Quant Mark Econ (2011) 9:301–337
DOI 10.1007/s11129-011-9102-7

A new multivariate count data model to study multi-category physician prescription behavior

Xiaojing Dong · Pradeep K. Chintagunta · Puneet Manchanda

Received: 24 September 2010 / Accepted: 10 May 2011 / Published online: 26 May 2011

Abstract Multivariate count models represent a natural way of accommodating data from multiple product categories when the dependent variable in each category is represented by a positive integer. In this paper, we propose a new simultaneous equation multi-category count data model–the Poisson-lognormal simultaneous equation model–that allows for the Poisson parameter in one equation to be a function of the Poisson parameters in other equations. While generally applicable to any situation where simultaneity is an issue and the dependent variables are measured as counts, such a specification is particularly useful for our empirical application where physicians prescribe drugs in multiple categories. Accounting for the endogeneity of detailing in such situations requires us to explicitly allow for pharmaceutical firms' detailing activities in one category to be influenced by their activities in other categories. Estimation of such a system of equations using traditional maximum likelihood method is cumbersome, so we propose a simple solution based on using Markov Chain Monte Carlo methods. Our simulation study demonstrates the validity of the solution algorithm and the biases that would result if such simultaneity is ignored in the estimation process.

The authors would like to thank ImpactRx for providing the data, Michael Li, Sridhar Narayanan, Nitin Mehta, Gunter Hitsch, Renna Jiang and V. "Seenu" Srinivasan for their helpful suggestions. Dong would like to acknowledge financial support from the Presidential Research Grant awarded by Santa Clara University and Chintagunta thanks the Kilts Center for Marketing at the Booth School of Business for financial support. All errors are the authors'

X. Dong (✉)
Leavey School of Business, Santa Clara University, Santa Clara, CA, USA
e-mail: xdong1@scu.edu

P. K. Chintagunta
Booth School of Business, University of Chicago, Chicago, IL, USA
e-mail: pradeep.chintagunta@chicagobooth.edu

P. Manchanda
Ross School of Business, University of Michigan, Ann Arbor, MI, USA
e-mail: pmanchan@umich.edu

We apply our methodology to study multi-category physician prescription behavior, while accounting for the endogeneity and simultaneity of firms' detailing efforts within and across categories, at individual physician level. Substantively, we show that detailing responsiveness estimates, as well as their implications for physician segmentation and firms' profits are significantly affected when we leverage data from multiple categories to account for endogeneity in detailing decisions.

Keywords Simultaneous Poisson model · Cross-category · MCMC · Endogeneity

JEL Classification C30 · C13 · C51 · M31 · C11

1 Introduction

There has been considerable research in recent years that has focused on understanding the cross-category behavior of consumers across a wide range of contexts (see for example the surveys by Russell et al. (1997), Russell et al. (1999), Elrod et al. (2002), and Seetharaman et al. (2005) as well as the studies by Erdem (1998) on umbrella branding; Manchanda et al. (1999) on the shopping basket, etc.). There are several substantive reasons motivating this line of work including the understanding of consumer response to marketing activity across categories (Ainslie and Rossi 1998), understanding the effect of marketing activities in one category on behavior in other categories (Erdem and Sun 2002); pooling information across categories to better pin down consumer heterogeneity (Hansen et al. 2006); providing a useful input for decision makers interested in allocating resources across categories (Gupta and Steenburgh 2008), etc. An important building block in this research is a demand model that characterizes the behavior of consumers, or more generally, of agents across the multiple categories of interest.

A vast majority of the above research uses a discrete choice framework for the underlying consumer demand model. Researchers have also demonstrated how one can accommodate unobserved sources of heterogeneity as well as endogeneity stemming from the presence of unobservable demand characteristics that could be correlated with included marketing activities such as prices and promotions (Song and Chintagunta 2006; Ma et al. 2009). For estimation, these studies either use maximum likelihood methods (with a control function approach when endogeneity is a concern; Ma et al. 2009) with household-level data or generalized method-of-moments based estimators with aggregate data and a BLP-type (Berry et al. 1995) approach to account for endogeneity in regressors.

