detailing captures the role of reminding physicians of the most updated information about drugs.

This research also contributes to the literature of structural consumer learning models. In addition to the pioneer work by Erdem and Keane (1996), the following 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. Neither Mullainathan (2002) nor Mehta et al. (2004) model the effect of marketing communication mix. Ackerberg (2003) estimates a model in which a consumer infers the quality of the product from the advertising intensity (implicitly through a signaling equilibrium). His model does not allow for consumer forgetting. Moreover,

[1] We provide a formal definition of forgetting in our context in Section 3.2.

TEVA_CAOC_14205462

TE-SF-02127.00003

similar to Erdem and Keane (1996), he assumes manufacturers know the true mean quality of their products. Ching (2000, 2008, 2009) estimates a structural learning model to examine the equilibrium pricing strategies and diffusion pattern empirically in the US prescription drug market after patent expiration. However, since brand-name firms usually cut their detailing efforts dramatically after patent expiration, he does not model detailing.

As far as we know, this is the first paper that develops an empirical structural model to study the effects of detailing on demand, under the environment where both manufacturers and physicians/patients are uncertain about the quality of drugs. This new framework allows us to quantify the marginal impact of detailing as a function of the measure of well-informed physicians and the current information sets, including new clinical evidence. To illustrate the empirical implications of our model, we estimate its structural parameters using a panel data on sales volume, prices, detailing efforts and clinical trials outcomes for ACE-inhibitors with diuretics. We follow the approach proposed by Ching (2000, 2008) to control for the potential endogeneity problem of detailing. We also 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. The results demonstrate that the empirical and managerial implications of our model can be very different from those of previous models. But we emphasize that this does not mean that our model is necessarily better than the previous learning models. Rather, our results point out the importance of developing a structural model of detailing that would capture the mechanism of how detailing/advertising conveys information in the market under study.

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, its empirical implications and identification. Section 4 describes the data and shows empirical evidence that supports our model. It also discusses the estimation strategy, including how we handle the potential endogeneity problem of detailing. In Section 5, we discuss the results of the estimation, policy experiment, and the managerial implications on detailing efforts. In Section 6, we conclude by discussing how a marketing manager could make use of our model to plan allocation/training of sales forces for future detailing activities. We also discuss the limitations of our model and future research directions.

2 Background

Why would the information about drugs' efficacies and side-effects change over time? 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

TEVA_CAOC_14205463

TE-SF-02127.00004

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).

Physicians are supposed to keep themselves updated of the latest information for drugs. However, with many new drugs entering the market each year, it is difficult for general physicians to keep up with the enormous amount of information that changes regularly.[2] Most primary care physicians therefore rely on three external sources of information: (1) peers who are opinion leaders (Haug 1997; Thompson 1997); (2) sales representatives (Schweitzer 1997; Coleman et al. 2004, p.179; Greider 2003, p. 67); and (3) medical journals.

According to the medical continuing education literature, opinion leaders are an important source of information for general physicians (e.g., Haug 1997; Thompson 1997). In Medicine, opinion leaders are physicians who specialize in doing research in a particular field (e.g., cardiovascular). The research focus of their career requires them to be much more updated about the current evidence about the drugs used in the field. In our model, we introduce a representative opinion leader to capture their role.

Among the three external sources listed above, sales representatives are the most time-saving source of information because they visit primary care physicians, compile information on clinical studies for them, and remind them of drug information. Primary care physicians are usually occupied with seeing patients.[3] Therefore, without detailing, it is plausible that they may forget the information about a drug's attributes (e.g., side-effects and efficacy profile) over time, and 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 1980s 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 should be noted that the presentations given by sales representatives may be biased towards the drugs they promote. This possibility appears to

[2]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.

[3]In a survey conducted by Salisbury et al. (1998), they found that more than 70% of general practitioners (GPs) agreed that there was too much information available on prescribing changes to assimilate, and they did not have enough time to keep themselves updated of recent recommendations in drug usage. Interestingly, when the GPs were asked which sources of advice about prescribing they found most useful, they ranked medical advisors (correspond to opinion leaders) and sales representatives the highest.

TEVA_CAOC_14205464

TE-SF-02127.00005

be well-recognized by health care professionals, and physicians are usually cautious when listening to the sales representatives' claims (e.g., Cooper et al. 2003; Ziegler et al. 1995). 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. Most importantly, 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.

Overall, the discussion above suggests that the impact of detailing on demand could depend on the actual effectiveness and side-effects of the drug. As mentioned earlier, the previous empirical literature also finds evidence to support this empirical implication. This motivates us to develop a new structural model of detailing. We now turn to discuss our model of detailing and prescribing decisions in detail.

3 Model

Our framework here extends Ching (2000, 2008). 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 and clinical trials' outcomes 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, D_{jt} is realized. Given D_{jt}, M_{jt} is determined. Each physician then makes prescribing decisions 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. At the same time, the results of some clinical trials may also be realized and published in academic medical journals. The representative opinion leader

TEVA_CAOC_14205465

TE-SF-02127.00006

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.[4]

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

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

where δ_{ijt} is the signal noise. We assume that δ_{ijt} is an *i.i.d.* normally distributed random variable with zero mean and variance σ^2_δ. 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.

In addition to patients' experiences, the representative opinion leader may receive new clinical evidence about drug j if the results of some new clinical trials are available or published in period t. Let $\tilde{q}^c_{ljt}$ be the experience signal of patient's l in a clinical trial published in period t. Then,

$$\tilde{q}^c_{ljt} = q_j + \eta_{ljt}, \tag{2}$$

where η_{ljt} is the signal noise and normally distributed with zero mean and variance σ^2_η. We expect that σ^2_η would be smaller than σ^2_δ because participants in clinical trials are usually monitored much closer than regular patients. As we will discuss later, our estimation results are consistent with this prediction.

The representative opinion leader's initial prior on q_j ($\underline{I}^o_j$) is also normally distributed:

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

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 and the outcomes of clinical trials. The updating is done in a Bayesian fashion.

[4] Obviously, drug qualities are multi-dimensional. Following Ching (2000), 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.

TEVA_CAOC_14205466

TE-SF-02127.00007

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

$$E[q_j|I(t+1)] = E[q_j|I(t)] + \iota_j^e(t) * (\bar{q}_{jt}^e - E[q_j|I(t)]) + I_t^c * \iota_j^c(t) * (\bar{q}_{jt}^c - E[q_j|I(t)]), \quad (4)$$

where $\bar{q}_{jt}^e$ is the sample mean of the consumption experience signals that are revealed in period t; I_t^c is the indicator function for whether a clinical trial is published in period t; and $\bar{q}_{jt}^c$ is the sample mean of the experience signals revealed in a clinical trial in period t. Let n_{jt} be the number of patients who take drug j, and n_{jt}^c be the number of participants who take drug j in the clinical trial. Then $\iota_j^e(t)$ and $\iota_j^c(t)$ can be expressed as:

$$\iota_j^e(t) = \frac{\frac{\kappa n_{jt}}{\sigma_\delta^2}}{\frac{1}{\sigma_j^2(t)} + \frac{\kappa n_{jt}}{\sigma_\delta^2} + \frac{n_{jt}^c}{\sigma_\eta^2}} \quad \text{and} \quad \iota_j^c(t) = \frac{\frac{n_{jt}^c}{\sigma_\eta^2}}{\frac{1}{\sigma_j^2(t)} + \frac{\kappa n_{jt}}{\sigma_\delta^2} + \frac{n_{jt}^c}{\sigma_\eta^2}}, \quad (5)$$

where κ is the proportion of experience signals revealed to the public. ι_j^e and ι_j^c can be interpreted as the weights that the representative opinion leader attaches to the information sources in updating its expectation about the level of q_j. In particular, $\iota_j^e(t)$ and $\iota_j^c(t)$ increase with $\sigma_j^2(t)$.

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

$$\sigma_j^2(t+1) = \frac{1}{\frac{1}{(\underline{\sigma}_j^o)^2} + \frac{\kappa N_{jt}}{\sigma_\delta^2} + \frac{N_{jt}^c}{\sigma_\eta^2}} = \frac{1}{\frac{1}{\sigma_j^2(t)} + \frac{\kappa n_{jt}}{\sigma_\delta^2} + \frac{n_{jt}^c}{\sigma_\eta^2}}, \quad (6)$$

where $N_{jt}(= \sum_{\tau=1}^t n_{j\tau})$ is the cumulative consumption of drug j; and $N_{jt}^c(= \sum_{\tau=1}^t n_{j\tau}^c)$ is the cumulative number of patients who took drug j in published clinical trials up to time t.

3.2 Detailing and measure of well-informed physicians

We now turn to discuss the physicians' choice problem and how detailing influences their choices. 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., $M_{jt} =$

TEVA_CAOC_14205467

TE-SF-02127.00008

$f(M_{jt-1}, D_{jt})$. We assume that $f(M_{jt-1}, .)$ is monotonically increasing in D_{jt}. To capture the possibility that physicians may forget about the exact outcomes of past clinical trials, which serve as the basis for assessing the effectiveness of drugs, 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 at time t. Such a modification will make the model computationally infeasible to estimate using product level data.[5] 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.

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

$$G_{jt} = (1 - \phi_G)G_{jt-1} + D_{jt}, \tag{7}$$

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

$$M_{jt} = \frac{exp(\beta_0 + \beta_1 G_{jt})}{1 + exp(\beta_0 + \beta_1 G_{jt})}. \tag{8}$$

Define the *average* rate of forgetting, $\phi_M \equiv (M - f(M, 0))/M$. Although ϕ_G is a constant, G_{jt} affects M_{jt} nonlinearly. In particular, the implied average forgetting rate, ϕ_M, will exhibit an inverted-U shape. This might first appear to be restrictive, but it is consistent with the following intuition. It is likely that

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

TEVA_CAOC_14205468

TE-SF-02127.00009

individual 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. Physicians with relatively higher rate of forgetting would probably rely more on sales representatives to 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. The arguments above suggest 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 potential implication of this inverted-U shape forgetting rate is that when manufacturers just launch a new drug, they may have an incentive to use a large amount of detailing to quickly build up a large stock of well-informed physicians. Because the forgetting rate is low for a large stock of well-informed physicians, manufacturers would be able to spend less detailing efforts to maintain its size afterwards. This prediction is consistent with the general detailing patterns that we observe in the industry during a drug's product lifecycle. It is typical that the detailing efforts are very high at the beginning of the product lifecycle, and then quickly decline. The detailing efforts may increase or decrease over time, but in general will be maintained at a much lower level than the introductory stage (e.g., Berndt et al. 1997; Narayanan et al. 2005).[6]

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:

$$E\left[\sum_{j\in\{0,1,\ldots,J\}} u_{ijt}\cdot d_{hij}(t)|I^h(t)\right], \tag{9}$$

where u_{ijt} is patient i's utility from consuming drug j at time t; $d_{hij}(t)=1$ 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

[6] We thank an anonymous referee who suggests to us to explore this prediction.

TEVA_CAOC_14205469

TE-SF-02127.00010

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:

$$u_{ijt} = \alpha - \exp(-r\tilde{q}_{ijt}) - \pi_p p_{jt} + \zeta_{ikt} + e_{ijt}, \tag{10}$$

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).[7] 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, 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:

$$\begin{aligned} E[u_{ijt}|I_h(t)] &= E[u_{ijt}|I_j(t)] \\ &= \alpha - \exp\left(-rE[q_j|I(t)] + \frac{1}{2}r^2(\sigma_j^2(t) + \sigma_\delta^2)\right) - \pi_p p_{jt} + \zeta_{ikt} + e_{ijt}. \end{aligned} \tag{11}$$

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

$$\begin{aligned} E[u_{ijt}|I_h(t)] &= E[u_{ijt}|\underline{I}_j^p] \\ &= \alpha - \exp\left(-r\underline{q}_j^p + \frac{1}{2}r^2(\underline{\sigma}_j^{p2} + \sigma_\delta^2)\right) - \pi_p p_{jt} + \zeta_{ikt} + e_{ijt}. \end{aligned} \tag{12}$$

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.

