**WORKING PAPER NO: 526**


# Ask Your Doctor If this Product is Right for You: A Bayesian Zero-Inflated Multinomial Joint Model for Patient Drug Requests and Physician Prescriptions

**Bhuvanesh Pareek**
*Indian Institute of Management Indore*
*Indore, Madhya Pradesh – 453331*
*bhuvaneshpareek@gmail.com*


**Qiang Liu**
*Assistant Professor*
*Purdue University – Kannert School of Management*
*West Lafayette, IN 47907 – 1310, US*
*liu6@purdue.edu*


**Pulak Ghosh**
*Professor*
*Decision Sciences & Information Systems*
*Indian Institute of Management Bangalore*
*Bannerghatta Road, Bangalore – 5600 76*
*Ph: 080-26993136*
*pulak.ghosh@iimb.ernet.in*

# Ask Your Doctor if This Product is Right for You: A Bayesian Zero-inflated Multinomial Joint Model for Patient Drug Requests and Physician Prescriptions

Bhuvanesh Pareek, Qiang Liu, and Pulak Ghosh

September, 2016

Bhuvanesh Pareek is Assistant Professor, Department of Operations Management and Quantitative Techniques, Indian Institute of Management, Indore, India; Qiang Liu is Assistant Professor of Marketing, Krannert School of Management, Purdue University, USA; Pulak Ghosh is Infosys Chair Professor of Economics, Presidency University, and Professor, Department of Quantitative Methods and Information Systems, Indian Institute of Management, Bangalore, India.

## Abstract

The goal of this research is to study physicians' prescription decisions and patients' drug request behaviors jointly. We have developed a new zero-inflated multinomial (ZiMNL) choice model to study patient drug request data with excessive zero requests and a standard multinomial logit (MNL) model to capture physician prescriptions decisions. The two models are joined by a flexible nonparametric multivariate distribution for their random effects. We also adopt an analytically consistent expression for the interaction effect in our non-linear and joint modeling framework. We apply our model to a unique physician panel data set from the Erectile Dysfunction category. Our key empirical findings include the following: (1) the triggering of drug requests by DTCA is complicated with category level DTCA reducing patients' probability of making drug requests and drug specific DTCA driving drug requests for the advertised drug; (2) patient characteristics may play a role in the impact of DTCA on drug requests and the impact of patient requests on physicians' prescription decisions; (3) patient drug requests have a significant impact on physicians' prescription decisions and patients can be consistent with physicians in choosing a drug based on patient diagnosis level and some unobserved factors; (4) there are significant correlations among physician-level random effects that drive both patients' drug requests and physicians' prescription decisions, which validates the joint modeling approach.

**Key Words**: zero-inflated, Bayesian, multinomial Logit, patient requests, physician prescriptions, pharmaceutical market

# 1 Introduction

Pharmaceuticals prescribed by more than 800,000 licensed physicians in the United States have generated tremendous benefits for patients by saving lives, increasing life spans, reducing suffering, preventing surgeries, improving life quality, and shortening hospital stays. In addition to improving public health, the U.S. pharmaceutical industry plays a critical role in the economy. According to the 2007 Economic Census, an estimated 1,552 companies in the U.S. developed, manufactured and marketed drug and biological products. In 2011, the U.S. pharmaceutical market, the world's largest, was estimated at $314 billion.[1] Although the pharmaceutical industry is driven by R&D, the total marketing spending to physicians and patients is much higher (e.g., the top 9 firms spend 2.5 times the amount on marketing than on R&D). Among various marketing resources expended, detailing (personal selling through representatives) targets the main decision maker, i.e., physicians, and is the single largest expenditure. In 2011, the U.S. pharmaceutical industry spent $6.5 billion on detailing.[2] In the US, unlike many other countries, not only the physicians, but also the patients are exposed to pharmaceutical firms' promotion. After detailing, the second largest expenditure is the direct-to-consumer advertising (DTCA), which targets the patients and has grown explosively since the Food and Drug Administration (FDA) lifted the restriction on pharmaceutical firms' use of TV advertising in 1997. In 2011, the industry spent $3.9 billion on DTCA.[2] DTCA is thought to have two effects on patient behaviors (Liu & Gupta 2011): (1) DTCA conveys medical information about health conditions and symptoms to patients and may encourage them to seek professional medical help. As a consequence, it may increase the number of patient visits for diagnosis, thus potentially expanding demand for the whole drug category; (2) a typical DTCA often urges patients to talk to the doctor about the advertised drug. Therefore DTCA may motivate patients to ask their physicians for the advertised drug.

---

[1]Source: IMS Health, National Sales Perspectives, Dec 2011
[2]Source: IMS Integrated Promotional Services[TM], 2011.

Although physicians have traditionally diagnosed diseases and decided which treatment is best, this paternalist model, where the patients are mostly passive participants, is changing because patients now have better access to medical information either via DTCA or other channels, and they are increasingly choosing to assert their perspectives especially for chronic diseases or lifestyle drug treatment. In particular, during medical visits patients are becoming increasingly prone to requesting their physicians to prescribe specific prescription drugs. By participating in the medical decision-making process, the patient exercises his or her most fundamental rights as a human being from a humanistic perspective. There is also a growing body of evidence showing that patients who are more informed and active have not only better clinical outcomes, but they also have lower costs for health care (Hibbard et al. 2013). On the other hand, concerns have arisen among medical professionals and public policy makers on patients' abilities in making optimal decisions about their care and treatment along the shift from paternalism to collaboration model. Because of this increasing social trend of patient involvement in medical decision making, it is important to better understand the interrelation of patient drug requests and physician prescription decisions in the context of promotion from pharmaceutical firms. More specifically, what is the impact of DTCA on patient drug request behaviors? How do physician prescription decisions get affected by patient drug requests? How can a patient's informativeness about a drug through DTCA and a physician's enhanced knowledge about a drug through detailing simultaneously motivate the prescription choices? Equally, it is imperative to understand how different covariates pertaining to physicians and patients alike can interact with the aforementioned interrelation. For example, are minority patients more responsive to DTCA in generating drug requests? Are physicians less likely to accommodate drug requests by minority patients? Although understanding all these questions is highly important to the pharmaceutical industry, as well as to the public policy makers, little research has been done on these topics, partially because of the lack of right behavior data at disaggregated level. With a unique data set including drug request behaviors, patient characteristics,

and physician prescriptions outcomes recorded at patient levels, we aim to bridge this gap. However, several analytical challenges rarely appearing in isolation have complicated the modeling of these important questions and therefore call for a comprehensive study of these questions in a unified framework.

The first challenge arises from the existence of an excess of zeros in the physician-reports on patent requests for brands of drugs. This excessive amount of zeros or zero inflations are being contributed by across physician levels, where a large number of physicians from the panel ($\approx 53\%$ in our data) never report patient requests, and within physician level, where physicians report that many patients do not request any drugs ($\approx 83\%$). Conceptually, no request from a patient may come if that patient requests no brands or as a result of the physician's recording error. Spurious over-dispersion occurs because of the presence of these extra zeros. Several studies have shown that if the data have indeed a large amount of zeros, not accounting for them in the model will result in biased estimates of the model parameters. As noted in Hatfield et al. (2012), in this kind of mixed data structure, "one should model the zero/one dichotomy to separately account for the absence/presence of the measured outcome." A typical zero-inflated model is modeled by mixing a degenerate point mass at zero and an appropriate distribution for the nonzero observation (Lachenbruch 2002, Rizopoulos et al. 2008, Ghosh & Albert 2009, Hatfield et al. 2012, Zhang et al. 2006). Although most of the zero-inflated data researches either model count data or positive continuous data, with the exception of Hatfield et al. (2012) who models proportions, we develop a zero-inflated multinomial model to better account for different aspects of the data.

The second challenge comes from the fact that patient decisions to request a particular drug might not be independent of their physician prescription decisions. For example, a patient informed via DTCA is more likely to request the advertised drug just if he/she thinks the drug is more useful for his/her condition through self-diagnosis. The self-diagnosis, if correct, will then be coincident with the physician's diagnosis. As a consequence, there is a chance that patients are more likely to request drugs that physicians are more likely to

prescribe to them. On the other hand, patients might be more likely to end up with a request for a particular drug just because they figure out that physicians are prone to the drug from the conversation. In other words, some patient characteristics or physician characteristics may simultaneously drive patient request behaviors and physician prescription decisions, therefore creating interdependence between these two behaviors. The interdependence, if it exists and is not appropriately accounted for, will then result in an overestimate of the impact of patients' drug requests on physicians' prescription decisions. We deal with this analytical challenge from two dimensions. First, we control for observed patient characteristics in both the patient request decision model and the physician decision model. Therefore if any observed patient characteristic drives both patient requests and physician prescriptions, its effect will be controlled while we disentangle the impact of requests on physicians' decisions. Second, we propose a joint modeling framework allowing for potential interdependence between patients' request behaviors and physician prescription decisions because of unobserved patient or physician characteristics. Specifically, in both the patient request models and the physician prescription decision model, unobserved patient and physician characteristics that drive requests or prescriptions are modeled as random effects of propensity for a drug at physician level. To control for the correlation between the request models and the prescription model, we specify a flexible joint multivariate semi-parametric distribution on the aforementioned random effects from the patient request models and the physician prescription model. With correlated random effects, the proposed approach allows for potential interdependence between behaviors of patient's requests and decisions of physicians at the disaggregate level.

The third challenge is about the interpretation of interaction coefficients in our nonlinear modeling framework. To investigate the role of patient characteristics in the triggering of drug requests and the physician's willingness to accommodate patient drug requests, we interact patient characteristics with DTCA in the patient drug request model and patient characteristics with patient requests in the physician prescription model. In a nonlinear

4

model, however, the interpretation of interaction coefficient is not as straightforward as in a linear model. Neither can the significance of interaction effects be obtained by testing the significance of interaction coefficients in a nonlinear model (Ai & Norton 2003, Norton et al. 2004). The clarification of this distinction has important implications in applied statistics, given the wide use of nonlinear models (e.g., the logit model and Poisson model). To deal with this challenge, we adopt an analytically consistent expression for the interaction effect between patient characteristics and DTCA on patient drug request probabilities and the interaction effect between patient characteristics and requests on physicians' prescription probability. In our application, we further demonstrate that the estimation of interactive effects is straightforward in a Bayesian Markov chain Monte Carlo (MCMC) framework. The conventional approach suggested by Ai & Norton (2003) involves a more complicated derivation of asymptotic variance for the interactive effects.

It should also be noted that not only from the modeling aspect, dealing with the aforementioned analytical challenges has important managerial and public policy implications because it provides a better understanding of the role of DTCA in generating patient requests for the advertised drug and the role of drug requests in motivating physician prescriptions. First of all, pharmaceutical firms can optimize their DTCA spending levels with a better understanding of the effectiveness of DTCA and patient drug requests. Without accounting for zero-inflation within and across physicians appropriately, a request model will result in a biased estimate of the effectiveness of DTCA in generating patient requests and mislead pharmaceutical firms to either overspend or underspend on DTCA. Similarly, without controlling for interdependence between patient request decisions and physician prescription decisions, a prescription model will result in a bias estimate of the impact of drug requests on physician prescriptions and mislead firms to either overspend or underspend their DTCA to generate drug requests. Second, an accurate understanding of the role of patient characteristics in the trigger of drug requests or physician accommodations to drug requests cannot be reached without the aforementioned challenges being handled, in particular, without an appropriate

interpretation of interaction coefficients in our nonlinear framework. A misunderstanding of these interactive effects will result in inappropriate allocations of DTCA resources to target specific segments of patients. For example, an overestimate of a minority's responsiveness to DTCA in generating drug requests will result in excessive DTCA being allocated to target minority patients. Similarly, an underestimate of physician accommodations to minority patient drug requests relative to nonminority patients may mistakenly make inference that minority patients are being discriminated against as they participate in medical decisions.

The remainder of this article is organized as follows. Section 2 describes a motivating example from an erectile dysfunction drug. Section 3 describes our proposed Bayesian joint modeling framework for patient requests and physician prescriptions. Section 4 discusses Bayesian inference for the proposed model framework and develops the idea of interaction effect in a nonlinear model. Section 5 discusses model selection criteria and presents empirical results, followed by a conclusion in Section 6.

# 2 Motivating Data

We focus on a unique physician panel data set from the Erectile Dysfunction category, combining advertising expenditure data. Because this is one of the categories with a significant amount of patient involvement, it provides an appropriate context for the study of patient drug request behaviors and physician prescription decisions jointly.