However, there are many empirical situations where researchers may be unwilling or unable to use a discrete choice model as the underlying consumer demand function. Examples of such situations are number of web pages visited in different categories of websites (Li et al. 2009), grocery purchases in multiple categories (Brijs et al. 1999) and health care usage in terms of ER visits and hospitalizations (Munkin and Trivedi 1999). In this paper, we are interested in one such situation—the prescription behavior of physicians across multiple therapeutic categories over time. Like previous multi-category research, our objective is simple–to understand

the impact of marketing activities by firms across categories on prescription behavior of physicians. Since our outcome of interest is the monthly number of prescriptions written by the physicians, a multivariate count data model is appropriate as the "demand" specification. In addition to this, it is also important to account for endogeneity of firms' marketing resource allocation in the pharmaceutical context, especially as these firms often have to allocate such resources across drugs in multiple categories (Sinha and Zoltners 2001). Thus, the level of detailing (the marketing resource we study) in one category is likely to influence the level of detailing to the physician for other categories.

In order to accommodate all these features of the institutional setting and the data, the ideal model would be a simultaneous equation count (Poisson) model with equations for prescriptions of each drug as well detailing for each drug. Besides the fact that all these count variables are correlated, the inclusion of the detailing equations leads to a model where the Poisson parameters in some equations (in our case the ones for detailing from the company's marketing in multiple categories) are functions of the Poisson parameters in other equations (detailing in other categories from the same companies) and of other exogenous variables and their associated effects (the "structural" parameters). In other words, this model (comprising a simultaneous system of Poisson equations) is a general version of the multivariate count models proposed in the existing literature.

The general strategy followed in the literature for dealing with a system of count variables has been to obtain the joint model by deriving and solving the marginal distributions. For example, King (1989) uses a Seemingly Unrelated Poisson Regression model while Munkin and Trivedi (1999) use a Multivariate Poisson Regression model. For the general model proposed here, this is not feasible due to the simultaneity feature (where Possion parameters are direct functions of each other). We propose a solution for this general class of models using Markov Chain Monte Carlo methods (Gelfand and Smith 1990). Specifically, we first solve the system of simultaneous Poisson parameters to obtain a transformed set of Poisson parameters wherein each Poisson parameter is a function only of the exogenous variables and the structural parameters of the model. We then use these transformed Poisson parameters to characterize the likelihood function. In estimation, we simulate the structural parameters of the model, and use these in conjunction with the exogenous variables to obtain the transformed Poisson parameters and then, the likelihood function. Parameter identification depends on the number of exogenous and endogenous variables in the system and an associated order condition that needs to be satisfied. A simulation study demonstrates the ability of our estimation approach to recover the model parameters.

We then use both prescription and detailing data from two large pharmaceutical categories (with one firm selling drugs in both categories) to estimate our model. We also demonstrate how ignoring the simultaneity in the estimation could impact the estimated parameters of the system and lead to biased detailing elasticities; and how this affects implications for response segmentation and firm profits.

In sum, our paper makes the following contributions. First, we generalize the literature on multivariate count data models by specifying a multivariate count model where the Poisson parameter in one equation depends on the Poisson parameters from other equations. We then propose an approach to estimate such a model. Our

approach is straightforward in that it needs no additional parameters and uses the standard likelihood function. This approach can be applied to other situations where multivariate count data are prevalent and where endogeneity concerns may exist. Next, we take a step towards adding to the growing body of literature that quantifies the effects of marketing (especially detailing) to physicians by leveraging information on both the response and the levels of marketing activities *across* categories. We use our results to provide substantive insights for the pharmaceutical firm, especially with respect to firm's segmentation strategies and their profit consequences.

2 Institutional setting, model and econometric specification

2.1 Institutional setting

Pharmaceutical companies that market drugs across multiple therapeutic classes might be interested in understanding the impact of their marketing activities on prescriptions for their entire portfolio of medicines at the level of the individual physician. If physician prescription and firm detailing behavior are independent across categories, then pharmaceutical companies can simply repeat the type of physician level analysis carried out for single product categories across each of the categories in their portfolios.

However, such an independent category-by-category level analysis could lead to erroneous conclusions regarding the sensitivities of physicians to detailing activities. There are three reasons for this. First, from the demand side, one would expect a physician's prescriptions across categories to be correlated due to multiple factors. One might expect a correlation in a physician's responsiveness to detailing across categories—a physician who is responsive to detailing in one category could also be responsive in other categories (Ainslie and Rossi (1998) and Chib et al. (2002) document this for consumers in the packaged goods category). Further, one would expect the preferences for brands; and the detailing sensitivities for the brands marketed by the *same* firm in these categories to be correlated (e.g., Erdem 1998). Additionally, a physician's prescriptions *across* categories can also be correlated due to unobserved (by the researcher) factors that drive these prescriptions. This latter set of factors could include patient co-morbidities across therapeutic classes for patients treated by a physician.[1] Correlation in prescription behavior across therapeutic classes, by itself, is not of concern when studying the sensitivities of physicians to detailing across categories (Ainslie and Rossi 1998). However, in conjunction with the other factors below, their presence could be problematic when carrying out single category analyses.