[7] 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.

TEVA_CAOC_14205470

TE-SF-02127.00011

Equations 10–12 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:

$$E[u_{i0t}|I_h(t)] = \alpha_0 + \pi_t t + \zeta_{i0t} + e_{i0t}. \tag{13}$$

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,

$$n_{jt} = \text{Size}_t \cdot S(j|D_t, (E[q_j|I(t)], \sigma_j(t), M_{jt-1})_{j=1}^{J}; \theta_d) + \epsilon_{jt}, \tag{14}$$

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 and identification

3.4.1 Differences between our model and previous models

To illustrate some empirical implications of our model for the effectiveness of detailing, 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,

$$\begin{aligned} S_{jt} = {} & M_{jt}M_{kt}s_{jt}(I_j(t), I_k(t)) + M_{jt}(1 - M_{kt})s_{jt}(I_j(t), \underline{I}_k^p) \\ & + (1 - M_{jt})M_{kt}s_{jt}(\underline{I}_j^p, I_k(t)) + (1 - M_{jt})(1 - M_{kt})s_{jt}(\underline{I}_j^p, \underline{I}_k^p), \end{aligned} \tag{15}$$

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

$$\frac{\partial S_{jt}}{\partial D_{jt}} = \frac{\partial M_{jt}}{\partial D_{jt}} \times \{M_{kt}\Delta s_{jt}(I_k(t)) + (1 - M_{kt})\Delta s_{jt}(\underline{I}_k^p)\}, \tag{16}$$

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 his/her information set for drug k being I_k. Equation 16 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). Moreover, since $\partial S_{jt}/\partial D_{jt}$ depends on $\partial M_{jt}/\partial D_{jt}$, the marginal return of detailing also depends on the measure of well-informed physicians (M_{jt-1}), its depreciation rate or forgetting rate (ϕ_G), and $f(M_{jt-1}, .)$.

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

TEVA_CAOC_14205471

TE-SF-02127.00012

true quality of the incumbent, i.e., $I_k(t) \rightarrow I_k(\infty)$). Conditional on M, Eqs. 15 and 16 imply that the entrant's marginal return of detailing will increase with its market share. Moreover, the marginal impact of detailing 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 $\phi_G > 0$. Moreover, its impact on demand depends on $I(t)$, $\underline{I}^p$ and M_{jt-1} (i.e., G_{jt-1}). On the contrary, previous models of learning and informative detailing/advertising, which follow the framework of Erdem and Keane (1996), imply that the marginal impact of informative detailing/advertising *diminishes* over time as the uncertainty about product quality is slowly resolved.[8] This demonstrates that the empirical implications of our model are quite different from those of the previous models.

We next turn to discuss how the outcomes of new clinical evidence would affect the marginal return of detailing. Suppose that in period t, the result of a clinical trial shows that drug j is able to treat a new problem. As a result, it improves $I_j(t)$. It follows from Eq. 16 that both $\Delta s_{jt}(I_k(t))$ and $\Delta s_{jt}(\underline{I}_k^p)$ will increase. This in turn raises $\frac{\partial S_{jt}}{\partial D_{jt}}$. Similarly, if a clinical trial reveals bad news about drug j, this would reduce $\frac{\partial S_{jt}}{\partial D_{jt}}$ in our framework. On the contrary, in the previous learning models, information signals from clinical trials are substitute for consumption experience and detailing signals. Therefore, $\frac{\partial S_{jt}}{\partial D_{jt}}$ will decrease as more outcomes of clinical trials come out, regardless of whether they are good or bad news during the product lifecycle. As we mentioned earlier, this is inconsistent with the previously documented evidence in a few therapeutic categories (Azoulay 2002; Venkataraman and Stremersch 2007).

Finally, we discuss how the long-run implication of the effect of detailing in our model differs from that in the previous models with persuasive effects (detailing goodwill stock in the utility function). By incorporating a goodwill stock of detailing into the utility function, the previous learning models can generate long-run effects of detailing after the uncertainty about the drug quality has been completely resolved. Previous research argued that this approach to model the long-run effects of detailing allows researchers to capture its reminding role (e.g., Narayanan et al. 2005, p. 278). However, if detailing is to remind physicians about the true quality of drugs (e.g., clinical evidence and side-effects), its long-run effectiveness should also depend on the true mean qualities of the drugs. Nevertheless, such an intuitive relationship

[8] Anand and Shachar (2005) introduced a framework of informative advertising, which could also lead to marginal return of detailing either increasing or decreasing over time when combined with persuasive effects. The logic is that consumers are uncertain about whether they match well with some unobserved product attributes. Suppose that the consumers do not match well with the product. The marginal return due to informative advertising would then be negative, and diminishing over time. Now suppose that the marginal return due to persuasive advertising is constant over time. Then the total marginal return (informative + persuasive) would be increasing over time and converges to the marginal return due to persuasive advertising alone.

TEVA_CAOC_14205472

TE-SF-02127.00013

cannot be explained by the previous approach, while our model captures it in a natural way.[9]

Our model also has an empirical implication on firm's detailing efforts over time. If the information reveals that a drug turns out to be superior (inferior) than another drug in terms of efficacy or side-effects profile, the manufacturer that markets the superior (inferior) drug would have a stronger (weaker) incentive to do detailing under our framework. As we will show in the next section, the data for ACE-inhibitors with diuretics is consistent with this implication.

3.4.2 Identification

In this subsection, we provide some intuitions about how the parameters of our model are identified. To identify the parameters that determine the measure of well-informed physicians (β_0, β_1, ϕ_G), the best source should be the data variation during the later stage of the product lifecycle when the public information sets for both drugs converge. This is because in the long run, the true drug qualities are revealed after accumulating sufficient evidence and experiences. As a result, the variation of market shares should be mainly driven by the variation of the measures of well-informed physicians, which in turn depends on detailing efforts.[10] The long-run steady state market shares should help identify the true mean qualities. The diffusion paths depend on both the evolution of the public information sets and the measures of well-informed physicians. After controlling the measure of well-informed physicians (relying on the long run variation in the data), the diffusion path, which is closely tied to the rate of learning, should help identify the learning parameters. In particular, its non-linear nature should help identify the initial prior mean qualities and variance, experience signal noise variance and risk aversion parameter. The timing of the release of clinical trials and the corresponding change in market shares should identify the signal noise variance of clinical trials. The fluctuations of initial market shares should also help identify the initial prior mean qualities and variance.

To identify the difference between the expected utilities due to the initial prior of the representative opinion leader and the physicians, the best source should be the beginning stage of the product lifecycle when there are very little patients' experiences revealed to the public. Intuitively, the magnitude of the impact of detailing on the demand should increase with their difference. In other words, if the initial impact is very small, this would tell us that the initial

[9] Alternatively, to capture the intuitive relationship that we discuss here, one could allow a detailing goodwill stock interacting with the expect quality to enter the utility function. Such a modification can be viewed as a reduced form of our structural model.

[10] But even before we reach the long run, the fluctuation of market shares and detailing efforts (current and past) would also help identify (β_0, β_1, ϕ_G), as we point out in the previous subsection about the relationship between $\frac{\partial S_{jt}}{\partial D_{jt}}$ and ($\frac{\partial M_{jt}}{\partial D_{jt}}$, M_{kt}).

TEVA_CAOC_14205473

TE-SF-02127.00014

expected utility between informed and uninformed physicians are very close to each other. If the initial impact is positive and large, this would indicate that the initial prior for the representative opinion leader is "better" than that for physicians. If the impact is negative, this would suggest that physicians may have a more optimistic prior than the representative opinion leader.[11]

In simulated maximum likelihood, parameter values will be chosen to fit the growth rate of market shares, and impacts of detailing over time. Intuitively, one can imagine that we start with an initial guess of $(\underline{I}^o, \underline{I}^p, \sigma_\eta^2, \sigma_\delta^2, q_j)$, simulate a sequence of $\{E[q_j|I(t)], \sigma_j^2(t)\}_{t=0}^T$, treat them as the data, and estimate the rest of parameters $(r, \pi_p, \beta_0, \beta_1, \phi_G)$. Then we fix the interim estimate of $(r, \pi_p, \beta_0, \beta_1, \phi_G)$, generate $\{M_{jt}\}_{t=1}^T$, and estimate $(\underline{I}^o, \underline{I}^p, \sigma_\eta^2, \sigma_\delta^2, q_j)$. The estimation procedure essentially goes through this reiterative process many times until it finds the set of parameter vector that generate model predictions that fit the data best.

With the CARA utility function and the logit relationship between M and (G, D), we can identify all the parameters of the model if we normalize $\underline{q}_j^p$ (or $\underline{\sigma}_j^p$) $\forall j$, and one of the q_j's. In Appendix A, we provide a more formal discussion about the identification.

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 and heart failure. 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 the therapy is under way.

We choose Canada and ACE-inhibitor with diuretic for five reasons. First, most of the patients who have high blood pressure are elderly, and their prescription drugs are covered by the Canadian government. This suggests that prices may not play an important role in determining demand. Second, 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 (Elgie 2001). In order to change the prices, a brand-name firm has to submit an application to the Patented Medicine Price Review Board, and they may or may not approve it. As a result, the

[11] We believe that the latter case is very unlikely. To our knowledge, most of the existing studies find positive correlation between detailing and demand. Moreover, it is likely that opinion leaders or informed physicians are more familiar with the clinical studies that allow the manufacturer to gain the approval of selling a new drug.

TEVA_CAOC_14205474

TE-SF-02127.00015

pharmaceutical prices usually change quite infrequently in Canada. These institutional details allow us to treat prices as exogenous and focus on modeling the effects of detailing. Third, the market of ACE-inhibitor with diuretic does not have direct-to-consumer (DTC) advertising. DTC advertising has increased dramatically in the USA since 1997. It is believed that it plays a significant role in the demand for prescription drugs. However, the way that DTC advertising influences physicians' choice is likely different from detailing. Modeling the effects of DTC advertising is beyond the scope of this paper. Fourth, 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. Fifth, ACE-inhibitors is also a class of drugs about which medical advisory boards had consistently recommended changes in their prescribing guidelines to physicians to reflect the findings from clinical trials in the early 1990s (Salisbury et al. 1998), when ACE-inhibitors with diuretics were introduced. This suggests that uncertainty and learning about the quality of this class of combination drugs is likely important for both manufacturers and physicians.

Sales and detailing data for this study come from IMS Canada, a firm that specializes in collecting sales and detailing 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 sales data, we plot the number of prescriptions filled for Vaseretic and Zestoretic in Fig. 1. The sales of both drugs increase over time. The monthly sales of Vaseretic grow slowly and steadily from 2,500 to 4,500 prescriptions, while Zestoretic's monthly sales grow at a much faster rate from around 300 to more than 14,000 prescriptions. Being the incumbent of

TEVA_CAOC_14205475

TE-SF-02127.00016

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

the ACE-inhibitor with diuretic, the sales of Vaseretic is about eight times as many as that of Zestoretic at the beginning of the sample period (March 1993). It took Zestoretic more than 2 years to overtake Vaseretic's sales. By the end of the sample period (February 1999), the sales of Zestoretic is more than three times as many as 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, ACE-inhibitor with diuretic, and Thiazide Diuretic. It increases from 655,000 to 860,000 during the sample period.