## 2.1 Erectile Dysfunction (ED) Category

Erectile dysfunction is a condition that affects 15 to 30 million men in the United States (NIH Consensus Conference 1993). Only three oral drugs, of one category called PDE5 inhibitor, have been approved by the U.S. Food and Drug Administration to manage this disease condition. These drugs are Viagra (by Pfizer), Levitra (by GSK Pharmaceuticals and Bayer), and Cialis (by Eli Lilly). The patent for Viagra received approval in 1998, whereas

the other competitors got approval in 2003. The annual sales of Viagra reached $1.1 billion in 2012 and was the 53rd best selling drug among drugs for cancer, HIV, cholesterol, asthma, and anxiety disorder. Even after being launched in 2003, Cialis reached $0.9 billion sales in 2012. The ED market has become competitive over the years. Advertising for ED drugs grew to $337 million in 2009 from $200 million in 2006.[3]

### 2.1.1 Physician-Reported Patient Visits Data

Our first dataset comes from a physician panel managed by a marketing research firm, ImpactRx Inc. The panel consists of a representative sample of the universe of physicians in the United States. Each sampled physician reports prescriptions written for each of his/her patients, any drug requests made by the patient, detailing visits, and each patient's characteristics, including insurance status, diagnosed level of severity, ethnicity/race backgrounds, and age. We observed the data for 44 months from May 2002 to December 2005. Since Cialis and Viagra were launched in 2003, to avoid the complication from the new drug launch and to solve the initial condition for advertising carry-over effects, we are modeling the prescription and request data from January 2004 to December 2005 and use the beginning 20 months to calculate initial stock for DTCA. Because pharmaceutical firms normally focus on acquiring new customers in their marketing competition, we therefore, focus on new patient visits. The result leaves us with 8,053 patient prescriptions from 567 physicians over a 24-month (month 21 - month 44) period.

The variable *Minority* captures a patient's race/ethnicity, with 1 indicating the patient is African-American or Hispanic and 0 indicating non-minority. *Mild* captures a physician's diagnosis of a patient's condition, with 1 indicating mild and 0 indicating a severe condition. The third variable, *U/M*, captures the insurance status of a patient, with 1 being uninsured or Medicaid-insured and 0 being others. The fourth variable, *Age*, represents a patient's age. In the study sample, 22.3% are minority patients; 30% patients have mild disease

---

[3]Source: TNS/Kantar Media Intelligence Ad$pender.

**Patients Drug Requests**

Number of Patients

Patient Requests for

**Physician Reporting Requests**

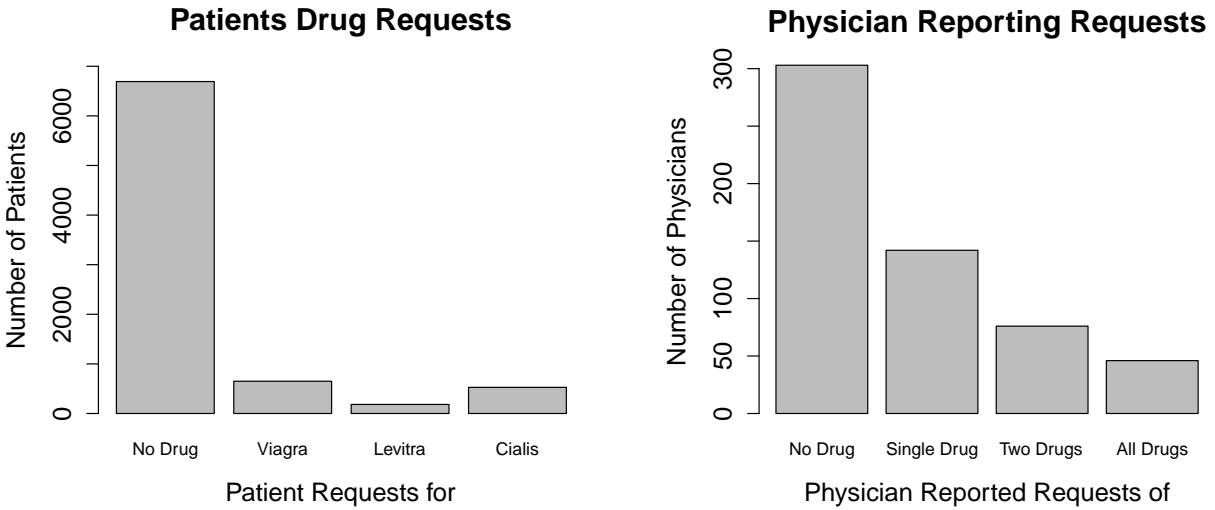Number of Physicians

Physician Reported Requests of

Figure 1: Physicians Reported Patient Request

condition, and 17% are either uninsured or insured by Medicaid. The insurance variable here approximates the socio-economic status of patients because persons with low incomes are more likely to be Medicaid recipients or uninsured (Becker & Newsom 2003). The average age of patients in our sample is 54.

In Figure 1, we show the distribution of patient drug requests. Among 8053 patient visits, the majority (6692) requested no drugs, 651 requested Viagra, 183 requested Levitra, and 523 requested Cialis. Among 567 physicians in our sample, 306 never reported drug requests from their patients, and 261 physicians reported drug requests for at least one of these three drugs. Therefore, a large number of no requests occur within and across physicians, which complicates the modeling.

In Figure 2, we plotted the number of detailing visits aggregated across physicians by brand over time. It is evident that the two relative new drugs delivered more detailing visits than Viagra, the oldest drug. This pattern is especially pronounced for the months after entry of the two new drugs. The figure also shows that all three brands are reducing their detailing levels along with each other overtime, which suggests that the firms in this category compete head-to-head in detailing.
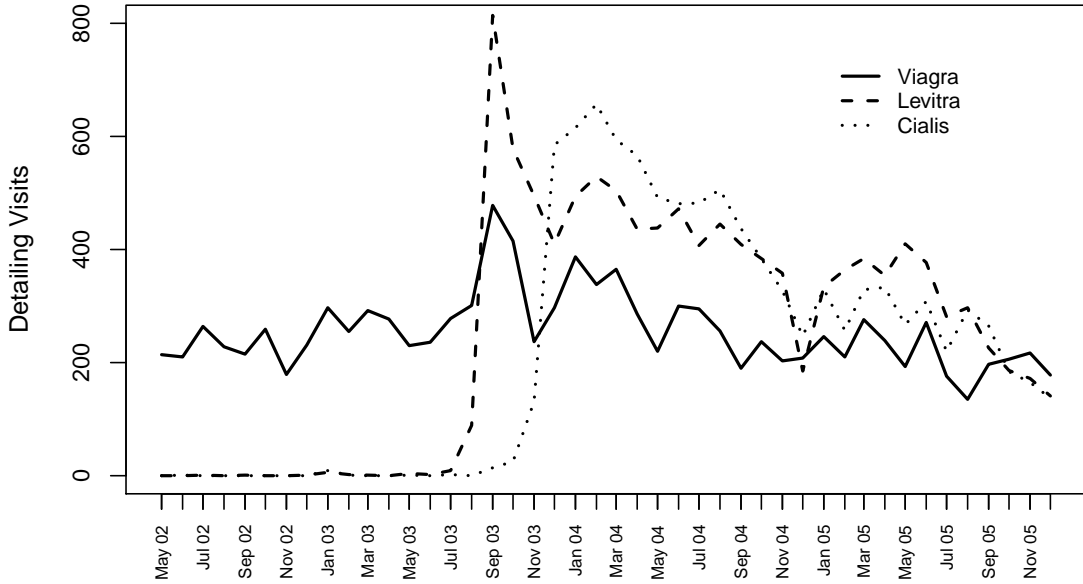
Figure 2: Total Number of Detailing Visits by Brand Name

### 2.1.2 Advertisement Expenditure Data

Our second dataset contains monthly DTCA expenditures for each brand in the category at the Designated Media Area (DMA) level. DMAs are large, and each one contains more than 200 zip codes on average. We map the DTCA expenditure, using DMA to zip mapping, to each physician's practice location. One would expect that pharmaceutical firms spent more on DTCA in more populous DMAs. To control for this difference because of DMA population size, we used DTCA per capita in this study. In Figure 3, we plotted the DTCA per capita expenditure averaged over DMAs for all three brands. Similar to the detailing visits, the two new drugs spent heavily on DTCA during the months after their entries into the market.
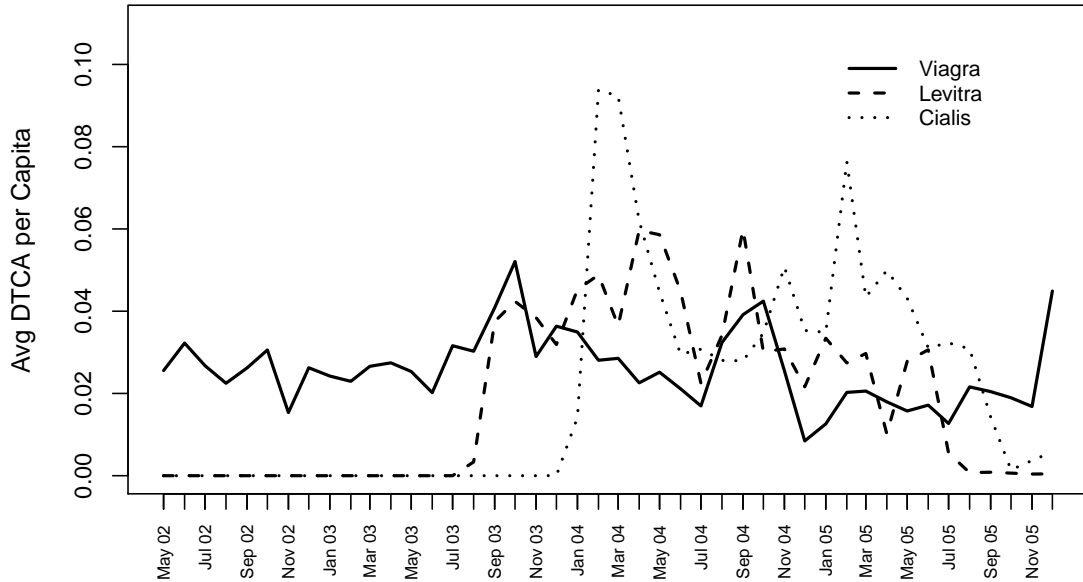
9

Figure 3: Average DTCA per capita by Brand Name

# 3 Zero-inflated Multinomial Joint Model

In this section, we present our modeling framework which consists of two models estimated jointly. First, we propose a zero-inflated mutlinomial model for the drug requests by patients, after which we present a model for physicians' prescriptions decisions under the influence of drug requests and detailing. We finally joined the two models with flexibly correlated random effects at physician level.

## 3.1 Modeling Patient Requests: A Zero-inflated Multinomial Model

Let $y_{ij}$ denote the drug request behavior of the $j$th patient as reported by $i$th physician during the patient's visit to the physician, $j = 1, 2, \cdots, n_i; \ i = 1, 2, \cdots, m$; where $m$ represents the number of physicians in the study, and $n_i$ is the number of patients visited the $i$th physician. $y_{ij} = k$ indicates that the patient requests for drug $k$ $(k = 1, 2, \cdots, K)$ and $y_{ij} = 0$ indicates that the patient does not make any drug request according to the physician's record. Since a large number of zero drug requests observed in $y_{ij}$, our model needs to account for extra

zeros to avoid biased estimates. However, the occurrence of zero requests might be because of two different incidents: zero request because of physicians not reporting and zero request because of patients making no requests. A great many physicians ($\approx 53\%$) reporting no drug requests in our sample suggests that the former incidents are not negligible. Ideally, while modeling patients drug requests, one should distinguish the occurrence of zeros resulting from the two different incidents. However, as there are no covariates to distinguish the two kinds of zeros, we use a binary logit with physician-level random intercepts to model all the occurrences of zero requests and account for the different kinds of zeros in the distribution of physician heterogeneity.

For each observed response, $y_{ij}$, we define:

$$
y_{ij} \begin{cases} = 0, & \text{with probability } (1 - q_{ij}), \quad \text{accounting for the zero observations} \\ \neq 0, & \text{with probability } q_{ij}, \end{cases} \tag{1}
$$

where $q_{ij}$ is the probability of making a request for any drugs by patient $j$ in her/his visit to physician $i$.