We also plot detailing minutes in Fig. 2. The average detailing minutes of Zestoretic are about the same as those of Vaseretic before $t = 30$. But after $t = 30$, about the time when Zestoretic overtakes Vaseretic, the average detailing minutes of Zestoretic become higher than Vaseretic. It should also be noted that there is significant fluctuation in detailing minutes, which should

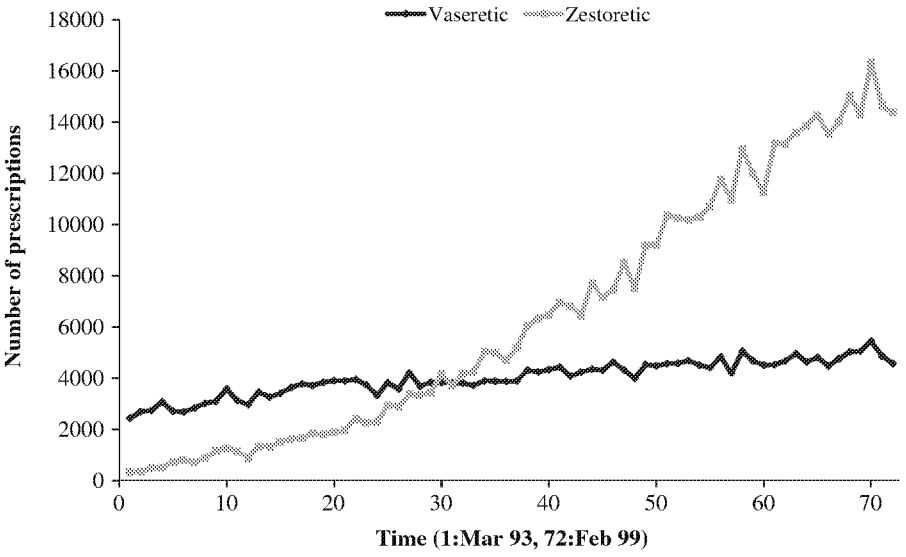

Fig. 1 Total sales vs time

TEVA_CAOC_14205476

TE-SF-02127.00017

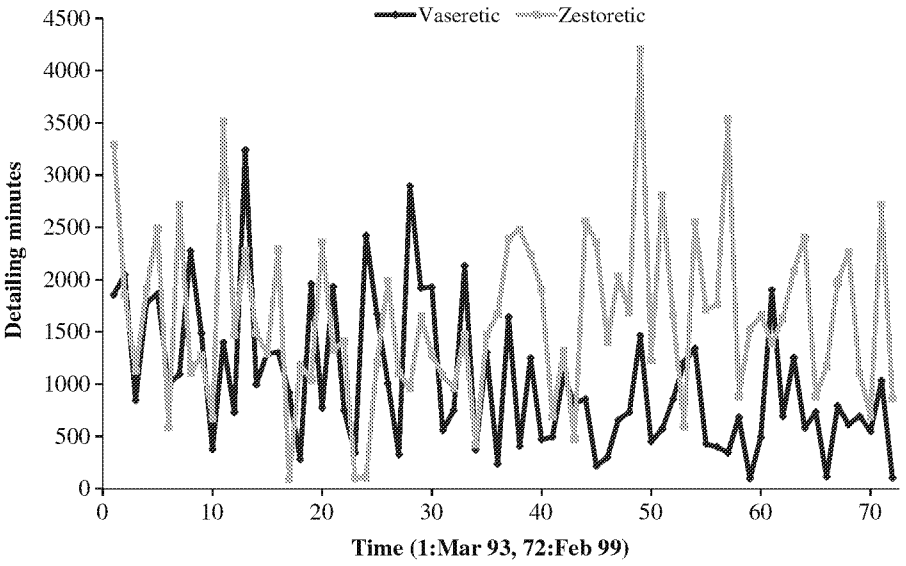

Fig. 2 Detail minutes vs. time

help identify the parameters that determine the measure of well-informed physicians (i.e., β_0, β_1, and ϕ_G).

In addition, we collected data on clinical trials that compare the efficacy of Vaseretic and Zestoretic from medical journal articles archived in PubMed.[12] We focus on the clinical trials that involve *direct comparison* between Vaseretic and Zestoretic because they should be of first order importance in affecting physician's choice between these two drugs. We collect the clinical trials data from September 1990 to February 1999. We started in September 1990 because this is the inception date of the incumbent drug. For each clinical trial, the *relative* outcome could either be (1) positive for Vaseretic (i.e., negative for Zestoretic), (2) positive for Zestoretic (i.e., negative for Vaseretic), (3) no difference between Vaseretic and Zestoretic. In addition, we also obtain the number of participants in each clinical trial. It should be emphasized that even when a direct comparison clinical trial is "neutral" or "negative" about a drug, the updating process could still revise the beliefs about its true mean quality upwards. In fact, each comparison trial contains two quality signals, one for each drug. A "neutral" outcome simply means that these two signals are the same (i.e., the trial finds that these two drugs are equally effective), and a "positive" or "negative" outcome simply means that the signal for one drug is better than another's. Even though a comparison trial finds that one drug

[12] PubMed (www.pubmed.gov) is a service of the U.S. National Library of Medicine (NLM) that includes over 18 million citations from MEDLINE, the NLM's premier bibliographic database for the life sciences, and other life science journals for biomedical articles back to 1948.

TEVA_CAOC_14205477

TE-SF-02127.00018

Table 2 Comparison clinical trials published between September 1990 and February 1999

Chemical	No. of clinical trials	No. of clinical trails with positive outcome	No. of patients in a clinical trial			
			Mean	S.D.	Max	Min
Enalapril	22	1	53.09	79.19	321	3
Lisinopril	22	8	86.73	168.74	620	3

is inferior, its reported results might still be better than the prior belief about its mean quality. As a result, the signal for an inferior drug could still lead to an upward revise in the expected quality. In Appendix B, we explain how to update the prior of the public information set, based on the outcomes of direct comparison clinical trials in detail.

Table 2 shows the summary statistics for the clinical trial data. As shown, the number of clinical trials with positive outcomes for Zestoretic is larger than that for Vaseretic. If we limit our attention to our sample period, there are four clinical trials that found evidence favorable for Zestoretic whereas there is only one clinical trial that is favorable for Vaseretic. Figure 3 shows the cumulative number of clinical trials since the inception date of Vaseretic. For each clinical trial, we code its outcome as +1 (positive), −1 (negative), and 0 (no difference), and show how the cumulative outcome of clinical trials for Vaseretic evolves over time in Fig. 4 (note that the cumulative outcome of clinical trials for Zestoretic is just the mirror image of Vaseretic). The cumulative outcome of clinical trials for Vaseretic generally decreases over time. Interestingly, this is consistent with: (1) Zestoretic dominates the market

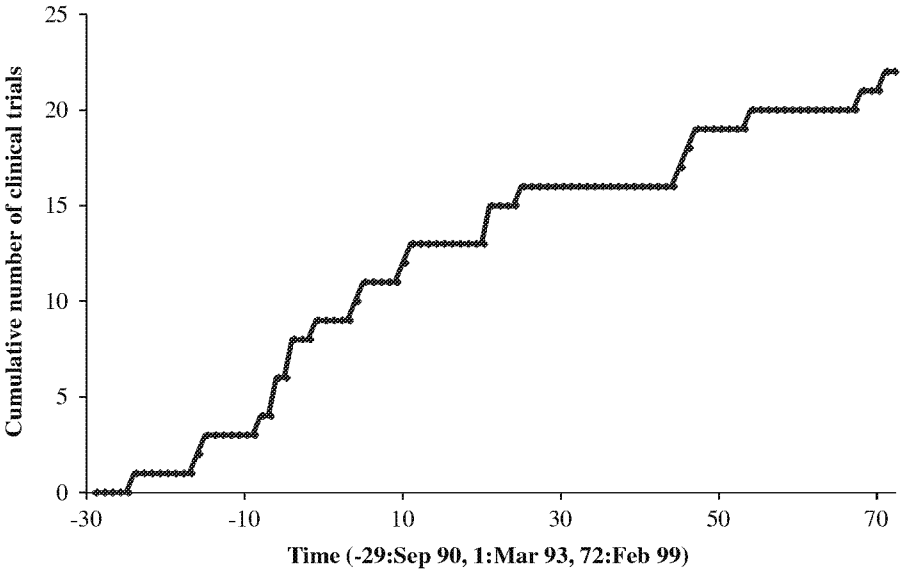

Fig. 3 Cumulative number of clinical trials

TEVA_CAOC_14205478

TE-SF-02127.00019

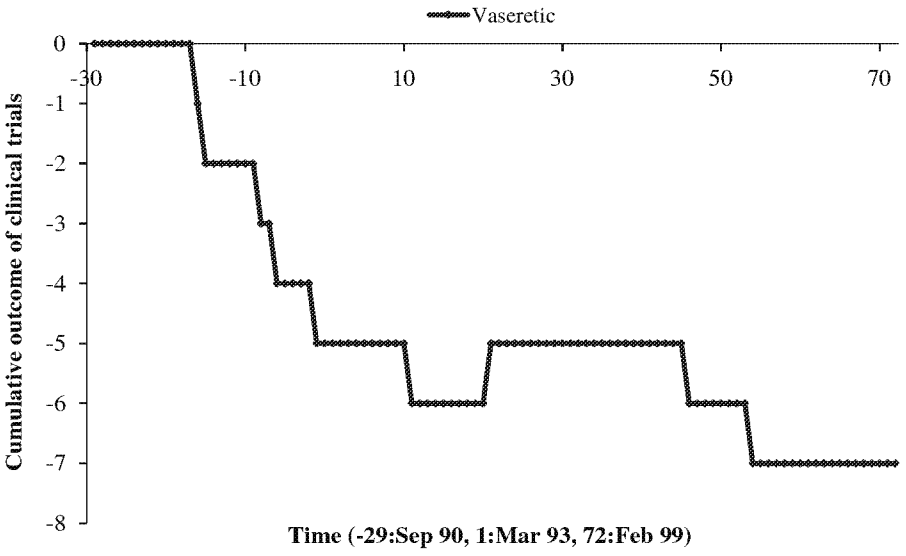

Fig. 4 Cumulative outcomes of clinical trials

later on (see Fig. 1); (2) Vaseretic decreases its detailing efforts in the later part of the sample period (see Fig. 2). In particular, according to our model, we expect that the marginal return of detailing for Vaseretic should decrease over time as more information reveals that it is inferior to Zestoretic. This could potentially explain observation (2).

To provide further evidence that the effectiveness of detailing could be affected by the clinical trial outcomes, we regress the number of prescriptions on the interaction between the cumulative clinical outcomes and detailing (or cumulative detailing), controlling for other factors, and assuming the coefficients are the same across both drugs. Table 3 reports the results for five specifications. We find that the interaction terms are positive and statistically significant across all regressions. The results support our hypothesis that the effectiveness of detailing depends on the current information set of a drug. This also confirms our prior belief that our model should be applicable to ACE-inhibitors with diuretics.

4.2 Simultaneity problem

As we argued above, although we are willing to assume price is exogenous, we feel that detailing could be potentially endogenous. It is plausible that manufacturers observe $I(t)$ before detailing takes place in each period. Consequently, detailing could be a function of $I(t)$, and D_{jt} might 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

TEVA_CAOC_14205479

TE-SF-02127.00020

Table 3 OLS regression of the number of prescriptions on the interaction between the cumulative clinical outcomes and detailing (or cumulative detailing)

DV: Number of prescriptions$_{jt}$	Specification				
variable	(1)	(2)	(3)	(4)	(5)
Det_{jt}	**−0.701**	**−0.535**		−0.06	
	(0.326)	(0.205)		(0.483)	
Cum_Det_{jt}		**0.468**	**0.306**	−0.615	**−0.424**
		(0.032)	(0.019)	(0.343)	(0.199)
$Price_{jt}$	**−61.157**	**−95.304**	**−37.922**	**−72.720**	**−38.564**
	(26.900)	(17.040)	(9.298)	(27.458)	(23.832)
$Cum\_Clinical_{jt}$	418.499	**−907.713**	**−6,652.769**	905.323	**−2,272.400**
	(401.410)	(267.645)	(318.394)	(482.032)	(597.321)
$Det_{jt} \times Cum\_Clinical_{jt}$	**1.142**	**0.789**		**1.048**	
	(0.280)	(0.177)		(0.282)	
$Cum\_Det_{jt} \times Cum\_Clinical_{jt}$			**0.247**		**1.029**
			(0.012)		(0.125)
Constant	**7,999.760**	**−3,705.890**	**−2,955.975**	**10,356.690**	**7,273.512**
	(1,146.685)	(1,074.798)	(536.504)	(1,738.670)	(1,446.416)
Adjusted R-squared	0.395	0.762	0.935	0.404	0.562
No. of observations	144	144	144	144	144

Standard errors are in parentheses; estimates shown in bold are significant at 5% level. In Specification (2) and (3), we follow Berndt et al. (1997) and set the depreciation rate of the cumulative detailing stock at 4.2%. In Specification (4) and (5), we follow Narayanan et al. (2005) and set the depreciation rate at 30%.

Definition of variables: Det_{jt} detailing minutes for drug j at time t, Cum_Det_{jt} cumulative detailing minutes for drug j at time t, $Price_{jt}$ price of drug j at time t, $Cum\_Clinical_{jt}$ cumulative outcomes of direct comparison clinical trials for drug j at time t

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 of estimating demand models using product level data is the GMM approach developed by Berry et al. (1995) (BLP). However, in our model, the unobserved product characteristic (i.e., $E[q_j|I(t)]$) differs across physician types. Consequently, the BLP estimation approach cannot be applied. We therefore estimate our model using the approach developed by Ching (2000, 2008), who used it to control for the price endogeneity problem in a model that study the demand for brand-name drugs and their generic counterparts. To take the endogeneity of detailing into account, Ching's method requires us to approximate manufacturers' detailing 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. Similar to BLP, this approach does not require making any strong assumptions about the equilibrium solution, and whether drug manufacturers maximize their total discounted profits (i.e., forward-looking) or current profits (i.e., myopic). So we can avoid some risks of misspecifying the supply side, which may result in biased estimates. Moreover, it allows us to avoid the computational burden of solving a dynamic oligopoly model when estimating the demand model. Nonetheless, there are two drawbacks in

TEVA_CAOC_14205480

TE-SF-02127.00021

Ching's approach: (1) It increases the number of parameters to estimate due to the pseudo-detailing policy functions; (2) The estimates are not as efficient as full-information maximum likelihood because the supply side model is not explicitly modeled in the estimation. But we should note that even though the computational power has been improved rapidly these days, currently it is still too computationally burdensome to incorporate the solution of a dynamic oligopoly model in the estimation procedure.

Regardless of whether manufacturers are forward-looking or myopic, 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. In addition, it may also depend on factors that we do not explicitly model. For instance, a manufacturer specific shock could also affect the detailing amount of any drugs produced by that manufacturer. To approximate this manufacturer specific shock, we use the total detailing minutes for the set of drugs in the cardiovascular category, which are produced by the manufacturer of the focal drug, but not explicit substitutes for ACE-inhibitors with diuretics.[13] We denote this variable by F_{jt}. By construction, F_{jt} is correlated with D_{jt} through the manufacturer specific shock, but uncorrelated with the demand shocks for drugs that belong to the potential market size for ACE-inhibitor with diuretics. Essentially, F_{jt} serves as an instrumental variable for D_{jt}. Note that Berndt et al. (1997) use a similar variable as the instrument for detailing in their reduced form model.

When specifying the pseudo-detailing policy function, ideally one would use a flexible high order polynomial to do the approximation if the sample is large. In practice, however, one usually needs 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 $j \neq k$,

$$\begin{aligned} log(D_{jt}) = {} & \lambda_{j0} + (\lambda_{j1} + \lambda_{j2} * M_{kt-1}) * (1 - M_{jt-1}) * |\Delta u^q_{jkt}| * \mathbb{I}(\Delta u^q_{jkt} > 0) \\ & + (\lambda_{j3} + \lambda_{j4} * M_{kt-1}) * M_{jt-1} * |\Delta u^q_{jkt}| * \mathbb{I}(\Delta u^q_{jkt} < 0) \\ & + \lambda_{j5} * F_{jt} + \nu_{jt}, \end{aligned} \tag{17}$$

where

$$\Delta u^q_{jkt} = E[u^q_{jt}|I(t)] - E[u^q_{kt}|I(t)], \tag{18}$$

$$E[u^q_{jt}|I(t)] = -\exp\left(-rE[q_j|I(t)] + \frac{1}{2}r^2(\sigma_j^2(t) + \sigma_\delta^2)\right), \tag{19}$$

[13] This set of drugs includes Alpha blockers, Antiarrhythmic agents, Anticoagulants, Antiplatelets, Thrombolytics, Beta blockers, Calcium channel blockers, Centrally acting drugs, Cholesterol-lowering agents, Digitalis drugs, Direct Vasodilators, Nitrates, Peripheral adrenergic antagonists, AII Receptor Antagonists, and other combination drugs. In our estimation, we treat ACE-inhibitor and Thiazide Diuretic as explicit substitutes for ACE-inhibitor with diuretics and use the total sales of these three sub-categories of drugs to construct the potential market size for ACE-inhibitor with diuretics.

TEVA_CAOC_14205481

TE-SF-02127.00022

v_{jt} is the prediction error, $\mathbb{I}(\cdot)$ is an indicator function. Note that $E[u^q_{jt}|I(t)]$ is part of the expected utility that depends on $E[q_j|I(t)]$ and $\sigma^2_j(t)$. Δu^q_{jkt} is the difference between this partial expected utility from choosing drug j and k.

Our model suggests that manufacturer j has an incentive to increase detailing if $\Delta u^q_{jkt} > 0$. 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 $|\Delta u^q_{jkt}|$ when $\Delta u^q_{jkt} > 0$. We expect the coefficient associated with the interaction term to be positive (i.e., $\lambda_{j1} > 0$). 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 still detail more in order to build up M_j earlier even though $\Delta u^q_{jkt} < 0$. This is because manufacturer j may take into consideration the stochastic nature of Δu^q_{jkt}, which could become positive later. The sign of the coefficient for the interaction term (i.e., λ_{j3}) is therefore ambiguous.

As shown in Eq. 16, the static marginal return of detailing depends on the measure of well-informed physicians for a competing drug as well. This implies that the dynamic marginal return of detailing for drug j will also depend on M_{kt}, $j \neq k$. Therefore, we also allow M_{kt-1} to interact with M_{jt-1} and Δu^q_{jkt}. Following from Eq. 16, if manufacturers are myopic, the sign of λ_{j2} and λ_{j4} would be positive if $\Delta s_{jt}(I_k(t)) > \Delta s_{jt}(\underline{I}^p_k)$, and vice versa. If manufacturers are forward-looking, they will take the future stochastic evolution of $I(t)$ into account, and the sign of λ_{j2} and λ_{j4} would be ambiguous.

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, v_{jt}, in Eq. 17 is normally distributed, we obtain the conditional likelihood of observing D_t,

$$f_d(D_t|(E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}; \theta_s), \tag{20}$$

where θ_s is the vector of parameters.

Assume further that the measurement error, ϵ_{jt}, in Eq. 14 is normally distributed, and denote $f_n(n_t|D_t, (E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}, \text{Size}_t; \theta_d)$ as the likelihood of observing n_t conditional on $(D_t, (E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}, \text{Size}_t)$. The joint likelihood of observing (n_t, D_t) is simply the product of $f_n(n_t|D_t, .)$ and $f_d(D_t|.)$:

$$\begin{aligned} &l(n_t, D_t|(E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}, \text{Size}_t; \theta_d, \theta_s) \\ &\quad = f_n(n_t|D_t, (E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}, \\ &\qquad \text{Size}_t; \theta_d)\, f_d(D_t|(E[q_j|I(t)], \sigma_j(t), M_{jt-1})^2_{j=1}; \theta_s). \end{aligned} \tag{21}$$

TEVA_CAOC_14205482

TE-SF-02127.00023

Now note that $\sigma_j(t)$ is a function of $\{n_{j\tau}\}_{\tau=1}^{t-1}$ (see (6)). Therefore, one can rewrite (21) as,

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

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

$$\begin{aligned} &L(n, D|\{E[q|I(\tau)], M_{\tau-1}, Size_\tau\}_{\tau=1}^T; \theta_d, \theta_s) \\ &= \prod_{t=1}^T l(n_t, D_t|E[q|I(t)], \{n_\tau\}_{\tau=1}^{t-1}, M_{t-1}, Size_t; \theta_d, \theta_s). \end{aligned} \tag{23}$$

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 (2008). The only difference is that we also need to draw the quality signal of the direct comparison clinical trials. We explain how to simulate these draws in Appendix B.

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}, $E[q_j|I(1)]$ 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 to compute. It requires us to explicitly incorporate a 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 $(Size_{t_j^I}, ..., Size_0)$, and the actual clinical trial outcomes between the inception of the incumbent drug and $t = 0$, we use our physician's choice model to simulate the unconditional joint distribution of $(G_{j0}, E[q_j|I(1)], \sigma_j(1))$, which is then incorporated in our likelihood function.

TEVA_CAOC_14205483

TE-SF-02127.00024

5 Results

5.1 Parameter estimates

We now discuss the parameter estimates. The total number of structural demand parameters is 15. 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. We use q_1 and q_2 to denote the true mean qualities for Vaseretic (incumbent) and Zestoretic (entrant), respectively. For identification reasons, we need to normalize q_1, the scaling parameter for the number of consumption experience signals, κ, and the intercept term for the utility of the outside good, α_0. We set $q_1 = 1$, $\kappa = 1/30000$, and $\alpha_0 = 0$. In addition, we restrict $\underline{q}_j^o = \underline{q}_j^p \equiv \underline{q}_j$ 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 differences. These extra restrictions should also help avoid overfitting the model. For instance, we would like to avoid fitting the data very well, but only due to the implausible initial belief of the drug qualities (e.g., the estimated initial prior variance of the new entrant may be much smaller than that of the incumbent). Our approach is similar to Coscelli and Shum (2004) and Crawford and Shum (2005) who imposed extra restrictions (they assume agents have rational expectation), or more recently, Buera et al. (2008), who used informative priors on several parameters.

Table 4 shows the parameter estimates. Model 1 refers to the model presented above. 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 Fig. 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 5.24, which is higher than that of Vaseretic (q_1). The initial prior mean qualities of Vaseretic and Zestoretic are -7.37 and -10.85, 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, the reverse order of their true mean qualities. The initial prior perceived variance is 0.68, which is quite small. This suggests that although the public has a wrong initial prior expectation about the drug qualities, they are quite confident about their beliefs.

Variances associated with consumption experience signals (σ_δ^2) and clinical trial signals (σ_η^2) are both significant. To make these two parameter estimates comparable, we have multiplied σ_η^2 by κ when reporting its estimate in Table 4. The estimate of $\sigma_\eta^2 = 0.04$ is much smaller than that of $\sigma_\delta^2 = 1.30$. This indicates that clinical trial signals provide much more precise information about the true

TEVA_CAOC_14205484

TE-SF-02127.00025

Table 4 Parameter estimates

	Model 1		Model 2	
	Estimates	Standard errors	Estimates	Standard errors
Learning parameters				
σ_δ^2	**1.301**	0.099	**1.248**	0.095
σ_η^2	**0.038**	0.003	**0.035**	0.003
$\underline{q}_1$	**−7.366**	0.134	**−8.187**	0.179
$\underline{q}_2$	**−10.848**	0.220	**−12.102**	0.168
$\underline{\sigma}^2$	**0.683**	0.049	**0.652**	0.048
q_1	1		1	
q_2	**5.241**	0.281	**5.786**	0.326
κ	1/30,000		1/30,000	
Preference parameters				
α	**−3.651**	0.032	**−3.662**	0.041
r	**0.168**	0.007	**0.153**	0.005
π_p	3.51E-04	3.29E-04	2.53E-04	3.28E-04
π_t	**−0.005**	3.32E-04	**−0.005**	3.52E-04
Detailing stock parameters				
Φ_G	**0.045**	0.001	**0.041**	0.001
β_0	**−0.155**	0.098	**−0.083**	0.093
β_1	**9.47E-05**	1.71E-06	**9.90E-05**	1.67E-06
Other parameters for error terms				
s.d.(ε)	**158.547**	9.365	**158.627**	9.358
s.d.(ζ)	1		1	
s.d.(e)	**0.489**	0.014	**0.482**	0.015
Pseudo-detailing policy functions				
λ_{10}	**5.255**	2.194		
λ_{11}	7.904	4.212		
λ_{12}	**−5.910**	1.237		
λ_{13}	0.831	1.066		
λ_{14}	1.054	1.108		
λ_{15}	0.169	0.216		
λ_{20}	**6.245**	0.360		
λ_{21}	**−343.292**	45.208		
λ_{22}	**421.838**	47.891		
λ_{23}	−26.887	18.902		
λ_{24}	28.413	20.541		
λ_{25}	0.080	0.046		
s.d.(υ)	**0.683**	0.036		
Log likelihood	−2,103.116		−953.326	

Estimates shown in bold are significant at 5% level.