When a patient makes a drug request, $y_{ij}$ can take any of the values among $k = 1, 2, \cdots, K$, where $K$ is the number of available drugs for the disease. Thus, we have:

$$
y_{ij} = k | y_{ij} \neq 0 \quad \sim \quad \text{multinomial } (1; \pi_{ij1}, \pi_{ij2}, \cdots, \pi_{ijK}) \tag{2}
$$

It follows that

$$
P(y_{ij} = 0) = (1 - q_{ij}) \tag{3}
$$

$$
P(y_{ij} = k) = q_{ij}\pi_{ijk} \qquad k = 1, 2, \cdots, K. \tag{4}
$$

where $\pi_{ijk}$ denotes the probability that the $j$th patient making request for the $k$th drug to the $i$th physician given that she/he makes a drug request. To model the probability of

observing a request, we use a binary logit model. To model the probability of observing a request for a particular drug given that the patient made a request, we use a multinomial logit model.

Thus, the zero-inflated multinomial (ZiMNL) choice model cane be written as :

$$P(y_{ij} = k) \;=\; (1 - q_{ij})^{I_{\{y_{ij}=0\}}} + (q_{ij}\pi_{ijk})^{I_{\{y_{ij}\neq0\}}} \quad k = 0, 1, 2, \cdots, K \tag{5}$$

and

$$\text{logit}(q_{ij}) \;=\; \beta_0^q + \omega_1 t(j) + X_{ij}^T \beta^q + \delta_0^q TAdS_{it(j)} + (X_{ij}^T \times TAdS_{it(j)})\delta^q + b_i^q \tag{6}$$

$$\pi_{ijk} \;=\; \frac{exp(\eta_{ijk})}{\sum_k exp(\eta_{ijk})} \tag{7}$$

$$\eta_{ijk} \;=\; \beta_{k,0}^\pi + \omega_{2k} t(j) + X_{ij}^T \beta_k^\pi + \delta_0^\pi AdS_{ikt(j)} + (X_{ij}^T \times AdS_{ikt(j)})\delta^\pi + b_{ik}^\pi \tag{8}$$

where $\beta_0^q$ is the intercept for the binary logit model and $\beta_{k,0}^\pi$ is a drug-specific intercept for the multinomial logit (MNL) model; $t(j)$ represents the month in which patient $j$ visits physician $i$; $\omega$ captures a linear time effect on drug requests;[4] $X_{ij} = \{Minority_{ij}, Mild_{ij}, (U/M)_{ij}, Age_{ij}\}$ denotes $j$th patient's (of $i$th physician) race/ethnicity, diagnosis level, insurance status, and age, respectively. Because patient characteristics are common to all drugs, we specify alternative specific coefficients to capture their impacts on patients decisions about which drug to request; to capture the unobserved heterogeneity of drug requests across physicians, we introduce the physician-level random effects, $(b_i^q, b_{i1}^\pi, \cdots, b_{iK}^\pi)$, into the models. Because we do not differentiate two kinds of zero request observations in our model, $b_i^q$ captures two types of physician-level unobserved effects: (1) unobserved characteristics of patient base that drive

---

[4]We also tried alternative time effects, such as quadratic time effects. Our main results remain robust to alternative time effect set ups.

patients to make drug requests or not; (2) some physicians are reporting active while others are not. As a consequence, one may expect the multimodality and skewness in the latent distribution characterizing the physician-level heterogeneity in Equation (6). To address this problem, we introduce a flexible nonparametric distribution for random intercepts in our model, which we will discuss in detail in Section 3.3.

$AdS_{ikt}$ is the accumulated and depreciated DTCA stock for drug $k$ in month $t$ in the DMA where physician $i$ practices. The impact of direct-to-consumer advertising (DTCA) by pharmaceutical companies is expected to carry from one period to the next with deteriorating effectiveness. To capture this long-term dynamic effect, we follow the exponential decaying process advertising model of Nerlove & Arrow (1962) and formulate $AdS_{ikt}$ as follows:

$$\begin{aligned} \text{AdS}_{ikt} &= \text{DTCA}_{ikt} + \lambda_a \text{AdS}_{ik,t-1}; \quad 0 < \lambda_a < 1, t = 1, \cdots, 44 \qquad (9) \\ &= \sum_{\tau=0}^{t} \lambda_a^{(t-\tau)} DTCA_{ik\tau} \qquad (10) \end{aligned}$$

where, $DTCA_{ikt}$ is the dollar-per-capita expenditure of drug $k$'s DTCA in month $t$ in the DMA where physician $i$ practices; and $\lambda_a$ is a carryover parameter that captures the DTCA stock carried over from the previous period. $TAdS_{it(j)} = \sum_{k=1}^{K} AdS_{ikt(j)}$ representing category level accumulated and depreciated DTCA. With this category level DTCA, we can understand whether DTCA expands the whole ED market via patient drug requests or not. Previous studies of consumer purchases and no purchase decisions have used a same or similar approach in modified brand choice multinomial logit model (Ching et al. 2009, Liu et al. 2015). In models (6) and (8), we also include interaction terms between patient characteristics and DTCA stock, which help us to answer some interesting questions from our data.

It is worth to note that DTCA can suffer from possible endogeniety if pharmaceutical companies decide their DTCA expenditures based on market factors not included in our model. Therefore we might end up with a biased estimate of the effect of DTCA on pa-

tient requests because of the endogeniety problem. For linear models, in general two-step approach is utilized, in which instrument variables are used to regress endogenous variable and its predicated values are used to get the unbiased estimator. In our case we utilize a control function approach (Wooldridge 2015, Petrin & Train 2010), which is also a two stage regression with instrument variables. However the residuals instead of the predicted values from the first stage regression are used at the second stage regression to get unbiased estimators.

For $TAdS_{it}$ we use the total DTCA stock by ED firms on drugs other than ED in the DMA where physician $i$ practices in month $t$ as the instrument variable. The first stage regression is specified as the following:

$$TAdS_{it} = \alpha_1^I + \beta_1^I \times TODTCAS_{it} + r_{it}^I \tag{11}$$

where $ODTCAS_{ikt}$ are accumulated and depreciated DTCA expenditure by ED firms on all drugs other than ED and formulated the same way as $TAdS_{ikt}$. To calculate the stock variables for the control functions, we fix carry-over parameters at 0.85 based on previous studies (Berndt et al. 1995).[5]

For $AdS_{ikt}$ we use DTCA stock on drugs other than ED drugs from the pharmaceutical firm $k$ in the DMA where physician $i$ practices in month $t$ as the instrument variable. So the first stage regression is given as:

$$AdS_{ikt} = \alpha_2^I + \gamma_k^I + \beta_2^I \times ODTCAS_{ikt} + \beta_3^I 1(k=1) + \beta_4^I 1(k=2) + s_{ikt}^I \tag{12}$$

where $ODTCAS_{ikt}$ are formulated the same way as $AdS_{ikt}$, $\beta_3^I$ and $\beta_4^I$ captures drug specific fixed effects. Again, the carry-over parameter is fixed at 0.85.

We use the residuals $\hat{r}_{it}^I$ and $\hat{s}_{ikt}^I$ obtained from the first set of regressions in Equation 6

---

[5]We also tried carry-over values between 0.8 and 0.9 for robustness check. Our results remain qualitatively the same for these alternative values.

and 8 as independent variables. We can then rewrite the equations as follows:

$$\text{logit}(q_{ij}) = \beta_0^q + \omega_1 t(j) + X_{ij}^T \beta^q + \delta_0^q T AdS_{it(j)} + (X_{ij}^T \times T AdS_{it(j)})\delta^q + b_i^q + \vartheta_1 \times \hat{r}_{it}^I \quad (13)$$

$$\eta_{ijk} = \beta_{k,0}^\pi + \omega_{2k} t(j) + X_{ij}^T \beta_k^\pi + \delta_0^\pi AdS_{ikt(j)} + (X_{ij}^T \times AdS_{ikt(j)})\delta^\pi + b_{ik}^\pi + \vartheta_2 \times \hat{s}_{ikt}^I \quad (14)$$

Our regressions of total ED DTCA stock and brand specific ED DTCA stock on the instrumental variables yield positive coefficients. This suggests that ED DTCA expenditures are correlated with instrumental DTCA expenditures over time. We believe that the positive correlations are driven by common underlying factors such as media costs. The strength of the instruments are demonstrated by an F-statistic of 7328 for Equation (11) and 758 for Equation (11) against the restricted models that the instruments are irrelevant in the first-stage regressions. A common rule of thumb is that this F-statistic should be larger than 10 (Staiger & Stock 1997). Therefore, the residuals we include in Equations (11) and (12) help us control for the possible endogeneity.

## 3.2   Modeling Physician Prescriptions: A Multinomial Model

Let $R_{ij}$ denote the prescription made by the $i$th physician for the $j$th patient and $\rho_{ijk}$ be the probability of prescribing drug $k$. We use a multinomial discrete choice model to study physicians' prescription decisions under the influence of patient drug requests, physician prescription habits, patient characteristics, detailing, and the word-of-mouth among physicians. Thus a model is given as the following:

$$R_{ij} \sim \text{Multinomial } (1; \rho_{ij1}, \rho_{ij2}, \cdots, \rho_{ijK}) \quad (15)$$

$$\rho_{ijk} = \frac{exp(U_{ijk})}{\sum_{l=1}^K exp(U_{ijl})} \quad (16)$$

where

$$U_{ijk} = \beta^\rho_{k,0} + X^T_{ij}\beta^\rho_k + \delta^\rho_1 H_{ij} + \delta^\rho_2 DetS_{it(j)k} + \kappa^\rho_0 I_{(y_{ij}=k)} + (X^T_{ij} \times I_{(y_{ij}=k)})\kappa^\rho + b^\rho_{ik} \quad (17)$$

where, $\beta^\rho_{k,0}$ is a drug-specific intercept for the multinomial logit model; $X_{ij}$ includes patient characteristics, the same as described above for drug request models; $H_{ij}$ is the prescription decision for a visit proceeding patient $j$'s visit capturing physician $i$'s prescription habits; $I_{(y_{ij}=k)}$ is an indicator variable that assumes a value of 1 if drug $k$ is requested by the patient, otherwise zero; $(b^\rho_{i1}, \cdots, b^\rho_{iK})$ are drug-specific physician-level random effects that captures the unobserved physician preferences for different drugs, including the preferences resulting from physicians' unobserved patient base characteristics. To capture the interactive effects of drug requests and patient characteristics, we include corresponding interaction terms in the model.

$DetS_{ikt}$ is the cumulative and depreciated measurement of detailing visits to physician $i$ at time $t$ from drug $k$. Previous studies have shown that the detailing visits by pharmaceutical sales representatives influence physician prescription decisions in both current and future periods (Ching & Ishihara 2010, 2012, Liu et al. 2015, 2016). Similar to DTCA, we assume an exponential decay process to the detailing-stock $(DetS_{ikt})$ as follows:

$$\begin{aligned} \mathrm{DetS}_{ikt} &= \mathrm{Det}_{ikt} + \lambda_e \mathrm{DetS}_{ik,t-1}; \quad 0 < \lambda_e < 1, t = 1, \cdots, 44 \quad (18) \\ &= \sum_{\tau=0}^{t} \lambda_e^{(t-\tau)} Det_{ik\tau} \quad (19) \end{aligned}$$

where, $\mathrm{Det}_{ikt}$ is the number of detailing visits to physician $i$ at time $t$ from drug $k$. Our end variable $\mathrm{DetS}_{ikt}$ is obtained by summing $\mathrm{Det}_{ikt}$ over months, i.e., $\mathrm{DetS}_{ikt} = \sum_{\tau=0}^{t} \lambda_e^{(t-\tau)} Det_{ik\tau}$.

## 3.3 Correlation Structure and Heterogeneity: Centered Dirichlet Process

As we discussed in the introduction, the models of drug requests and the model of physician prescription decisions might not be independent of each other because certain unobservable factors can potentially drive both patient request and physician prescription decisions. To account for this dependence, we must combine these effects by correlating the multiple outcomes. However, since these outcomes are measured on a variety of different scales (viz., multinomial, ZI), it is not possible to directly model the joint predictor effects because of the lack of any natural multivariate distribution for characterizing such a dependency. A flexible solution is to model the association between different responses (patient drug requests and physician prescriptions) by correlating the random heterogeneous effects from each response. In our joint modeling approach, random effects are assumed for each response process and the different processes are associated by imposing a joint multivariate distribution on the random effects. Such a model not only gives us a covariance structure to assess the strength of association between the responses, but it also borrows information across the outcomes and offers an intuitive way of describing the dependency between the responses.