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

qualities of the drugs. This is what we expect because participants in clinical trials are usually monitored much closer than regular patients.

All of the preference parameter estimates are statistically significant except for the price coefficient. The insignificant price coefficient 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 expected utility of choosing it. However, the estimate

TEVA_CAOC_14205485

TE-SF-02127.00026

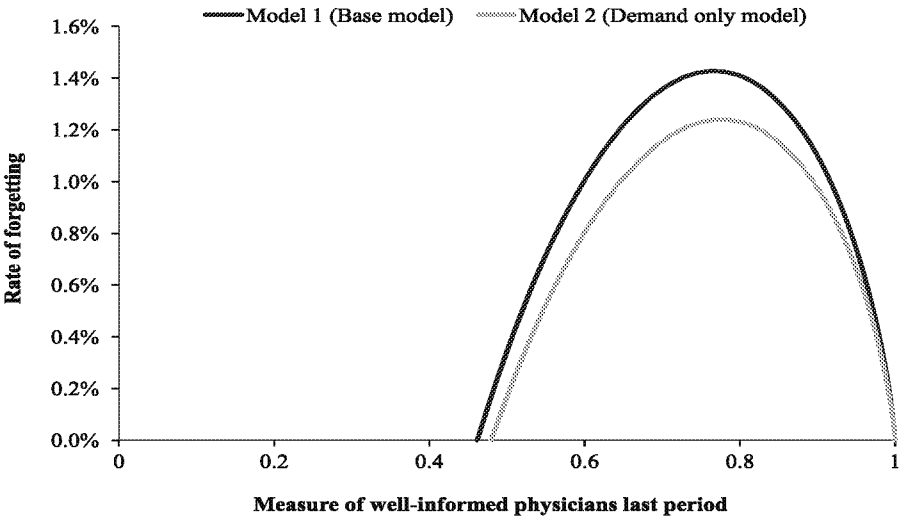

Fig. 5 Rate of forgetting

for r is 0.17, 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 -0.16, which implies that nearly 46% of physicians will be well-informed about $I_j(t)$ (i.e., $M_j = 0.46$) when $G_j = 0$. This represents the percentage of physicians who keep up with the most updated information about ACE-inhibitors with diuretics themselves even without any help from detailing. The estimate of ϕ_G is 4.5%. The implied average rate of forgetting is shown in Fig. 5. As we discussed before, it exhibits an inverted-U shape. The average rate of forgetting starts from 0% at around $M_{jt-1} = 0.46$. It increases and reaches the maximum of 1.4% at around $M_{jt-1} = 0.75$, and then declines. The estimate of β_1 is 9.47e-05. To get a sense of the economic significance of β_1, in Fig. 6 we plot its implied rate of building M_{jt} without forgetting (i.e., $\phi_G = 0$), 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 around 7.0% when M_{jt-1} is around 0.46 (i.e., $G = 0$). Then it declines almost linearly at the rate of 1.3% 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 they influence physicians' choice, we are able to recover them from the evolution of market shares and detailing data. Figure 7 shows the evolution of the measures of well-informed physicians

TEVA_CAOC_14205486

TE-SF-02127.00027

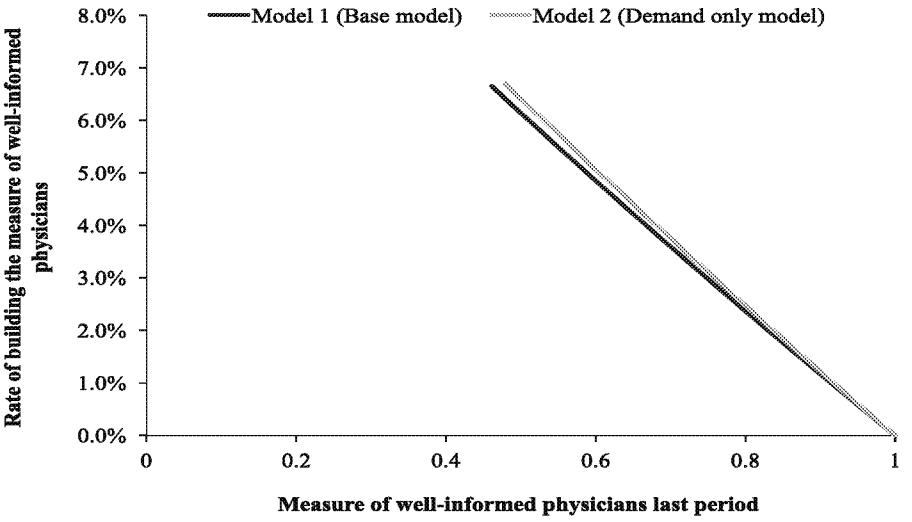

Fig. 6 Rate of building the measure of well-informed physicians

during the sample period. For Vaseretic, the measure of well-informed physicians starts off at around 0.90. It increases to 0.94 after 30 months, and then gradually reduces to around 0.82 at the end of the sample period. For Zestoretic, the measure of well-informed physicians increases from 0.69 to around 0.97. Figure 8 shows how $E[q_j|I(t)]$ evolves over time, since the

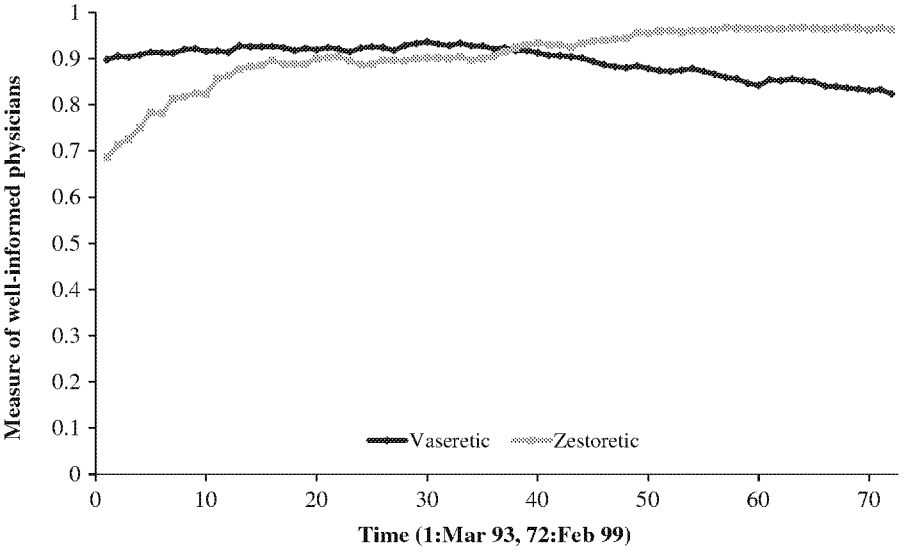

Fig. 7 Measure of informed physicians

TEVA_CAOC_14205487

TE-SF-02127.00028

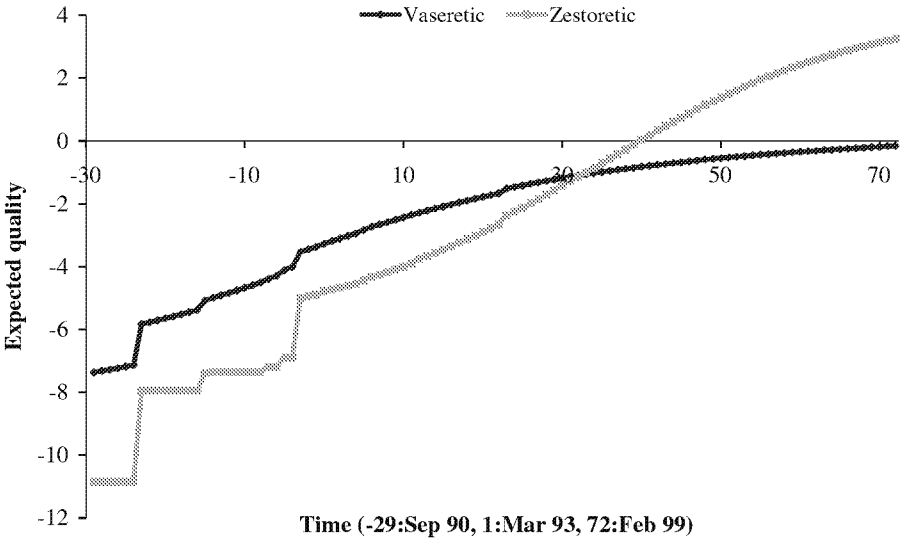

Fig. 8 Expected qualities

inception date of the incumbent drug, Vaseretic. The discrete jumps represent the impacts of clinical trials. Note that the magnitude of a jump partly depends on the number of participants in a clinical trial. As we can see, their impacts are quite strong prior to $t = 0$. Some clinical trials have stronger effects for one drug because they find that it is more effective than the other drug. After $t = 0$, $E[q_j|I(t)]$ increases slowly from around -3.2 to -0.2 for Vaseretic, and it increases at a much faster rate from -4.7 to 3.3 for Zestoretic. Moreover, the impact of clinical trials on $E[q_j|I(t)]$ becomes much smaller. This is mainly because $\sigma_j^2(t)$ has been reducing over time, and as a result, the updating process is putting less weight on the clinical trial signals.[14]

As for the pseudo-detailing policy functions, λ_{j0} and λ_{j2} for $j = 1, 2$, and λ_{21} are significant. Remember that λ_{j1} captures the incentive to detail when the quality of own drug is superior, and λ_{j2} captures its interaction effect with the measure of well-informed physicians for the opponent drug. For Vaseretic ($j = 1$), λ_{11} is positive and λ_{12} is negative. These suggest that the manufacturer responds to favorable information about its drug by increasing the amount of detailing, but the incentive decreases as the measure of well-informed physicians for Zestoretic increases. For Zestoretic ($j = 2$), λ_{21} is negative. This may appear to be inconsistent with our earlier argument that manufacturer j has an incentive to increase detailing when $\Delta u_{jkt}^q > 0$. However, the total effect, $(\lambda_{21} + \lambda_{22} \cdot M_{1t-1}) = -343.3 + 421.8 \cdot M_{1t-1}$, is positive given that M_{1t-1}

[14]Due to the space constraint, we did not plot how $\sigma_j^2(t)$ changes over time. It is available upon request.

TEVA_CAOC_14205488

TE-SF-02127.00029

takes values close to one (see Fig. 7). Thus, the estimates here are still consistent with our earlier argument. Finally, both instrumental variables for Vaseretic (λ_{15}) and Zestoretic (λ_{25}) are not significant at the 5% level, but the one for Zestoretic is significant at 10% level. This may indicate that these are weak instruments.

Our estimated model provides a good fit to the data. To illustrate this, we simulate 5,000 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 9 and 10 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 quite well, in particular, for Zestoretic. This indicates that even though we only have four types of physicians in our model, it is flexible enough to fit the data. Figures 11 and 12 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.