Let $\mathbf{b}_i = (b_i^q, b_{i1}^\pi, \cdots, b_{iK}^\pi, b_{i1}^\rho, \cdots, b_{iK}^\rho)^T$ be the vector representing the random effects associated with the $i$th physician. Usually, a parametric normal distribution is considered for $\mathbf{b}_i$, though the choice of the normality is often a result of computational tractability, an assumption that may not always hold in reality. It also provided limited flexibility because a normal distribution is limited to a symmetrical and unimodal distribution. In many problems, particularly in our setting, this may result in misleading inferences about the magnitude of effects and nature of heterogeneity. In our setting, we anticipate the random effects to have at least a bimodal shape because of the presence of two kinds of zero requests, viz., zero requests resulting from physicians not reporting, and zero requests resulting from patients making no requests. One common way would be to use a finite mixture of normal distributions as an alternative choice. However, rather than handling the very large number

of parameters resulting from the finite mixture models, it may be easier to work with an infinite dimensional specification by assuming a random mixing distribution which is not restricted to a specific parametric family (Li & Ansari 2014, Voleti et al. 2015, Braun & Bonfrer 2011). We use the richer nonparametric model by assuming a Dirichlet process prior for the $\mathbf{b}_i$ (Ferguson 1973, 1974). Thus, in the context of our proposed models, we assume an unknown distribution $G$ for the random effects which in turn is assumed to be random and a Dirichlet process (DP) is placed on the distribution of $G$. The model for $\mathbf{b}_i$ then can be written as

$$\mathbf{b}_i \sim G, \qquad G \sim DP(\alpha G_0) \tag{20}$$

where $\alpha$ is a positive scalar precision parameter and $G_0$ is a parametric baseline distribution. We assume a multivariate normal distribution for $G_0$, i.e., $G_0 \sim \mathbf{N}(\mathbf{0}, \Sigma)$. Realizations from the DP are discrete with probability one, implying that the estimated $\mathbf{b}_i$ that will be drawn from $G$ will be grouped into cluster, thus allowing for possible multimodality in the distribution of $\mathbf{b}_i$. The discrete nature of the DP can be seen to be obvious from the popular stick-breaking formulation pioneered by Sethuraman (1994). The stick-breaking formulation implies that $G \sim DP(\alpha G_0)$ is equivalent to

$$G = \sum_{h=1}^{\infty} \pi_h^D \delta_{\mathbf{b}_h}, \quad \mathbf{b}_h \sim G_0, \quad \text{and} \quad \sum_{h=1}^{\infty} \pi_h^D = 1 \tag{21}$$

where $G$ is a mixture of countably but infinite atoms, and these atoms are drawn independently from the base distributions $G_0$, and $\delta_{\mathbf{b}}$ is a point mass at $\mathbf{b}$. An atom is like a cluster (i.e., a subgroup of random effects), $\pi_h^D$ is the probability assigned to the $h$th cluster, $\mathbf{b}_h$ is the value of that cluster and all random effects in a cluster share the same $\mathbf{b}_h$. In (21) $\pi_h^D = V_h \prod_{l<h}(1 - V_l)$, which is formulated from a stick-breaking process, with $V_h \sim \text{Beta}(1, \alpha)$. For small values of $\alpha$, $V_h \to 1$ and thus $\pi_h^D \to 1$ assigning all probability weight to few clusters and thus the $G$ is far off from $G_0$. On the contrary for large value of

$\alpha$, number of cluster can be as many as the number of random effects implying the sampled distribution of $G$ is close to the base distribution $G_0$. For practicality, researcher use a finite truncation to approximate the $G$, i.e., $G \sim \sum_{h=1}^{H} \pi_h^D \delta_{\mathbf{b}_h}$.

While the above formulation looks good, there is an issue of identifiability in it in the sense although the prior expectation of the mean of $G$ is 0, the posterior expectation can very well be nonzero an thus can bias inference (Yang 2010, Li & Ansari 2014). In parametric hierarchical models, it is a standard practice to place a mean constraint on the latent variable distribution for the sake of identifiability and interpretability. In nonparametric DP, Yang (2010) proposed to use a entered DP to tackle the identifiability. Li & Ansari (2014) has shown the utility of entered DP in modelling heterogeneity in choice models. Following Yang (2010) and Li & Ansari (2014), we centre the DP to have zero mean. We estimate the mean and variance of the process, i.e., $\mu_G^m$ and $\Sigma_G^m$ at the $m$th MCMC iteration as

$$\mu_G^m = \sum_{h=1}^{H} V_h^m \prod_{l<h}(1 - V_l^m)\mathbf{b}_h^m \tag{22}$$

$$\Sigma_G^m = \sum_{h=1}^{H} V_h^m \prod_{l<h}(1 - V_l^m)(\mathbf{b}_h^m - \mu_G^m)(\mathbf{b}_h^m - \mu_G^m)' \tag{23}$$

where $V_h^m$ and $\mathbf{b}_h^m$ are the posterior samples from the uncentered process defined in (21) and $(\mathbf{b}_h^m - \mu_G^m)$ is the centered estimate for random effects at the $m$th iteration. The above entered DP implies that $E(\mathbf{b}_i|G) = \mathbf{0}$ and $\text{Var}(\mathbf{b}_i|G) = \Sigma_G$.

# 4    Bayesian Inference

## 4.1    Likelihood

Let $\mathbf{y}_i = (y_{i1}, y_{i2}, \cdots, y_{in_i})^T$, and $\mathbf{R}_i = (R_{i1}, R_{i2}, \cdots, R_{in_i})^T$ be the response vectors. Further, we define $\beta^\pi = (\beta_1^\pi, \cdots, \beta_K^\pi)^T$, $\beta_0^\pi = (\beta_{10}^\pi, \cdots, \beta_{K0}^\pi)^T$, $\beta_0^\rho = (\beta_{10}^\rho, \cdots, \beta_{K0}^\rho)^T$, and $\beta^\rho = (\beta_1^\rho, \cdots, \beta_K^\rho)^T$, $\omega_2 = (\omega_{21}, \cdots \omega_{2K})$. Let $\Omega = (\Omega_1, \Omega_2, \Omega_3)$ be the parameter space. Here,

$\Omega_1 = (\beta_0^q, \beta^q, \delta_0^q, \delta^q, \beta_0^\pi, \beta^\pi, \delta_0^\pi, \delta^\pi, \lambda_a, \omega_1, \omega_2)$ is the parameter vector for the patient request model, $\Omega_2 = (\beta_0^\rho, \beta^\rho, \delta_1^\rho, \delta_2^\rho, \kappa_0^\rho, \kappa^\rho, \lambda_e)$ is the parameter vector from the physician prescription model, and $\Omega_3 = (\alpha, \Sigma)$ is the parameter from DP prior for the random physician effects $\mathbf{b}_i$. Then under the assumption that conditional on the correlated random effects, $\mathbf{b}_i$, the models are independent, the joint likelihood can be written as

$$L(\mathbf{y}_i, \mathbf{R}_i, \mathbf{b}_i|\Omega) \propto L(\mathbf{y}_i|\Omega_1, \mathbf{b}_i) \times L(\mathbf{R}_i|\Omega_2, \mathbf{b}_i) \times L(\mathbf{b}_i|\Omega_3) \tag{24}$$

where,

$$L(\mathbf{y}_i|\Omega_1, \mathbf{b}_i) \propto \prod_j^{n_i} [1 - q_{ij}]^{I_{\{y_{ij}=0\}}} \times \left[ (q_{ij})\pi_{ij1}^{I_{\{y_{ij}=1\}}} \pi_{ij2}^{I_{\{y_{ij}=2\}}} \cdots \pi_{ijK}^{I_{\{y_{ij}=K\}}} \right]^{\left(1 - I_{\{y_{ij}=0\}}\right)} \tag{25}$$

with $q_{ij}$ and $\pi_{ijk}$ given in Equations (13) and (14), and

$$L(\mathbf{R}_i|\Omega_2, \mathbf{b}_i) \propto \prod_j^{n_i} \rho_{ij1}^{I_{\{R_{ij}=1\}}} \rho_{ij2}^{I_{\{R_{ij}=2\}}} \cdots \rho_{ijK}^{I_{\{R_{ij}=K\}}} \tag{26}$$

with $\rho_{ijk}$ given in Equation (16), and $L(\mathbf{b}_i|\Omega_3)$ is the likelihood corresponding to the DP prior assumed to characterise the physician-level random effects.

## 4.2  Prior Specification and Posterior Inference

It is worthwhile to note that $\Omega_1$ contains 30 parameters and $\Omega_2$ contains 17 parameters. Since the number of parameters is relatively high, we use shrinkage prior. As discussed in Belloni et al. (2012), this could lead to nonreliable estimates because of the high-dimensionality of the parameter space. In recent times, researchers have being relying on shrinkage methods. Let us assume that $\theta = \{\theta_j; \; j = 1, 2, \cdots, 47\}$ is the set of high-dimensional covariates. A general hierarchical formulation of shrinkage prior would then look like

$$\theta_j | \tau_j^2 \sim N(0, \tau_j^2); \quad \tau_j^2 \sim F$$

People make different choices of $F$ and thus results in different families of shrinkage prior. Belloni et al. (2012) use $F \sim \exp(\lambda^2/2)$ resulting in the famous Lasso prior. Here $\lambda$ is the shrinkage parameter. Lasso is a common and popular shrinkage prior that gives high probability for an estimated parameter to be near zero and in the meantime gives each coefficient a chance to take a large effect. However, a major disadvantage of the Lasso shrinkage method is that it fails to account for the possible multicollinearity between the covariates. This is a serious drawback of Lasso methods because (a) it is hard to check for multicollinearity using variance inflation factor for all possible pairwise covariates; (b) another issue of high-dimensional covariates is spurious correlation that can impose a sample multicollinearity even if there is no theoretical basis for the presence of correlation. Thus we need a prior that not only does the shrinkage as Lasso, but is also robust in the presence of multi-collinearity. The Bayesian elastic net as proposed by Zou & Hastie (2005) does shrinkage even when there are unknown groups of multicollinear predictors.

Thus, we use a Bayesian elastic net prior as follows:

$$\theta_j | \tau_j^2 \quad \sim \quad N(0, \tau_j^2); \quad \tau_j^2 \sim F \tag{27}$$

$$F \quad = \quad (w_j^{-2} + \lambda_2)^{-1}; \quad w_j^2 \sim \exp(\lambda_1^2/2) \tag{28}$$

$$\lambda_1^2 \quad \sim \quad \mathrm{Gamma}(a, b); \quad \lambda_2^2 \sim \mathrm{Gamma}(c, d) \tag{29}$$

For each component of $\Omega_1$ and $\Omega_2$, we specify a Bayesian elastic net prior as mentioned above. The full Bayesian model in the present context is completed by prior assumptions on $\alpha$ and $G_0$.

$$\alpha \sim \text{Gamma}(\nu_1, \nu_2), \quad G_0 \sim N(\mathbf{0}, \Sigma), \text{and} \quad \Sigma^{-1} \sim \text{W}(e, (eE)^{-1})$$

where $W$ stands for an Wishrat distribution. the hyper parameters $a$, $b$, $c$, $d$, $\nu_1$, $\nu_2$, $e$, $E$ are assumed to be known.

The joint posterior distribution of the parameters of the models conditional on the data are obtained by combining the likelihood in (24) and the prior densities using Baye's theorem:

$$f(\Omega, \mathbf{b}|\mathbf{y}, \mathbf{R}) \quad \propto \quad \prod_{i=1}^{m}\{L(\mathbf{y}_i, \mathbf{R}_i, \mathbf{b}_i|\Omega)\} \times \pi(\alpha) \times \pi(\Sigma) \times \prod_{l=1}^{47} \pi(\theta_l) \qquad (30)$$

The posterior distributions are analytically intractable. However, the model can be fitted by using Markov chain Monte Carlo (MCMC) methods, such as the Gibbs sampler, which allows us to generate a sequence of draws from the full-conditional distributions for each parameter conditional on all other parameters. We implemented the model estimation in the WinBUGS, a free available software at www.mrc-bsu.cam.ac.uk/bugs. Convergence was monitored via MCMC chain histories, autocorrelation and cross correlation, density plots, and Brooks–Gelman–Rubin statistics (Brooks & Gelman 1997). We ran two chains of the Gibbs sampler with widely dispersed initial values. The initial values for the fixed parameters were selected by starting with prior mean and covering ±3 standard deviations. For the ZiMNL, we used the "ones trick" (Spiegelhalter et al. 2004). However, it led to very long computational times; thus we used the WinBUGS Development Interface (WBDev; Lunn (2003)) to compile functions and distributions in component Pascal. This is efficiently implemented via the R2WinBUGS (Sturtz et al. 2005) package for R.