5.2 Effectiveness of detailing

5.2.1 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

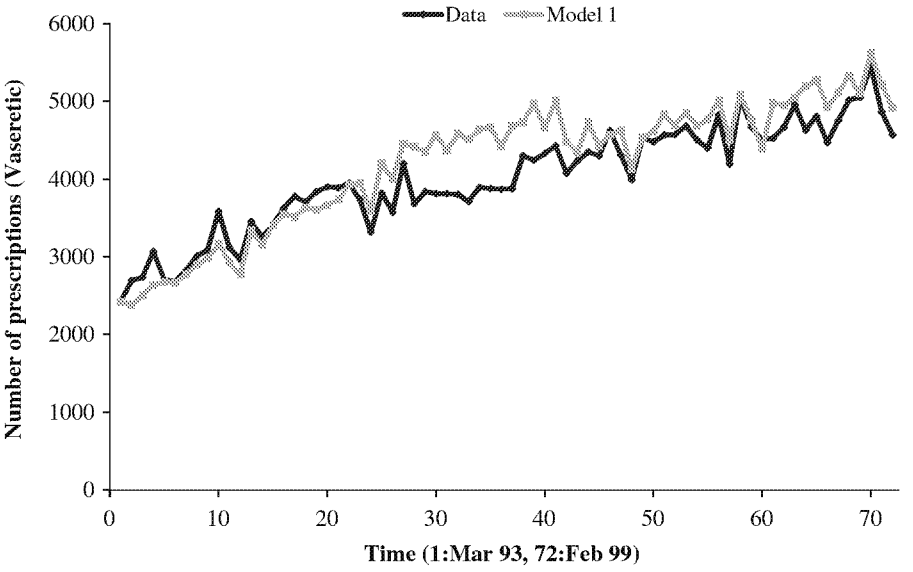

Fig. 9 Predicted and actual demand for Vaseretic

TEVA_CAOC_14205489

TE-SF-02127.00030

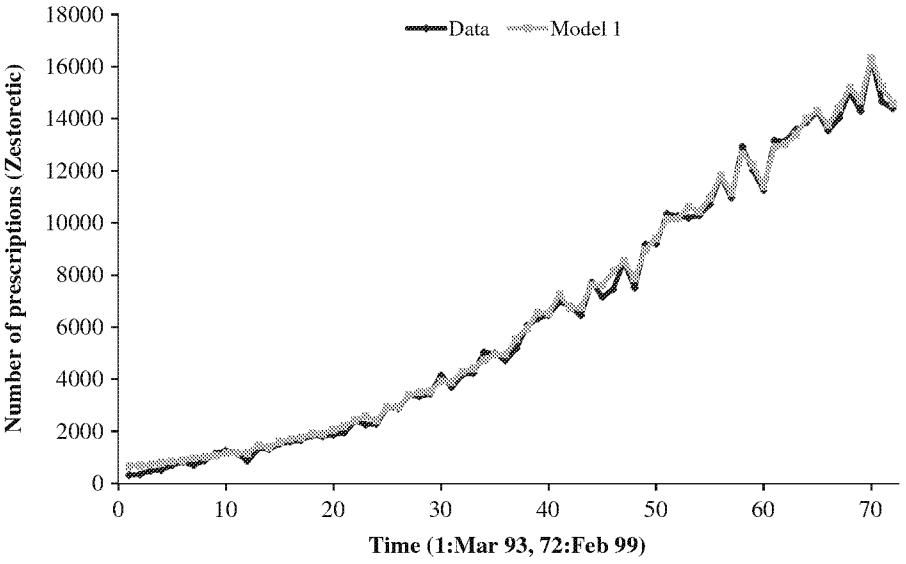

Fig. 10 Predicted and actual demand for Zestoretic

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. We will first

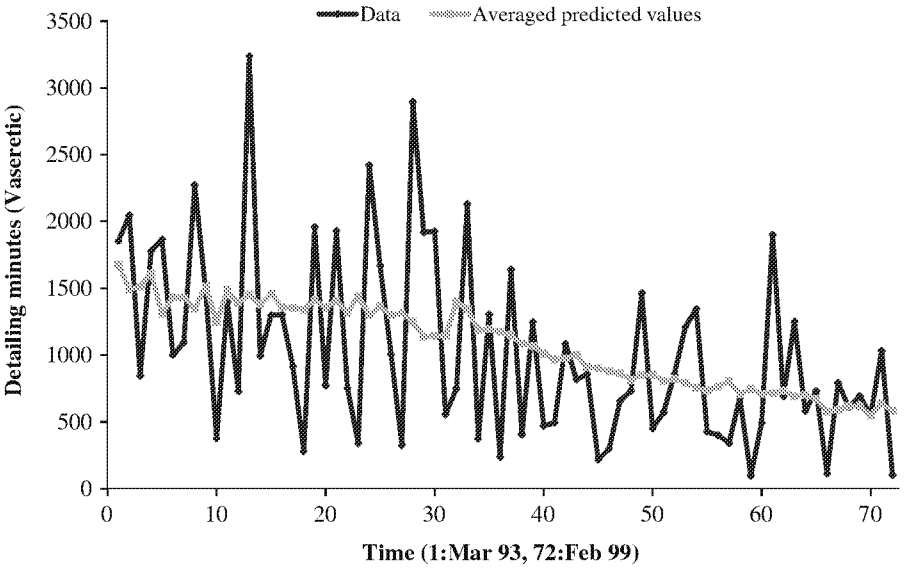

Fig. 11 Predicted and actual detailing minutes for Vaseretic

TEVA_CAOC_14205490

TE-SF-02127.00031

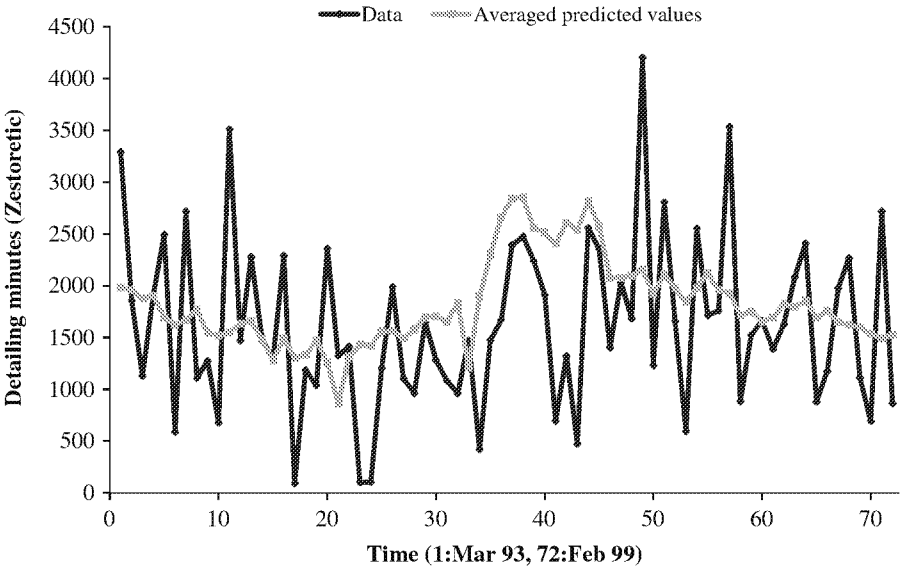

Fig. 12 Predicted and actual detailing minutes for Zestoretic

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 will simulate the model by keeping $M_{1t} = M_{2t}$ for all t in the baseline case, so that the changes in the marginal return of detailing across drugs over time will be mainly driven by $I(t)$. Recall that Vaseretic and Zestoretic entered the market before $t = 1$ (when our sample begins), and Vaseretic entered two years earlier than Zestoretic. To ensure $M_{1t} = M_{2t}$ and obtain the initial value of the information sets at $t = 1$ (i.e., $I(t = 1)$), we set $M_{1t} = M_{2t} = 0.50$ for $t < 1$. For $t \geq 1$, we set $D_{1t} = D_{2t} = 1300$ (which is the average observed amount of detailing across both drugs) and let D_{jt} determine M_{jt}. This ensures $M_{1t} = M_{2t}, \forall t$ in the baseline case. We also set p_{jt} at its average observed values for all t, and assume the outcomes of the clinical trials are from the data. 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: (1) $t = 1$ when the average expected quality for Vaseretic is higher; (2) $t = 26$ when the average expected qualities are about the same for both drugs; (3) $t = 60$ 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. In each scenario, we simulate 5,000 histories of demand.

Panel 1 of Table 5 shows the results. For Vaseretic, the changes in current demand (the number of prescriptions) are 32.3, 26.7, and 22.4 at $t = 1$, 26,

TEVA_CAOC_14205491

TE-SF-02127.00032

Table 5 Effect of a one-time increase in detailing by 50% on current demand

Time	Increase in detailing for Vaseretic		Increase in detailing for Zestoretic		Base demand		Average I(t) ($E[q_j|I(t)]$, σ_j^2 (t))		Measure of well-informed physicians in the last period
	Change in demand		Change in demand						
	Vaseretic	Zestoretic	Vaseretic	Zestoretic	Vaseretic	Zestoretic	Vaseretic	Zestoretic	
Panel 1 (baseline, model 1 estimates)									
1 (Mar 93)	32.302	−6.002	−6.332	20.375	1,410.8	704.5	(−3.55, 0.37)	(−4.68, 0.36)	0.50
26 (Apr 95)	26.654	−9.265	−9.911	27.303	3,268.6	3,211.3	(−1.78, 0.23)	(−1.80, 0.24)	0.85
60 (Feb 98)	22.360	−9.214	−17.763	52.973	4,688.0	10,878.4	(−0.52, 0.12)	(2.49, 0.09)	0.92
Panel 2 (Without controlling for endogeneity of detailing, model 2 estimates)									
1 (Mar 93)	33.415	−6.367	−6.690	21.321	1,398.7	709.0	(−3.55, 0.37)	(−4.68, 0.36)	0.50
26 (Apr 95)	27.045	−9.359	−9.989	27.578	3,254.4	3,180.5	(−1.78, 0.23)	(−1.80, 0.24)	0.85
60 (Feb 98)	22.434	−9.239	−17.808	53.126	4,683.2	10,861.8	(−0.52, 0.12)	(2.49, 0.09)	0.92
Panel 3 (Policy experiment, model 1 estimates with doubled kappa)									
1 (Mar 93)	48.011	−8.014	−6.123	20.570	1,949.3	711.2	(−2.42, 0.28)	(−4.48, 0.35)	0.50
26 (Apr 95)	31.285	−11.327	−15.770	46.181	3,793.5	5,431.5	(−0.60, 0.13)	(0.57, 0.16)	0.85
60 (Feb 98)	25.080	−10.385	−20.661	61.929	5,241.0	12,717.3	(0.22, 0.06)	(4.04, 0.04)	0.92

TEVA_CAOC_14205492

TE-SF-02127.00033

and 60, respectively.[15] The effect decreases monotonically despite the fact that Vaseretic's average $E[q_1|I(t)]$ improves from -3.55 to -0.52. One reason is that Zestoretic's average $E[q_2|I(t)]$ improves even more from -4.68 to 2.49. This reduces the attractiveness of Vaseretic to physicians. Another reason is that there is a diminishing return in building up the measure of well-informed physicians. M_{1t} increases from 0.50 to 0.92. According to Eq. 16, a lower return in building up M results in a smaller effect of detailing on current demand. For Zestoretic, we find a monotonically increasing pattern: The changes in current demand are 20.4, 27.3, and 53.0 at $t = 1, 26$, and 60, respectively. Unlike Vaseretic, the improvements of the information set outweigh the decreasing return in building up M. Overall, the results show that the marginal impact of detailing exhibits a pattern that captures the evolution of the relative qualities of the two drugs. In contrast, as we discussed earlier, the marginal impact of informative detailing necessarily decreases over time for models based on the framework of Erdem and Keane (1996).