## 4.3   Interaction Effects

One of our objectives in this paper is to investigate the interaction effects between patient characteristics and DTCA on patient drug request behaviors and the interaction effects between patient characteristics and patient drug requests on physician prescription decisions. Since we have a nonlinear modeling setups for both patients' drug request behaviors and physicians' prescription decisions, the regression coefficients of interaction terms in our models are not the same as the interaction effects.

To accurately interpret the estimated interaction term coefficients in Equations (14) and (17), we extended the logistic formulation of Norton et al. (2004) to a Multinomial logit model. Let's assume that a dependent variable $y$ can take $k$ values $C_0, C_1 \ldots, C_{k-1}$ (no natural ordering of choices), and $F^i$ is the probability that $y = C_i$ (taking $C_0$ as the base case). Thus $F^i$ has a multinomial choice. In a simplified set-up with $x_1$ and $x_2$ interacted, we have

$$F^i(\mathbf{u}) = F^i(u_0, u_1, \ldots, u_{k-1}) \quad = \quad \frac{\exp(u_i)}{1 + \sum_{j=1}^{k-1} \exp(u_j)}$$

For simplicity, we write the alternative specific latent utility in a general form as $u_i = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + X\beta$ where a coefficient is alternative specific (e.g., $\beta_1^i$) if the covariate is common to all alternatives (e.g., $x_1$), and the coefficient is common to all alternatives (e.g., $\beta_1$) if the covariate is alternative specific (e.g., $x_1^i$). The formulation of interaction effect between $x_1$ and $x_2$ will change based on the nature of the covariates of $x_1$ and $x_2$.

In the first scenario, both $x_1$ and $x_2$ are binary variables (e.g., patient request and patient characteristics). Thus the interaction effect can be written as a discrete double difference:

$$
\frac{\Delta^2 F^i(\mathbf{u})}{\Delta x_1 \Delta x_2} = F^i(\mathbf{u}|_{x_1=1,x_2=1}) - F^i(\mathbf{u}|_{x_1=1,x_2=0})
$$
$$
-F^i(\mathbf{u}|_{x_1=0,x_2=1}) + F^i(\mathbf{u}|_{x_1=0,x_2=0})
$$

In the second scenario, $x_1^i$ is an alternative specific continuous covariate (e.g., drug specific DTCA) and $x_2$ is a binary covariate that is common to all alternatives (e.g., patient Minority). Thus the interaction effect can be calculated by first taking partial derivative with respect to the continuous variable and then taking a discrete difference for the binary variable.

$$
\begin{aligned}
\frac{\Delta}{\Delta x_2} \frac{\partial F^i(\mathbf{u})}{\partial x_1^i} &= \frac{\Delta}{\Delta x_2} \left\{ \frac{\partial u_i}{\partial x_1^i} \cdot \frac{\partial F^i(\mathbf{u})}{\partial u_i} + \sum_{l \neq i} \frac{\partial u_l}{\partial x_1^i} \cdot \frac{\partial F^i(\mathbf{u})}{\partial u_l} \right\} \\
&= \frac{\Delta}{\Delta x_2} \left\{ \frac{\partial u_i}{\partial x_1^i} \cdot \frac{\partial F^i(\mathbf{u})}{\partial u_i} \right\} \\
&= \frac{\Delta}{\Delta x_2} \left\{ (\beta_1^i + \beta_{12}^i x_2) f_i^i(\mathbf{u}) \right\} \\
&= (\beta_1^i + \beta_{12}^i) f_i^i(\mathbf{u}|_{x_2=1}) - \beta_1^i f_i^i(\mathbf{u}|_{x_2=0})
\end{aligned}
$$

where $f_i^i(\mathbf{u}) = \frac{\partial F^i(\mathbf{u})}{\partial u_i}$.

In the third scenario, $x_1^i$ is an alternative-specific continuous covariate (e.g., drug-specific DTCA) and $x_2$ is a continuous covariate that is common to all alternatives (e.g., patient age). Thus the interaction effect can be calculated by first taking partial derivative with respect to $x_1^i$ and then taking partial derivative with respect to $x_2$.

$$
\begin{aligned}
\frac{\partial^2 F^i(\mathbf{u})}{\partial x_2 \partial x_1^i} &= \frac{\partial}{\partial x_2} \left\{ (\beta_1^i + \beta_{12}^i x_2) f_i^i(\mathbf{u}) \right\} \\
&= \beta_{12}^i f_i^i(\mathbf{u}) + \sum_{m=0}^{k-1} (\beta_1^i + \beta_{12}^i x_2)(\beta_2^m + \beta_{12} x_1^m) \tilde{f}_{im}^i(\mathbf{u})
\end{aligned}
$$

where $f_i^i(\mathbf{u}) = \frac{\partial F^i(\mathbf{u})}{\partial u_i}$ and $\tilde{f}_{im}^i(\mathbf{u}) = \frac{\partial^2 F^i(\mathbf{u})}{\partial u_i \partial u_m}$.

In the last scenario, $x_1^i$ is an alternative specific binary covariate (e.g., drug specific request) and $x_2$ is a continuous covariate that is common to all alternatives (e.g., patient Age). Thus the interaction effect can be calculated by first taking partial derivative with respect to the continuous variable, $x_2$ and then hen taking a discrete difference for the binary variable, $x_1^i$.

$$
\begin{aligned}
\frac{\Delta}{\Delta x_1^i} \frac{\partial F^i(\mathbf{u})}{\partial x_2} &= \frac{\Delta}{\Delta x_1^i} \left\{ \sum_{l=0}^{k-1} (\beta_2^l + \beta_{12} x_1^l) f_l^i(\mathbf{u}) \right\} \\
&= (\beta_2^l + \beta_{12}) f_l^i(\mathbf{u}|_{x_1^i=1}) - (\beta_2^l) f_l^i(\mathbf{u}|_{x_1^i=0}) \\
&\quad + \sum_{l \neq i} (\beta_2^l + \beta_{12} x_i^l) \{ f_l^i(\mathbf{u}|_{x_1^i=1}) - f_l^i(\mathbf{u}|_{x_1^i=0}) \}
\end{aligned}
$$

where $f_l^i(\mathbf{u}) = \frac{\partial F^i(\mathbf{u})}{\partial u_l}$.

# 5    Results

We apply the proposed joint modeling framework to the data described in Section 2. In this section, we first compare our models with a couple of competing models, and we then discuss estimation results for each component of our joint modeling framework.

## 5.1    Bayesian Model Selection

In addition to the proposed model (Model-1), we also consider the following two alternative models in terms of model fit:

- Model-2: A joint modeling framework includes a zero-inflated MNL model for patient drug requests and an MNL model for physician decisions, such as those in Model-1, but the models are joined with a parametric normal distribution on random effects, $\mathbf{b_i}$.

- Model-3: A joint modeling framework includes an MNL model for patient drug requests where zero drug request is not modeled, and an MNL model for physician prescription decisions as that in Model-1. The two MNL models are then joined as in Model-1 by a nonparametric multivariate distribution for $\mathbf{b_i}$.

The selection of a model with the best fit is a challenging task in our case because the possible models are non-nested and have different structures and difficulty in integrating out the latent random effects to achieve the marginal model selection criteria. To tackle the above-mentioned problems, we adopt a DIC based on the observed likelihood. The DIC as a model selection approach has been used in several previous studies involving zero-inflated data (Neelon et al. 2013, Montagna et al. 2012). The DIC is a natural generalization of the Akaike Information Criterion (AIC) (Akaike 1973) and interpreted as a Bayesian measure of fit penalized for increased model complexity. Let $\mathbf{D} = (\mathbf{y}, \mathbf{R})$ be the observed data, $\theta$ be the set of parameters and $\mathbf{b}$ is the set of latent random effects variable.

DIC in its basic form is defined by ((Spiegelhalter et al. 2002))

$$\text{DIC}(\mathbf{D}) = \overline{\mathbf{D}(\theta)} + \mathbf{p_D} = -4\mathbf{E}_\theta[\log \mathbf{p}(\mathbf{D}|\theta)|\mathbf{D}] + 2\log \mathbf{p}(\mathbf{D}|\mathbf{E}_\theta(\theta|\mathbf{D})).$$

Although AIC and BIC are well suited for fixed effects models (since the number of parameters are easily determined), DIC is better suited for the hierarchical random effects model because the dimension of the parameter space in a hierarchical random effects model is less clear and depends on the degree of heterogeneity between subjects. However, the above definition of DIC tends to be unreliable and can produce negative $p_D$ for random-effects/mixture models (Celeux et al. 2006). We have a similar issue. Also in our setting, with latent variable $\mathbf{b}$, $p(\mathbf{D}|\theta)$ is not in a closed form. Hence, we follow the approach in Jiang et al. (2015), and Celeux et al. (2006), and calculate $\text{DIC}_4(\mathbf{D})$, by first considering DIC measure with "complete data" with $\mathbf{b}$ and then integrating out the unobserved $\mathbf{b}$.

$$\text{DIC}_4(\mathbf{D}, \mathbf{b}) = -4E_\theta[\log p(\mathbf{D}, \mathbf{b}|\theta)|\mathbf{D}, \mathbf{b}] + 2\log p(\mathbf{D}, \mathbf{b}|E_\theta(\theta|\mathbf{D}, \mathbf{b})).$$

Integrating out **b** gives

$$\text{DIC}_4(\mathbf{D}) \;=\; -4E_{\mathbf{b}}[\{\log p(\mathbf{D}, \mathbf{b}|\theta)|\mathbf{D}\} + 2\log\{(\mathbf{D}, \mathbf{b}|E_\theta(\theta|\mathbf{D}, \mathbf{b})\}] \tag{31}$$

$$=\; -4E_{\mathbf{b},\theta}\{\log p(\mathbf{D}, \mathbf{b}|\theta)|\mathbf{D}\} + 2E_{\mathbf{b}}[\log\{\mathbf{D}, \mathbf{b}|E_\theta(\theta|\mathbf{D}, \mathbf{b})\}|\mathbf{D}]. \tag{32}$$

where integration over **b** is obtained via numerical methods (Jiang et al. 2015).

In Table 1, we report the value of $\text{DIC}_4$ for the three models. As shown in the table, we can see that $\text{DIC}_4$ value for the proposed model is the lowest. Thus among the models considered, our proposed modeling framework is best in terms of the model selection criteria, which enhances our confidence on the robustness of the results reported subsequently.

Table 1: $\text{DIC}_4$ for Three Models

| Model | $\text{DIC}_4$ |
|---|---|
| Model-1 | 14190 |
| Model-2 | 15392 |
| Model-3 | 14882 |

## 5.2 Carryover Effects

We first present the estimates of carryover parameters from Equation (12, 13) and their 95% Bayesian posterior intervals in Table 2. Hereafter in our tables, parameter estimates with * are significant beyond 0.1 level and those with ** are significant beyond 0.05 level.

Table 2: Carryover Effects of DTCA and Detailing

| Variable | Parameter | Mean | 2.5% | 97.5% |
|---|---|---|---|---|
| DTCA | $\lambda_a$ | 0.886** | 0.818 | 0.958 |
| Detaling | $\lambda_e$ | 0.754** | 0.706 | 0.794 |

Our estimate of DTCA carryover effect (0.886) is close to the value (0.85) reported by Berndt et al. (1995). The estimate suggests that it takes around 175 days for one dollar of DTCA stock to depreciate to half of its value. The carryover effect of detailing is found to

be 0.754, which is similar to the value estimated by Yi (2008), and suggests that it takes around 75 days for one detailing visit to depreciate to half of its current value.

## 5.3   Patients' Drug Request Model

We next report empirical results in our patient drug request model. In Table 3, we present estimates of the model of whether a patient makes a drug request or not. We find that patients without insurance or with Medicaid insurance are more likely to request a drug, suggesting those patients are more active in voicing their opinions in life-style drug treatments. On the other hand, patients with milder condition are less likely to make a drug request. This might be because more severe patients are more concerned and more knowledgeable about the treatment. We also find that age is positively related to the likelihood of making a drug request, though the effect is not significant. Elder patients might be more experienced and more comfortable in making requests from physicians. Interestingly, we find that category level advertising stock has a significant negative impact on a patient's intention to request a drug from the category. This might be a surprise for the pharmaceutical industry that expects DTCA to generate more drug requests. However, on the other hand, it is intuitive. With a bombardment of ED advertisements from all competing firms, patients might get confused about which drug to request and become concerned about the side effects statement from DTCA; therefore they are more likely to rely on physicians to make drug choices. In September 2005, the American College Physician Chair said, "DTCA leaves patients confused and misinformed ... " (D. 2004). Even Hoek & Gendall (2003), in their study on New Zealand participants, report one third of the sampled population became confused about what medicine is right for them, and more than 60% got confused about risk/benefit information. Also, excessive competition on DTCA may drive companies to focus on attacking inferior characteristics of each other and drive away patients from requesting any ED drugs. Similarly, Ansolabehere & Iyengar (1997) find that a large fraction of political advertising campaign turns voters off and keeps people away from the polls, thus reducing overall turn

outs and "shrinking the market.".