5.2.2 The importance of endogeneity of detailing

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 4, under Model 2 (demand only model). The estimate for β_1 is 9.90e-05. This is slightly higher than the estimate from the base model (i.e., Model 1), which is 9.47e-05. The depreciation rate of the detailing stock, ϕ_G, is 0.041, which is lower than the estimate 0.045 in the base model. A likelihood ratio test rejects the hypothesis that the estimates of $(\beta_0, \beta_1, \phi_G)$ in the base model are the same as those in Model 2 at 5% significance level. 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 in terms of economic significance, we plot the implied average rate of forgetting from the demand only model in Fig. 5, and the implied rate of building M in Fig. 6. The average rate of forgetting is slightly biased downward, with its peak at 1.2% instead of 1.4%; the rate of building M is also slightly biased upward.

To understand how the bias would affect the estimates of the effectiveness of detailing, we repeat the exercise in Section 5.2.1 by using the parameter estimates from Model 2. We use the same simulated values of $I(t)$ and M_{jt-1} at $t = 1$, 26, and 60 from the baseline simulation in Panel 1 of Table 5. 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

[15] Although the effect of a one-time increase in detailing might seem small, the impact in terms of elasticity is in line with estimates from some previous studies. For example, Berndt et al. (1997) estimated *the upper bound* of the elasticity of demand w.r.t. cumulative detailing minutes to be 0.67. We find that the point estimates of the elasticity range from 0.10 to 0.27 here (they differ across time). Moreover, it should be noted that a one-time increase in detailing will also have long-term impacts on demand, which we did not show here.

TEVA_CAOC_14205493

TE-SF-02127.00034

detailing. The results are reported in Panel 2 of Table 5. The changes of the current demand are 33.4, 27.0 and 22.4 for Vaseretic, and 21.3, 27.6 and 53.1 for Zestoretic, at $t = 1, 26,$ and 60, respectively. Compared with the baseline case (Model 1, Panel 1 of Table 5), this suggests that the effectiveness of detailing would be slightly biased upward if we do not take the endogeneity into account. However, the magnitude of the bias appears to be not too significant. One reason why we do not find much bias could be because the month-to-month variations in the observed detailing minutes are mainly driven by some exogenous shocks. The noisiness of the detailing data appears to be consistent with this hypothesis (see Fig. 2).

5.2.3 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? We will use our structural model to address this question. To illustrate this, we re-simulate the effects of detailing in our model using the procedure above by doubling the value of κ. This is equivalent to doubling the rate of reporting patients' experiences to the public domain (e.g., via the Health Canada on-line reporting system).[16] Panel 3 of Table 5 shows the results. Compared with the baseline case in Panel 1 of Table 5, the information set, $I(t)$, has improved much quicker, and the changes of current demand for both Vaseretic and Zestoretic are higher at $t = 1, 26,$ and 60. However, the pattern of the improvement appears to be quite different for these two drugs. For Vaseretic, the increases in the effectiveness of detailing are the highest at $t = 1$ and then diminish over time. For Zestoretic, the increase in the effectiveness is very small at $t = 1$, becomes much higher at $t = 26$ (from 27 to 46), and then drops at $t = 60$ (from 53 to 62).

The reason why the increase is the highest at the beginning for Vaseretic is because it entered the market much earlier than Zestoretic. As a result, when doubling κ, there are many more experience signals available to update the prior for Vaseretic. For Zestoretic, the detailing effectiveness improves significantly at $t = 26$ mainly because the difference between its true quality and the initial prior quality is much larger than that for Vaseretic. As a result,

[16]The main objective of this policy experiment is to illustrate the empirical implications of our model. In order to quantify the impact of this policy accurately, one needs to calibrate the baseline value of κ, which requires additional data on the rate at which patients report their experiences to the public domain.

TEVA_CAOC_14205494

TE-SF-02127.00035

doubling κ has stronger impact for Zestoretic, as it allows the public to learn about the truth faster. For both Vaseretic and Zestoretic, the improvements in the effectiveness eventually decline over time mainly because the marginal return of experience signals in improving the information set must diminish. Given these results, marketing managers should consider *increasing* the amount of detailing in this market, in particular for Zestoretic, if this campaign is carried out.

It is important to understand the intuition behind these results. They are partly 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 very well shift down under this campaign if the market has optimistic initial prior about drug qualities. In that case, the expected qualities will be revised downward more quickly over time, and the implications would be that marketing managers should reduce their detailing efforts under such a campaign.

The discussion above again highlights the difference between our model and the traditional learning models pioneered by Erdem and Keane (1996), 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 necessarily 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.

5.3 Remarks

It should be emphasized that our model is not necessarily better than the previous learning models. Clearly, if we consider a market where manufacturers indeed have complete information about their products throughout the product lifecycle, and advertising actually provides information about the true product quality, using our model to conduct policy experiments could very well generate misleading managerial implications. Rather, our results point out that it is crucial for researchers to investigate the mechanisms of how advertising/detailing convey information in the market that they study because different ways to model informative detailing/advertising could lead to very different managerial implications.

There are at least two ways to check which modeling approach is more appropriate for a therapeutic category. First, one can check how many clinical trials are conducted after drugs have been approved for marketing in that category. If the number of post-marketing clinical trials is large, this suggests that firms may be quite uncertain about drugs' qualities even though the public health agency has approved them. This would indicate that our model could be more applicable. In contrast, if firms hardly conduct clinical trials after they have gained the approvals to sell their drugs (and the potential size of the

TEVA_CAOC_14205495

TE-SF-02127.00036

market is large), this suggests that the approach of Erdem and Keane (1996) could be a better approximation. Second, one can run several reduced-form regressions to see if there is any evidence that the effectiveness of detailing is influenced by the outcomes of cumulative clinical trials (see, for example, Table 3). If the interaction terms between detailing and cumulative clinical trials (or cumulative detailing and cumulative clinical trials) are significant, this would provide support for our model. Otherwise, the Erdem and Keane (1996) framework may be more appropriate to be used.

We should also emphasize that our model is not necessarily better than a model that simply allows a detailing goodwill stock to enter the utility function. There are certainly situations where this simpler traditional approach is more appropriate, in particular for drugs that are mature and do not have much new discoveries in clinical trials. Detailing goodwill stock in the utility function is a reasonable way to capture the bribery effect which is widely discussed in the literature.[17] Our new modeling approach should therefore be treated as an alternative approach to capture the detailing effects.

6 Conclusion

Motivated by recent empirical findings on the relationship between new clinical evidence and the effectiveness of detailing, we develop a new structural model of physicians' prescribing decisions and detailing under the environment where both manufacturers and physicians are uncertain about drug qualities. We introduce a representative opinion leader, whose role is to update the most current information about drug qualities based on past consumption experiences and the outcomes of clinical trials. 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, and model physician forgetting by allowing the measure of well-informed physicians to decrease if current detailing efforts are too low. This allows our model to directly link the marginal return of detailing to the measure of well-informed physicians and current information sets. We also discuss the differences between our model and the previous models in terms of empirical implications. Researchers could potentially base on these differences to check which model is more suitable for them to use when analyzing their data.

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

[17] It should be noted that since we do not allow for the bribery effect of detailing in our model, we may overestimate the importance of informative effect.

TEVA_CAOC_14205496

TE-SF-02127.00037

the diffusion pattern well. Using our estimates, we demonstrate how the effectiveness of detailing depends on the current information set and the measure of well-informed physicians. Our results show that the effectiveness of detailing exhibits a pattern that captures the evolution of the relative qualities of the two drugs, which is consistent with recent empirical findings (e.g., Azoulay 2002; Venkataraman and Stremersch 2007). We also 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 their detailing efforts. The implications are diametrically different from the previous learning models, which imply that managers should reduce the detailing efforts under such a campaign. We emphasize that this does not mean that our model is necessarily better than the previous learning models. Rather, our results point out the importance of developing a structural model of detailing that would capture the essential institutional details of the market under study.

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.

Although we present our model in the context of pharmaceutical demand, the framework could 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 served as a starting point to analyze their roles and potentially improve our understanding about how information is transmitted in markets other than prescription drugs.

Most studies have limitations and ours is no exception. One 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 of them 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,

TEVA_CAOC_14205497

TE-SF-02127.00038

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 that can be estimated simply using product level data, which is the most commonly used data in this market. We hope future research will extend our framework to allow for multiple representative opinion leaders. An interesting research direction is to use individual level data to examine the role of opinion leaders. A recent study by Nair et al. (2009) is taking this important step to examine the effects of heterogeneous opinion leaders on physician decisions.

Another limitation is that our model does not take into account the "bribery" effect. Sales representatives often give away gifts during their visits. Critics argue that these gifts may affect physicians' prescribing behavior. The main difficulty of incorporating the bribery effect is that there is no data on the amount of gifts given by sales representatives. The traditional approach to handle this is to allow a detailing goodwill stock to enter the utility function directly (e.g., Anand and Shachar 2005; Narayanan et al. 2005). Unfortunately, with product level data, it is difficult to separately identify the bribery effect and the informative effect that we model here (other than relying on the functional form assumptions). If the bribery effect is important, we would overestimate the informative role of detailing in this paper. We therefore emphasize that the empirical exercise conducted here is mainly for illustrating the empirical implications of our model. Disentangling between the bribery and the informative effects of detailing will be an important topic for future research.[18] Lastly, 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 also leave modeling the role of these marketing communication mix in the environment we consider here for future research.

Acknowledgements We thank Dan Ackerberg, Pradeep Chintagunta, Avi Goldfarb, Tom Holmes, Ig Horstmann, Ahmed Khwaja, Nitin Mehta, Sridhar Moorthy, Mengze Shi, Wei Tan, various conferences and seminars participants, and in particular two anonymous referees, and the editor, Peter Rossi, for their helpful comments. We also thank Matt Shum for helping us obtain the data from IMS Canada. The usual disclaimer applies. Ching's work on this project is supported by Connaught New-staff matching grant at the University of Toronto.

Appendix A: Identification

A.1 Learning parameters

In this appendix, we show how learning parameters are identified in our model. Given our functional form assumption on the utility function, the subutility

[18] Ching and Ishihara (2009) have recently proposed a new identification strategy to measure these two effects separately. Their idea is to focus on markets where two firms sign a co-marketing agreement to market the same chemical using two different brand-names.

TEVA_CAOC_14205498

TE-SF-02127.00039

associated with the expected quality and perceived variance for informed physicians is written as

$$-\exp(-rE[q_j|I(t)] + \frac{1}{2}r^2(\sigma_j^2(t) + \sigma_\delta^2)).$$

At time t, the expected quality and perceived variance given the past consumption and clinical trial signals are expressed as

$$E[q_j|I(t)] = \frac{\sigma_\delta^2\sigma_\eta^2\underline{q}_j + \sigma_\eta^2\underline{\sigma}_j^2\sum_{s=1}^{N_{jt-1}}\mu_{js}^e + \underline{\sigma}_j^2\sigma_\delta^2\sum_{s=1}^{N_{jt-1}^c}\mu_{js}^c}{\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c}$$

$$\sigma_j^2(t) = \frac{\underline{\sigma}_j^2\sigma_\delta^2\sigma_\eta^2}{\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c},$$

where σ_δ^2 is the signal variance for consumption signals, σ_η^2 is the signal variance for clinical trial signals, $\underline{q}_j$ is the initial mean quality for drug j, $\underline{\sigma}_j^2$ is the initial perceived variance for drug j,[19] N_{jt} is the cumulative number of consumption signals for drug j at time t, N_{jt}^c is the cumulative number of clinical trial signals for drug j at time t, and μ_{js}^e and μ_{js}^c are the sth consumption and clinical trial signals for drug j, respectively. Note that $\mu_{js}^e = q_j + \sigma_\delta\delta_{js}$ and $\mu_{js}^c = q_j + \sigma_\eta\eta_{js}$ where q_j is the true mean quality for drug j, and δ_{js} and η_{js} are signal noises. Thus, for the subutility function, the terms inside the exponential function above can be rewritten as

$$\begin{aligned}
&rE[q_j|I(t)] - \frac{1}{2}r^2(\sigma_j^2(t) + \sigma_\delta^2)\\
&= \frac{r(\sigma_\delta^2\sigma_\eta^2\underline{q}_j + \sigma_\eta^2\underline{\sigma}_j^2\sum_{s=1}^{N_{jt-1}}(q_j + \sigma_\delta\delta_{js}) + \underline{\sigma}_j^2\sigma_\delta^2\sum_{s=1}^{N_{jt-1}^c}(q_j + \sigma_\eta\eta_{js}))}{\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c}\\
&\quad - \frac{\frac{1}{2}r^2(\underline{\sigma}_j^2\sigma_\delta^2\sigma_\eta^2 + \sigma_\delta^2(\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c))}{\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c}\\
&= \frac{\sigma_\delta^2\sigma_\eta^2(r\underline{q}_j - \frac{1}{2}r^2(\underline{\sigma}_j^2 + \sigma_\delta^2))}{\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c} + \frac{\sigma_\eta^2\underline{\sigma}_j^2(rq_j - \frac{1}{2}r^2\sigma_\delta^2)}{\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c}N_{jt-1}\\
&\quad + \frac{\underline{\sigma}_j^2\sigma_\delta^2(rq_j - \frac{1}{2}r^2\sigma_\delta^2)}{\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c}N_{jt-1}^c\\
&\quad + \frac{r\underline{\sigma}_j^2\sigma_\delta\sigma_\eta\left(\sigma_\eta\sum_{s=1}^{N_{jt-1}}\delta_{js} + \sigma_\delta\sum_{s=1}^{N_{jt-1}^c}\eta_{js}\right)}{\sigma_\delta^2\sigma_\eta^2 + \sigma_\eta^2\underline{\sigma}_j^2 N_{jt-1} + \underline{\sigma}_j^2\sigma_\delta^2 N_{jt-1}^c}
\end{aligned}$$

[19] To simplify notations, we drop the o superscript for $\underline{q}_j$ and $\underline{\sigma}_j^2$.

TEVA_CAOC_14205499

TE-SF-02127.00040

The nonlinear expression above implies that the variations in N_{jt} and N^c_{jt} across periods identify $\underline{\sigma}_j^2\sigma_\delta^2$, $\sigma_\delta^2\sigma_\eta^2$, $\sigma_\eta^2\underline{\sigma}_j^2$, and the following three parts:

$$\sigma_\delta^2\sigma_\eta^2(r\underline{q}_j - \frac{1}{2}r^2(\underline{\sigma}_j^2 + \sigma_\delta^2)), \qquad \sigma_\eta^2\underline{\sigma}_j^2(rq_j - \frac{1}{2}r^2\sigma_\delta^2), \qquad \underline{\sigma}_j^2\sigma_\delta^2(rq_j - \frac{1}{2}r^2\sigma_\delta^2).$$

We can thus pin down $\underline{\sigma}_j^2$, σ_δ^2, and σ_η^2. It is clear that we cannot separately identify r, $\underline{q}_j$, and q_j. In our application, we normalize $q_1 = 1$. This allows us to pin down r and $\underline{q}_j$. Note that since r is common across drugs, we only need to fix the true mean quality for one of the drugs.

Coscelli and Shum (2004) argue that under their functional form assumption on the utility function, $\underline{\sigma}_j$ and r cannot be separately identified. Thus, they consider two alternative normalizations: (1) assume consumers are risk neutral, (2) assume the initial prior mean quality is equal to the true mean quality for all drugs (rational expectation assumption). Our identification requirements are slightly different from Coscelli and Shum (2004) because they assume utility is linear in the expected quality and perceived variance whereas we assume a CARA specification.

A.2 β_0, β_1 and ϕ_G

It is worth discussing the identification of β_1 and ϕ_G (the parameters that determine M). It may first appear that it is hard to separately identify them, because intuitively the effect on M due to an increase in β_1 (which captures the role of building up M) could be canceled by increasing ϕ_G (which captures the depreciation rate of M) appropriately. However, a more careful examination of Eqs. (7) and (8) reveals that there are subtle differences in terms of how M is generated by β_1 and ϕ_G. Notice that for each M_{jt}, there is a unique $(\beta_0 + \beta_1 G_{jt})$ which can be written as

$$\beta_0 + \beta_1\left(\sum_{\tau=1}^{t-1}(1-\phi_G)^{t-\tau}D_{j\tau}\right) + \beta_1 D_{jt}.$$

It should be clear that the variation in D_{jt} across time and drugs would be sufficient to identify β_0, β_1 and ϕ_G. Intuitively, as long as we have more than three observations, there will be overidentifying restrictions that allow us to estimate these three parameters.

Appendix B: Incorporating the outcome of clinical trials in Bayesian updating

In this appendix, we show how to draw a quality signal from a direct comparison clinical trial. Each clinical trial tells us: (1) which drug does better in the trial, (2) the number of participants who take each drug (i.e., n_j^c, for $j = 1, 2$). How do we simulate a draw of $\bar{q}_j^c$, for $j = 1, 2$? Let's first define

TEVA_CAOC_14205500

TE-SF-02127.00041

$\Delta\bar{q}^c \equiv \bar{q}_1^c - \bar{q}_2^c$. Given the distribution of $\bar{q}_j^c$, $\Delta\bar{q} \sim N\left(q_1 - q_2, \sigma_\eta^2\left(\frac{1}{n_1^c} + \frac{1}{n_2^c}\right)\right)$. The outcome of clinical trials belongs to one of the three cases: (1) $\Delta\bar{q}^c > 0$, (2) $\Delta\bar{q}^c < 0$, and (3) $\Delta\bar{q}^c = 0$. In case (1), we make a draw for $\Delta\bar{q}^{c*}$ from $N\left(q_1 - q_2, \sigma_\eta^2\left(\frac{1}{n_1^c} + \frac{1}{n_2^c}\right)\right)$ truncated below at zero. Case (2) is similar. In case (3), we simply set $\Delta\bar{q}^{c*} = 0$. Then, we make a draw for drug 1, $\bar{q}_1^{c*}$, from $N(q_1, \frac{\sigma_\eta^2}{n_1^c})$, and set $\bar{q}_2^{c*}$ by $\bar{q}_2^{c*} = \bar{q}_1^{c*} - \Delta\bar{q}^{c*}$. With these simulated draws, we can update prior using the Bayesian updating rule.

References

Ackerberg, D. A. (2003). Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review, 44*(3), 1007–1040.

Anand, B. N., & Shachar, R. (2005). *Advertising, the matchmaker*. Working Paper, Faculty of Management, Tel Aviv University.

Azoulay, P. (2002). Do pharmaceutical sales respond to scientific evidence? *Journal of Economics and Management Strategy, 11*(4), 551–594.

Berndt, E., Bui, L. T., Reiley, D. H., & Urban, G. L. (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, pp. 277–322). Chicago and London: University of Chicago Press.

Berry, S. T., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica, 63*(4), 841–890.

Buera, F. J., Monge-Naranjo, A., & Primiceri, G. E. (2008). *Learning the wealth of nations*. NBER Working Paper No. 14595.

Caves, R. E., Whinston, M. D., & Hurwitz, M. A. (1991). Patent expiration, entry, and competition in the U.S. pharmaceutical industry. *Brookings Papers on Economic Activity, Microeconomics, 1*, 1–48.

Chan, T., Narasimhan, C., & Xie, Y. (2007). *Impact of treatment effectiveness and side-effects on prescription decisions: The Role of Patient Heterogeneity and Learning*. Working Paper, Olin Business School, Washington University in Saint Louis.

Ching, A. (2000). *Dynamic equilibrium in the US prescription drug market after patent expiration*. Ph.D. Dissertation, University of Minnesota.

Ching, A. (2008). *Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration*. Working Paper, Rotman School of Management, University of Toronto.

Ching, A. (2009). A dynamic oligopoly structural model for the US prescription drug market after patent expiration. *International Economic Review* (in press).

Ching, A., & Ishihara, M. (2009). *Measuring the informative and persuasive roles of detailing on prescribing decisions*. Working Paper, Rotman School of Management, University of Toronto. Available at SSRN: http://ssrn.com/abstract=1377255.

Coleman, J. S., Katz, E., & Menzel, H. (2004). *Medical innovation: A Diffusion Study*. Indianapolis, I.N.: The Bobbs-Merrill Company.

Cooper, R. J., Schriger, D. L., Wallace, R. C., Mikulich, V. J., & Wilkes, M. S. (2003). The quantity and quality of scientific graphs in pharmaceutical advertisements. *Journal of General Internal Medicine, 18*, 294–297.

Coscelli, A., & Shum, M. (2004). An empirical model of learning and patient spillovers in new drug entry. *Journal of Econometrics, 122*, 213–246.

Crawford, G. S., & Shum, M. (2005). Uncertainty and experimentation in pharmacetical demand. *Econometrica, 73*, 1137–1174.

DeGroot, M. H. (1970). *Optimal statistical decisions*. New York: McGraw-Hill.

Elgie, R. G. (2001). *Enhancing transparency in drug price regulation*. Mimeo, Patented Medicine Prices Review Board.

TEVA_CAOC_14205501

TE-SF-02127.00042

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

Greider, K. (2003). *The big fix: How the Pharmaceutical Industry Rips off American Consumers*. New York, N.Y.: PublicAffairs.

Haug, J. D. (1997). Physicians' preferences for information sources: A meta-analytic study. *Bulletin of the Medical Library Association, 85*(3), 223–232.

Heckman, J. (1981). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process and some Monte Carlo evidence. In C. Manski, & D. McFadden (Eds.), *Structural analysis of discrete data with econometric applications*. Cambridge: MIT Press.

Lasser, K., Allen, P., Woolhandler, S., Himmelstein, D., Wolfe, S., & Bor, D. (2002). Timing of new black box warnings and withdrawals for prescription medications. *Journal of American Medical Association, 287*(17), 2215–2220.

Mehta, N., Rajiv, S., & Srinivasan, K. (2004). Role of forgetting in memory-based choice decisions: A structural model. *Quantitative Marketing and Economics, 2*, 107–140.

Mukherji, P. (2002). *A structural model of learning for new products*. Mimeo, Marshall School of Business, University of Southern California.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *Quarterly Journal of Economics, 117*(3), 735–774.

Nair, H., Manchanda, P., & Bhatia, T. (2009). Asymmetric social interactions in physician prescription behavior: The role of opinion leaders. *Journal of Marketing Research* (in press).

Narayanan, S., Manchanda, P., & Chintagunta, P. K. (2005). Temporal differences in the role of marketing communication in new product categories. *Journal of Marketing Research, 42*(3), 278–290.

Salisbury, C., Bosanquet, N., Wilkinson, E., Bosanquet, A., & Hasler, J. (1998). The implementation of evidence-based medicine in general practice prescribing. *British Journal of General Practice, 48*, 1849–1851.

Schweitzer, S. O. (1997). *Pharmaceutical economics and policy*. New York, N.Y.: Oxford University Press.

Thompson, M. L. (1997). Characteristics of information resources preferred by primary care physicians. *Bulletin of the Medical Library Association, 85*(2), 187–192.

Venkataraman, S., & Stremersch, S. (2007). The debate on influencing doctors' decisions: Are drug characteristics the missing link? *Management Science, 53*(11), 1688–1701.

Ziegler, M. G., Lew, P., & Singer, B. C. (1995). The accuracy of drug information from pharmaceutical sales representatives. *The Journal of the American Medical Association, 273*, 1296–1298.

TEVA_CAOC_14205502

TE-SF-02127.00043

Copyright of Quantitative Marketing & Economics is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.