For the interaction terms, we find a significant positive coefficient for TAdS × Mild, a insignificant coefficient for TAdS × Age, TAdS × Minority, and TAdS × U/M. As discussed in Section 5, the interaction coefficient cannot be interpreted as the interaction effect on dependent variables in a nonlinear model; neither does its significance necessarily represent the significance of the interaction effect. With the simulated MCMC draws for all parameters, we are able to estimate these interaction effects and their credible intervals as shown in Table 4. We find that patients with mild conditions are more likely to make a drug request under the influence of category level DTCA than patients with nonmild conditions probably because patients with milder conditions are less knowledgeable about the treatments in this category initially and therefore the information delivered by DTCA has a higher marginal effect for them than for those with more severe conditions. On the other hand, insurance status, minority and age do not seem to influence category DTCA's effect on drug requests.

Table 3: Patients' Binary Drug Request Model (Eq. 9)

| Variable | Parameter | Mean | 2.50% | 97.50% |
|---|---|---|---|---|
| Intercept | $\beta_0^q$ | 2.439** | 1.115 | 3.396 |
| Age | $\beta_1$ | 0.375 | -0.140 | 0.915 |
| Mild | $\beta_2$ | $-0.992$** | -1.907 | -0.013 |
| Minority | $\beta_3$ | 0.152 | -0.993 | 1.218 |
| U/M | $\beta_4$ | 0.804** | 0.093 | 1.578 |
| TAdS | $\delta_0^q$ | $-2.290$** | -3.500 | -1.126 |
| TAdS × Minority | $\delta_1^q$ | -0.102 | -1.731 | 1.548 |
| TAdS × Mild | $\delta_2^q$ | 1.657** | 0.087 | 3.286 |
| TAdS × U/M | $\delta_3^q$ | -0.414 | -1.756 | 0.659 |
| TAdS × Age | $\delta_4^q$ | -0.153 | -0.944 | 0.606 |
| Linear Time Effect | $\omega_1$ | $-0.209$** | -0.245 | -0.163 |
| Residual Term | $\vartheta_1$ | $-3.122$** | -5.439 | -0.547 |

Next, we present our estimates of patients' MNL drug request model in Table 5. We find that patients with mild conditions are more likely to request the oldest drug, Viagra, probably because they feel Viagra is enough to solve their problems. Older patients, however, are more

Table 4: Interaction Effect in the Patients' Binary Drug Request Model

| Variables | Mean | 2.50% | 97.50% |
|---|---|---|---|
| TAdS × Minority | -0.003 | -0.078 | 0.042 |
| TAdS × Mild | 0.031** | 0.001 | 0.159 |
| TAdS × U/M | -0.007 | -0.060 | 0.028 |
| TAdS × Age | 0.001 | -0.005 | 0.011 |

likely to request the newest drug, Cialis. Patients without insurance or Medicaid insurance and minority patients are less likely to ask for Viagra relative to Cialis suggesting those patients prefer more powerful drugs. Although the category-level DTCA stock negatively impact patients' probability of making a request for ED drugs, we find that drug specific DTCA stock does have a significant positive impact on patient requests for the advertised drug (1.875). This result suggests that DTCA is an important marketing instrument for firms to gain a larger market share via patients' request for the advertised drug.

As to the interaction effect, we find negative interaction effects between drug-specific DTCA stock and patients' minority status. This suggests that minority patients are less likely to be influenced by drug-specific DTCA in requesting a specific drug than nonminority patients are. It is worth noting that both Viagra and Levitra's DTCA have significant interactive effect with minority status in generating drug requests even if the coefficient of AdS × Minority is insignificant. The discrepancy found in our empirical finding highlights the importance of an appropriate interpretation/calculation of interaction effects and their significance in a nonlinear model, as we have specified. Other patient characteristics including U/M, Age, and Mild have positive but insignificant effects with drug-specific DTCA on patient requests for a certain drug. We also visually display all interaction effects from the MNL drug request model in the left panel of Figure 4.

Table 5: Patients' Multinomial Drug Request Model (Eq. 10)

|  | Variable | Parameter | Mean | 2.50% | 97.50% |
|---|---|---|---|---|---|
| Viagra | Intercept | $\beta^\pi_{1,0}$ | 0.943** | 0.087 | 1.69 |
|  | Age | $\beta^\pi_{1,1}$ | $-0.162$** | -0.339 | -0.004 |
|  | Mild | $\beta^\pi_{1,2}$ | 0.698** | 0.407 | 0.980 |
|  | Minority | $\beta^\pi_{1,3}$ | $-0.346$* | -0.687 | 0.031 |
|  | U/M | $\beta^\pi_{1,4}$ | $-0.451$** | -0.905 | -0.063 |
|  | Linear Time Effect | $\omega_{21}$ | $-0.059$* | -0.089 | 0.007 |
| Levitra | Intercept | $\beta^\pi_{2,0}$ | 3.082** | 0.923 | 4.639 |
|  | Age | $\beta^\pi_{2,1}$ | $-0.129$** | -0.363 | -0.089 |
|  | Mild | $\beta^\pi_{2,2}$ | -0.053 | -0.440 | 0.368 |
|  | Minority | $\beta^\pi_{2,3}$ | 0.169 | -0.249 | 0.553 |
|  | U/M | $\beta^\pi_{2,4}$ | -0.476 | -1.048 | 0.161 |
|  | Linear Time Effect | $\omega_{22}$ | $-0.269$** | -0.342 | -0.131 |
|  | AdS | $\delta^\pi_0$ | 1.875** | 0.147 | 4.498 |
|  | AdS × Minority | $\delta^\pi_1$ | -1.756 | -4.199 | 0.626 |
|  | AdS × Mild | $\delta^\pi_2$ | 0.414 | -1.95 | 2.652 |
|  | AdS × U/M | $\delta^\pi_3$ | 1.56 | -1.065 | 4.449 |
|  | AdS × Age | $\delta^\pi_4$ | 0.507 | -0.875 | 1.784 |
|  | Residual Term | $\vartheta_2$ | -0.270 | -3.319 | 2.569 |

Table 6: Interaction Effects in Patients' Multinomial Drug Request Model

| Variables | Drug | Mean | 2.50% | 97.50% |
|---|---|---|---|---|
| AdS × Minority | Viagra | $-0.042$** | -0.412 | -0.006 |
|  | Levitra | $-0.054$** | -0.500 | -0.006 |
|  | Cialis | -0.082 | -0.599 | 0.029 |
| AdS × Mild | Viagra | 0.010 | -0.101 | 0.185 |
|  | Levitra | 0.015 | -0.122 | 0.287 |
|  | Cialis | 0.023 | -0.161 | 0.344 |
| AdS × U/M | Viagra | 0.034 | -0.058 | 0.349 |
|  | Levitra | 0.052 | -0.041 | 0.562 |
|  | Cialis | 0.077 | -0.076 | 0.598 |
| AdS × Age | Viagra | 0.037 | -0.003 | 0.110 |
|  | Levitra | 0.0001 | -0.032 | 0.033 |
|  | Cialis | 0.032 | -0.060 | 0.102 |

## 5.4 Physicians' Prescription Choices

We now discuss the estimation results from the model of physician prescription decisions. As shown in Table 7, physicians are more likely to give the older drug (Viagra) to patients with milder conditions. This is consistent with what we find from the patient drug request model where patients with milder conditions are more likely to ask for Viagra. We also find that physicians are more likely to give older patients Levitra relative to Viagra and Cialis. This finding is not a surprise considering that Levitra is more effective in diabetic patients than other erectile dysfunction medicines, and type 2 diabetes risk increases with age. It is worth noting such a prescription preference is different from patients' preferences for drug requests where older patients tend to favor Cialis versus Viagra and Levitra. This empirical finding suggests that patients and physicians are consistent in choosing the drug based on diagnosis diagnosis level but inconsistent in choosing the drug based on patient age. Consistent with previous studies (Gönül et al. 2001, Dong et al. 2009, Liu et al. 2016), we detect a significant and positive impact of detailing stock on physicians' prescription decisions (0.352). An estimate of the coefficient of prescription habit is significantly positive (3.126) and indicates that physicians tend to have an inertia in their prescription behaviors. The primary term of a patient drug request is significantly positive at the 95% level. We also find significant positive estimates for the coefficients of (Request × Minority) and (Request × Age).

We compute the interaction effects and their 95% Bayesian intervals in Table 8. For all three drugs, we find positive and significant interaction effects between patient drug requests and patients race/ethnicity. Our results suggest that physicians are more likely to accommodate drug requests made by minority patients - notably African Americans and Hispanics - than by nonminority patients. Despite improvements, racial and ethnic minorities are prone to poorer quality health care than white Americans are, even factors such as insurance status are controlled. Bias, stereotyping, and prejudice on the part of health-care providers may contribute to racial and ethnic disparities in health care. Therefore the

positive interaction effects may appear to be counterintuitive because a typical impression is that bias and stereotyping in health care may drive physicians - mainly non-minority themselves - to be more dominant and less likely to accommodate requests made by minority patients than those by nonminority patients. However, physicians' concerns for physician-patient relationship might provide a reasonable explanation for our interesting findings. Recent work shows that minority patients are commonly in ethnic-discordant relationships with health professionals and rate the quality of interpersonal care by physicians in general more negatively than white patients do (Blendon et al. 1995). Studies have also suggested that accommodating patient drug requests helps physicians improve patients' satisfaction and reduces their defection (Bell et al. 1999, Stevenson et al. 2000). In particular, Lee & Beqley (2011) found that African-American patients were significantly more likely to react to a physician's prescription drug refusal by switching physicians. Fearing that refusals may damage the already shaky relationships with these patients, physicians might be less likely to refuse requests by minority patients. Another important reason for the positive interaction effects might be the change of physicians' awareness of race/ethnicity discrimination. Over the years, the long-held historical impression about race/ethnicity discrimination by whites may have made physicians cautious of the race/ethnicity issue, causing them to dislike being perceived as discriminatory. As a result, physicians are less likely to refuse requests made by minority patients. This behavioral "counter-bias" might be rooted in physicians' concerns about the social and legal (also economic as we discussed above) consequences of refusing a member of a historically oppressed racial or ethnic group. Such a behavioral "counter-bias" has been documented in other fields, such as criminology. Through a high-fidelity laboratory simulation, James & Vila (2014) find that participants show a behavioral bias in favor of African-Americans by hesitating longer before deciding to shoot black suspects than white suspects, although they subconsciously bias that African-Americans were more threatening.

As for patient age, we find that physicians are more likely to accommodate drug requests made by older patients than those by younger ones. Our results are consistent with a previous

finding that physicians are more likely to have patient-centered encounters with elderly patients (Peck 2011). We also visually display all interaction effects from the prescription model in the right panel of Figure 4.

Table 7: Physician Drug Prescription Model (Eq. 15)

|  | Variable | Parameter | Mean | 2.50% | 97.50% |
|---|---|---|---|---|---|
| Viagra | Intercept | $\beta_{1,0}^{\rho}$ | 0.281 | -0.109 | 0.616 |
|  | Age | $\beta_{1,1}^{\rho}$ | -0.014 | -0.095 | 0.065 |
|  | Mild | $\beta_{1,2}^{\rho}$ | 0.235** | 0.098 | 0.384 |
|  | Minority | $\beta_{1,3}^{\rho}$ | 0.016 | -0.149 | 0.181 |
|  | U/M | $\beta_{1,4}^{\rho}$ | -0.138 | -0.417 | 0.118 |
| Levitra | Intercept | $\beta_{2,0}^{\rho}$ | -0.222 | -0.517 | 0.153 |
|  | Age | $\beta_{2,1}^{\rho}$ | 0.088** | 0.006 | 0.167 |
|  | Mild | $\beta_{2,2}^{\rho}$ | 0.099 | -0.053 | 0.248 |
|  | Minority | $\beta_{2,3}^{\rho}$ | -0.049 | -0.226 | 0.125 |
|  | U/M | $\beta_{2,4}^{\rho}$ | 0.146 | -0.120 | 0.376 |
|  | DetS | $\delta_1^{\rho}$ | 0.352** | 0.301 | 0.406 |
|  | Physician Habits | $\delta_2^{\rho}$ | 0.492** | 0.410 | 0.577 |
|  | Request | $\kappa_0^{\rho}$ | 3.126** | 2.369 | 3.806 |
|  | Request × Minority | $\kappa_1^{\rho}$ | 1.064** | 0.525 | 1.623 |
|  | Request × Mild | $\kappa_2^{\rho}$ | -0.280 | -0.672 | 0.104 |
|  | Request × U/M | $\kappa_3^{\rho}$ | -0.286 | -1.01 | 0.504 |
|  | Request × Age | $\kappa_4^{\rho}$ | 0.479** | 0.251 | 0.702 |

Table 8: Interaction Effects in Physician Drug Prescription Model

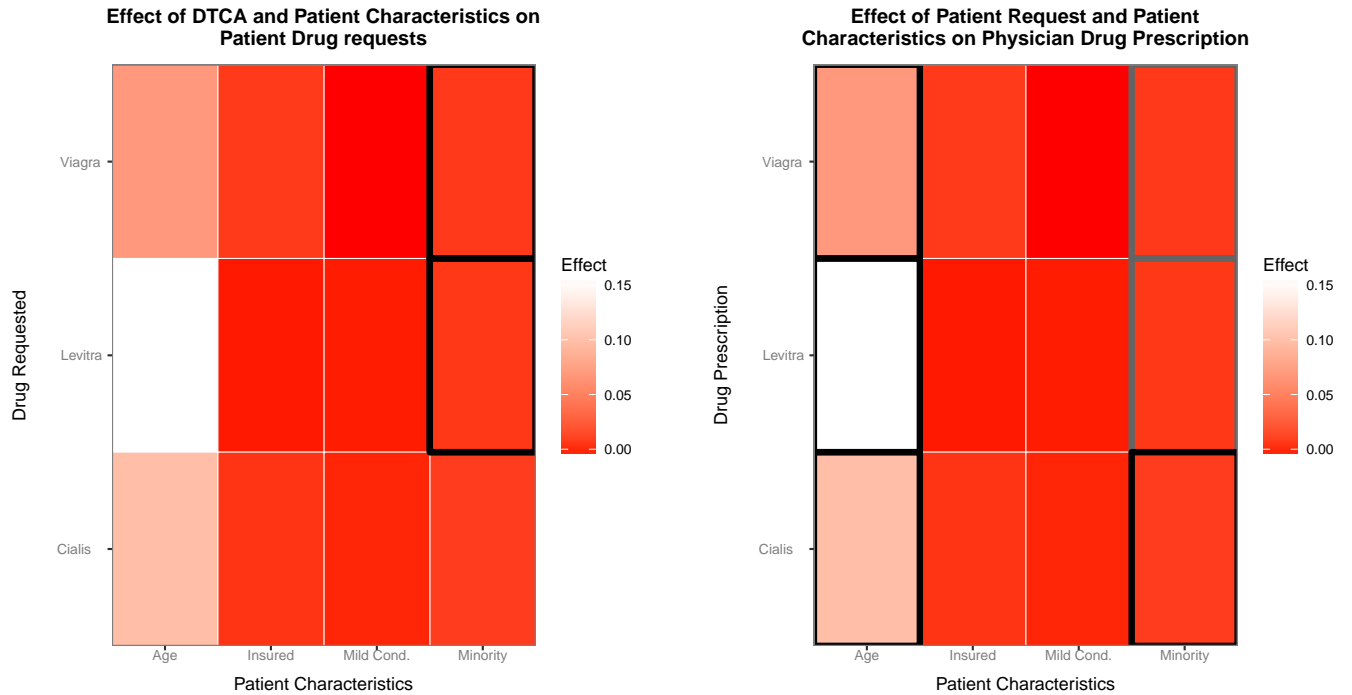| Variables |  | Mean | 2.50% | 97.50% |
|---|---|---|---|---|
| Request × Minority | Viagra | 0.007* | -0.002 | 0.017 |
|  | Levitra | 0.007* | -0.004 | 0.036 |
|  | Cialis | 0.009** | 0.001 | 0.021 |
| Request × Mild | Viagra | -0.009 | -0.048 | 0.002 |
|  | Levitra | -0.003 | -0.019 | 0.003 |
|  | Cialis | -0.001 | -0.007 | 0.007 |
| Request × U/M | Viagra | 0.007 | -0.002 | 0.038 |
|  | Levitra | -0.004 | -0.021 | 0.005 |
|  | Cialis | 0.004 | -0.015 | 0.024 |
| Request × Age | Viagra | 0.070** | 0.001 | 0.495 |
|  | Levitra | 0.155** | 0.008 | 0.598 |
|  | Cialis | 0.100** | 0.004 | 0.540 |

Figure 4: Interaction Effect

## 5.5 Heterogeneity and Correlation Structure

The estimated value of $\alpha$ for Dirichlet process is 2.820 with a 95% Bayesian posterior interval (1.033, 5.575). This indicates the presence of multimodality in the random effects distribution. To visually show the distribution of heterogeneity $\mathbf{b}_i$, we plot the density of each random effect in Figure 5.

As we can see from this figure, the estimated random effects pertaining to the binary logit model of Equation (11) is indeed multimodal and asymmetric. This empirical outcome verifies our assumption that there may be more than one kind of zero-generating process. The estimated random effects pertaining to patient request for Levtira also show some multimodal and asymmetric property that might be because Levitra is fit for ED patients with diabetic condition but our model does not explicitly account for patient diabetic condition because of the lack of such data. Other random effects tend to be more symmetric.

We finally discuss the joint multivariate distribution that links the patient drug request
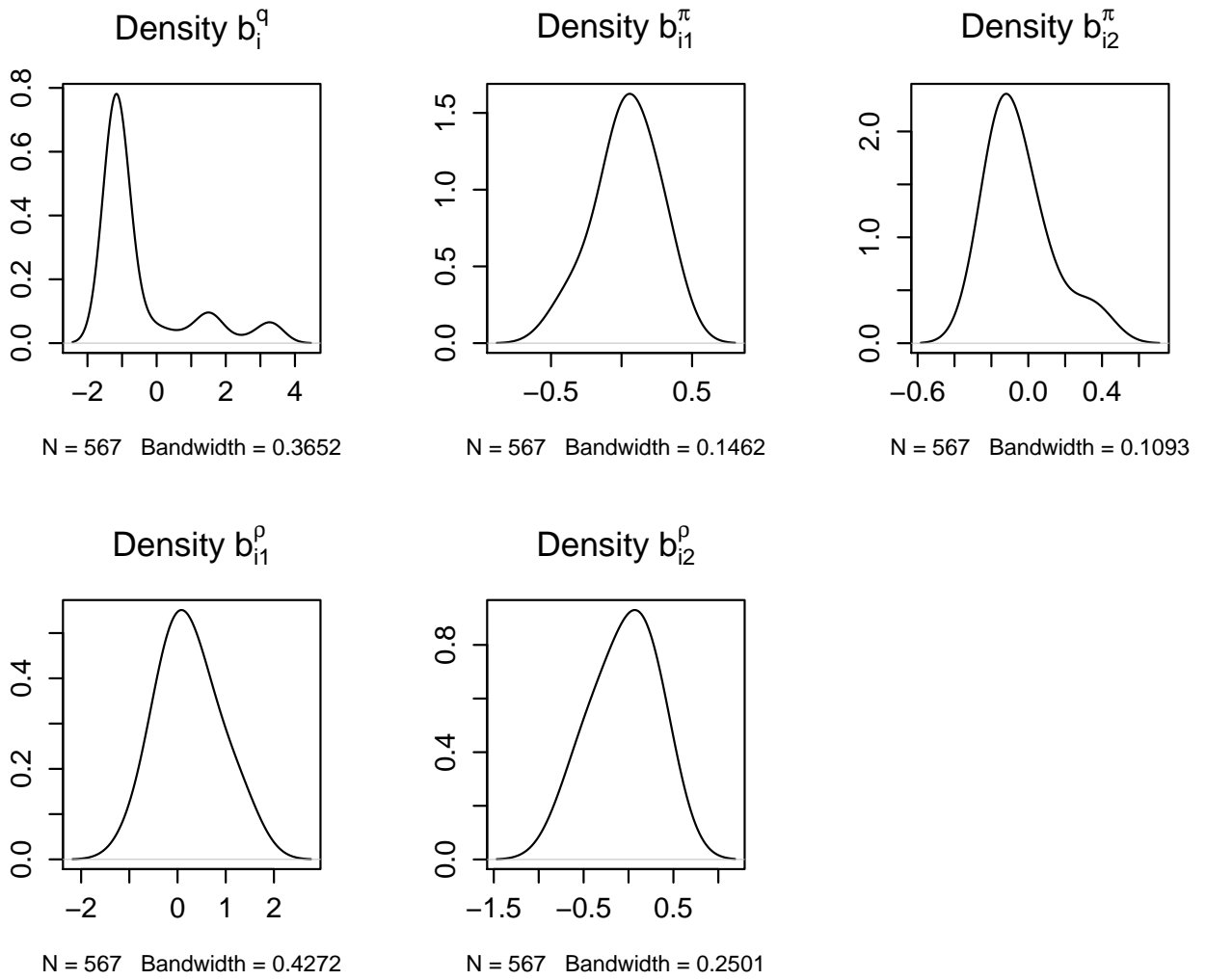
Figure 5: Density Plot for Random Effects

model and the physician prescription model. We estimate variance covariance structure between the random effects, using the MCMC samples. Our estimate of the variance-covariance matrix for the multivariate distribution for $\mathbf{b_i}$ is presented in Table 9.

Table 9: Empirical Variance-Covariance Matrix of Random Effects

|  | $b_i^q$ | $b_{i1}^\pi$ | $b_{i2}^\pi$ | $b_{i1}^\rho$ | $b_{i2}^\rho$ |
|---|---|---|---|---|---|
| $b_i^q$ | 2.097** | 0.054 | 0.210* | $-0.348$** | -0.037 |
| $b_{i1}^\pi$ | - | 0.108** | -0.007 | 0.153* | -0.007 |
| $b_{i2}^\pi$ | - | - | 0.091** | -0.004 | 0.054 |
| $b_{i1}^\rho$ | - | - | - | 0.635** | 0.202** |
| $b_{i2}^\rho$ | - | - | - | - | 0.272** |

The estimated variance covariance matrix indicates some significant correlations among physician-level unobserved factors that drive patient drug requests and those that drive physician prescriptions. This validates the necessity of a joint modeling approach. The positive covariance between $b_i^q$ and $b_{i2}^\pi$ suggests that a physician who tends to report more patient requests for Levitra also tends to report more drug requests in total. We find that the covariance between $b_i^q$ and $b_{i1}^\rho$ is negative and significant, which suggests that physicians who stick to the old drug (Viagra) are less likely to report patient drug requests in general. This might be because those physicians are less open-minded and care less about inputs from their patients; therefore, they either do not report their patient drug requests or do not create a comfortable environment for their patients to make drug requests. It is worth noting that the correlations of random effects from the brand-specific drug request model and physician drug prescription model are positive for both drugs, 0.153 for Viagra (significant at 90%) and 0.054 for Levitra. This means that there are some unobserved factors, e.g., certain disease condition, driving both the requests for Viagra and the prescriptions for Viagra relative to Cialis. If this positive interdependence is not accounted for as we do, the impact of drug requests on physician prescription decisions for Viagra may be overestimated. The estimated positive interdependence suggests that patients are more likely to request drugs that physicians are more likely to prescribe even after we control for some observable

patient characteristics, such as age, diagnosis level, and insurance status. In this sense, our empirical findings may suggest that patients are fairly consistent with physicians in making drug choices based on unobserved physician-level characteristics. The positive and significant covariance between $b_{i1}^{\rho}$ and $b_{i2}^{\rho}$ suggests that physicians who prefer Levitra to Cialis also tend to prefer Viagra to Cialis. Finally, a relative large value of variance of $b_i^q$ suggests a large heterogeneity of patient drug requests across physicians and is consistent with the data pattern that shows a large number of physicians never reported any drug requests.

# 6    Conclusion

Direct-to-consumer advertising by pharmaceutical companies is being argued to be effective to potential patients because it informs them of symptoms as well as availability of treatment options. Having better access to medical information, patients now are increasingly inclined to assert their preferences in health care via drug requests. Decision making by physicians on what drug to prescribe thus may become affected by drug requests from these informed patients as well as direct-to-physicians promotions. On the other hand, patients' decisions to request a particular drug might not be independent of their physicians' prescription decisions. These various factors form an interactive system in which the behaviors of interest are influenced by other factors, which in turn are influenced by the observed behaviors of physician reporting and a patient asking for a drug. Therefore, a unidimensional modeling approach with a narrower focus often fails to capture the full complexity of the situation and may produce an overly simplistic depiction of the behavior, which calls for a joint model to control for the interdependence of patient drug requests and physician prescriptions. In this paper we have developed a new zero-inflated multinomial (ZiMNL) joint model to account for the different sources of extra zeros and correlation between patient request and physician prescription behavior. The proposed joint model is flexible and new in separating the two sources of zeros and developing the interaction effect in a nonlinear setting.

Empirically, we mainly find that (1) the triggering of drug requests by DTCA is complicated with category level DTCA reducing patients' probability of making drug requests and drug specific DTCA driving drug requests for the advertised drug; (2) patient characteristics may play a role in the impact of DTCA on drug requests and the impact of patient requests on physician prescription decisions; (3) patient drug requests have a significant impact on physician prescription decisions and patients can be consistent with physicians in choosing a drug based on patient diagnosis level and some unobserved factors; (4) there are significant correlations among physician-level random effects that drive both patient drug requests and physician prescription decisions, which validates the joint modeling approach.

Our empirical findings have important managerial and public policy implications. For example, our study suggests that the overcompetition on DTCA may backfire the pharmaceutical firms and shrink the overall market (overall fewer requests for ED drugs), although a firm's heavy DTCA may help a firm to gain market share via patients' drug requests. Our finding also suggests that companies might benefit by focusing more on elderly patients, given that their drug requests are more likely to be accommodated by physicians and they are not less likely to be influenced by brand-specific DTCA than young patients. From a public policy perspective, patient requests for specific brand drugs represent an important right to participate in their medical decisions. Mechanisms that generate appropriate drug requests by minority patients will be helpful in reducing race/ethnicity disparity in health care because minority patient requests are more likely to be accommodated by physicians. Our empirical findings suggest that patients and physicians are fairly consistent in making drug choices based on observed patient diagnosis level and unobserved factors at the physician level, but are inconsistent in making drug choices based on patient age. On the other hand, the risk of worsening race/ethnicity disparity on health care is possible if minority patients' drug requests are inappropriate and physicians accommodate their requests for fear of being perceived as discriminatory or exacerbating relationships with patients. With these concerns, a public policy focusing on health providers aiming to reduce race/ethnicity disparity

probably should stick to a race/ethnicity blind to avoid possible undesirable outcomes.

A few limitations of our work must be underlined. One major issue is not having the physician demographic characteristics data, which limits our ability in modeling physician reporting behaviors. Another limitation is that we do not observe individual patient exposure to ED DTCA and have to rely on aggregate measurement of DTCA expenditure at the DMA level. With such micro-level data available, future research can better link the DTCA and patient drug requests. Notwithstanding these limitations, this research has developed a new road map for the analysis of patient requests and physician prescription behavioral data.

# References

Ai, C. & Norton, E. C. (2003), 'Interaction terms in logit and probit models', *Economics Letters* **80**, 123–129.

Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, *in* 'International Symposium on Information Theory', Springer New York, pp. 267–281.

Ansolabehere, S. & Iyengar, S. (1997), *Going Negative: How Political Ads Shrink and Polarize the Electorate*, The Free Press, New York.

Becker, G. & Newsom, E. (2003), 'Socioeconomic status and dissatisfaction with health care among chronically ill african americans', *Am J Public Health* **93(5)**, 742–748.

Bell, R. A., Wilkes, M. S. & Kravitz, R. L. (1999), 'Advertisement-induced prescription drug requests: Patients' anticipated reactions to a physician who refueses', *Journal of Family Practice* **48**, 446–452.

Belloni, A., Chen, D., Chernozhukov, V. & Hansen, C. (2012), 'Sparse models and methods for optimal instruments with an application to eminent domain', *Econometrica* **80**, 2369–2429.

Berndt, E. R., Bui, L., Reiley, D. R. & Urban, G. L. (1995), 'Information, marketing, and pricing in the U.S. antiulcer drug market', *The American Economic Review* **85**(2), 100–105.

Blendon, R. J., Scheck, A. C., K. Donelan, C. A. H., Smith, M., Beatrice, D. & Altman, D. (1995), 'How white and African Americans view their health and social problems. different experiences, different expectations', *J. Am. Med Assoc.* **273(4)**, 341–346.

Braun, M. & Bonfrer, A. (2011), 'Scalable inference of customer similarities from interactions data using dirichlet processes', *Marketing Sci.* **30(3)**, 513531.

Brooks, S. P. & Gelman, A. (1997), 'General methods for monitoring convergence of iterative simulations', *Journal of Compputational and Graphical Statistics* **7(4)**, 434–455.

Celeux, G., Forbes, F., Robert, C. P. & Titterington, D. M. (2006), 'Deviance information criteria for missing data models', *Bayesian Analysis* **1(4)**, 651–673.

Ching, A., Erdem, T. & Keane, M. (2009), 'The price consideration model of brand choice', *Journal of Applied Econometrics* **24(3)**, 393–420.

Ching, A. T. & Ishihara, M. (2010), 'The effects of detailing on prescribing decisions under quality uncertainty', *Quant. Marketing and Econom.* **8(2)**, 123–165.

Ching, A. T. & Ishihara, M. (2012), 'Measuring the informative and persuasive role of detailing on prescribing decisions', *Management Sci.* **58(7)**, 1374–1387.

D., S. (2004), *Statement of the American College of Physicians to the Senate Special Committee on Aging: The Hearing on the Impact of Direct-to-Consumer Advertising on Seniors Health and Health Care Costs.*

Dong, X. J., Manchanda, P. & Chintagunta, P. K. (2009), 'Quantifying the benefits of individual level targeting in the presence of firm strategic behavior', *J. of Marketing Res.* **46**, 207–221.

Ferguson, T. S. (1973), 'A Bayesian analysis of some nonparametric problems', *Annals of Statistics* **1(2)**, 209–230.

Ferguson, T. S. (1974), 'Prior distributions on spaces of probability measures', *Annals of Statistics* **2(4)**, 615–629.

Ghosh, P. & Albert, P. S. (2009), 'A Bayesian analysis for longitudinal semicontinuous data with an application to an acupuncture clinical trial', *Computational Statistics and Data Analysis* **53**, 699–706.

Gönül, F. F., Carter, F., Petrova, E. P. & Srinivasan, K. (2001), 'Promotion of prescription drugs and its impact on physicians' choice behavior', *J. Marketing* **65(3)**, 79–90.

Hatfield, L. A., Boye, M. E., Hackshaw, M. D. & Carlin, B. P. (2012), 'Multilevel Bayesian models for survival times and longitudinal patient-reported outcomes with many zeros', *J. Amer. Statist. Assoc.* **107**, 875–885.

Hibbard, J. H., Greene, J. & Overton, V. (2013), 'Patients with lower activation associated with higher costs; delivery systems should know their patients' 'scores'', *Health Affairs* **32(2)**, 216–222.

Hoek, J. & Gendall, P. (2003), 'Direct to consumer advertising of prescription medicines: A consumer survey', *Palmerston North: Department of Marketing, Massey University* .

James, L. D. & Vila, B. (2014), 'Racial and ethnic bias in decisions to shoot seen through a stronger lens: Experimental results from high-fidelity laboratory simulations', *Journal of Experimental Criminology* **10(3)**, 323–340.

Jiang, B., Elliott, M. R., Sammel, M. D. & Wang, N. (2015), 'Joint modeling of cross-sectional health outcomes and longitudinal predictors via mixtures of means and variances', *Biometrics* **71(2)**, 487–497.

Lachenbruch, P. A. (2002), 'Analysis of data with excess zeros', *Statictical Methods in Medical Research* **11**, 297–302.

Lee, D. & Beqley, C. E. (2011), 'Physician switching after drug request refusal', *Health Marketing Quarterly* **28(4)**, 304–316.

Li, Y. & Ansari, A. (2014), 'A Bayesian semiparametric approach for endogeneity and heterogeneity in choice models', *Management Science* **60(5)**, 1161–1179.

Liu, Q. & Gupta, S. (2011), 'Understanding the impact of direct-to-consumer advertising on

patients? behavior and the public policy implication', *International Journal of Research in Marketing* **28**, 205–217.

Liu, Q., Gupta, S., Venkataraman, S. & Liu, H. (2016), 'An empirical model of drug detailing: Dynamic competition and policy implications', *Management Sci.* **62(8)**, 2321–2340.

Liu, Q., Steenburgh, T. & Gupta, S. (2015), 'The cross attributes flexible substitution logit: Uncovering category expansion and share impacts of marketing instruments', *Marketing Sci.* **34(1)**, 144–159.

Lunn, D. J. (2003), 'WinBUGS Development Interface (WBDev)', *ISBA Bulletin* **10(3)**, 10–11.

Montagna, S., Tokdar, S. T., Neelon, B. & Dunson, D. B. (2012), 'Bayesian Latent Factor Regression for Functional and Longitudinal Data', *Biometrics* **68**(4), 1064–1073.

Neelon, B., Ghosh, P. & Loebs, P. F. (2013), 'A spatial Poisson hurdle model for exploring geographic variation in emergency department visits', *Journal of the Royal Statistical Society. Series A: Statistics in Society* **176**(2), 389–413.

Nerlove, M. & Arrow, K. J. (1962), 'Optimal advertising policy under dynamic conditions', *Economica* **29(May)**, 129–142.

NIH Consensus Conference (1993), 'NIH consensus development panel on impotence', *Journal of the American Medical Association* **270**, 83–90.

Norton, E. C., Wang, H. & Ai, C. (2004), 'Computing interaction effects and standard errors in logit and probit models', *The Stata Journal* **4(2)**, 154–167.

Peck, B. M. (2011), 'Age-related differences in doctor-patient interaction and patient satisfaction', *Current Gerontology and Geriatrics Research* **2011**, 1–10.

Petrin, A. & Train, K. (2010), 'A control function approach to endogeneity in consumer choice models', *J. of Marketing Res.* **47(1)**, 3–13.

Rizopoulos, D., Verbeke, G. & Molenberghs, G. (2008), 'Shared parameter models under random effects misspecification', *Biometrika* **95(1)**, 63–74.

Sethuraman, J. (1994), 'A constructive definition of dirichlet priors', *Statistica Sinica* **4**, 639–650.

Spiegelhalter, D. J., Abrams, K. R. & Myles, J. P. (2004), *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*, Wiley.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van Der Linde, A. (2002), 'Bayesian measures of model complexity and fit', *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **64**(4), 583–616.

Staiger, D. & Stock, J. (1997), 'Instrumental variables regression with weak instruments', *Econometrica* **65(3)**, 557–586.

Stevenson, F. A., Barry, A. C., Britten, N., Barber, N. & Bradley, C. P. (2000), 'Doctor-patient communication about drugs: The evidence for shared decision making', *Soc. Sci. Medicine* **50**, 829–840.

Sturtz, S., Ligges, U. & Gelman, A. (2005), 'R2WinBUGS: a package for running WinBUGS from R', *Journal of Statistical Software* **12**, 1–16.

Voleti, S., Kopalle, P. & Ghosh, P. (2015), 'An interproduct competition model incorporating branding hierarchy and product similarities using store-level data', *Management Sci.* **61(11)**, 27202738.

Wooldridge, J. M. (2015), *Introductory Econometrics: A Modern Approach*, Cengage Learning; 5th edition, Mason, OH.

Yang, M. (2010), 'Semiparametric Bayes hierarchical models with mean and variance constraints', *Computational Statistics and Data Analysis* **54**, 2172–2186.

Yi, J. C. (2008), 'An expert system to derive carryover effect for pharmaceutical sales detailing optimization', *Expert Systems with Applications* **34**(3), 1742–1753.
  **URL:** *http://www.sciencedirect.com/science/article/pii/S0957417407000516*

Zhang, M., Strawderman, R. L., Cowen, M. E. & Wells, M. T. (2006), 'Bayesian inference for a two-part hierarchical model: An application to profiling providers in managed health care', *J. Amer. Statist. Assoc.* **101**, 934–945.

Zou, H. & Hastie, T. (2005), 'Regularization and variable selection via the elastic net', *Journal of the Royal Statistical Society B* **67**, 301–320.