Second, when a firm that markets drugs across a number of therapeutic categories sets its detailing levels for its drugs in these categories, it is likely that the level of detailing for a drug in one category is influenced either positively or negatively by the levels of detailing chosen for the drugs in other categories. For example, there is

[1] "Comorbidity" describes the effect of all other diseases an individual patient might have other than the primary disease of interest.

usually a constraint on the total number of details that can be executed by the sales force resulting in the number of details in one category (e.g., an older drug) being influenced by those in another (e.g., a newer drug). In this case, the across category influence of detailing decisions is likely to be negative. On the other hand, certain PCPs who write more prescriptions across several categories could receive a larger number of details in all categories as opposed to others who write fewer prescriptions in these categories. This could result in a positive correlation in detailing decisions across categories. In an econometric sense, this would result in simultaneity in detailing decisions across categories for firms marketing drugs in those categories. In accounting for this simultaneity, it is important to accommodate empirical situations such as ours where we observe data only for a *subset* of categories in which the firms in our data market their products.

Third, in addition to correlation in prescription behavior across categories, other correlations can also play a role in the final prescription outcomes across therapeutic classes. For a single therapeutic class, we know that there are many unobservables that affect detailing (see Jones 2002). Sometimes, these unobservables are also correlated across categories. For example, if there are changes in the regulatory environment around pharmaceutical marketing–say the relaxation on advertising restrictions by the FDA in 1997—detailing across categories may be impacted. These correlations could also arise *within* a detail when the detailer is attempting to provide information on multiple categories to a physician. Further, as noted in previous research (e.g., Dong et al. 2009), these unobservables can also be correlated with those from the prescription decision discussed above and this correlation needs to be accounted for in the analysis. Thus there is a possibility that prescription behavior and detailing activity are correlated within and across categories because of both observables and unobservables and the simultaneity of detailing multiple drugs (across these categories). Ignoring these relationships across categories can lead to erroneous conclusions regarding the sensitivities of physicians to detailing activities.

2.2 The demand (prescription) model

We begin by noting that while our formulation is general, we describe it in the context of our empirical application to fix ideas. As the number of prescriptions written by a physician in a given time period is a nonnegative integer, we use a count model (Cameron and Trivedi 1998) to model physician prescription behavior. Let rx_{ijt} represent the number of prescriptions written by physician i, i=1,...,I, for brand j, j=1,..., J, in time period t, t=1,...,T. Let λ_{ijt} denote the mean number of prescriptions written by physician i for brand j in time period t. Then, given a Poisson distribution with parameter λ_{ijt}, we have:

$$P\left(rx_{ijt} = m \middle| \lambda_{ijt}\right) = \frac{\lambda_{ijt}^{m} \exp\left(-\lambda_{ijt}\right)}{m!}$$

We use the log-link function to account for the effect of covariates on the mean prescription rate:

$$\lambda_{ijt} = \exp\left(u_{ijt}\right) \tag{1}$$

u_{ijt} contains observed and unobserved factors that impact physicians' prescription decisions:

$$u_{ijt} = \beta_{ij0} + \frac{\beta_{ijj}}{1 + dtl_{ijt}} + \sum_{k \neq j} \beta_{ijk} dtl_{ikt} + \beta_{ijl} log\left(rx_{ij,t-1} + 1\right) + \in_{ijt} \quad (2)$$

The intercept β_{ij0} accounts for a physician-brand specific effect. To accommodate the presence of nonlinear effects of own detailing on prescriptions via β_{ijj}, we adopt a flexible log-reciprocal transformation for detailing dtl_{ijt} (Lilien et al. (1995), p. 656) that allows for concave, convex or S-shaped detailing effects. We add 1 to the denominator to include the cases when $dtl_{ijt} = 0$ in the data.[2] The model also incorporates the effects of competitive detailing in the category (β_{ijk}), with these effects differing across competitors as well as across brands. We use different specifications for the own and competitive effects as the linear specification of the competitive effects gave us the best model fit. $log(rx_{ij,t-1}+1)$ is the log-transformation of the number of prescriptions dispensed by the physician in the previous time period, and β_{ijl} is the associated impact on prescriptions. This variable is included in order to account for state dependence in a physician's prescription behavior, and for carry-over effects of detailing (e.g., Manchanda et al. 2004). Finally, $\in_{ijt}$ is an additive random error term, accounting for other physician-brand-time specific factors that could impact a physician's prescription decisions but are not observed or measurable by researchers. For example, previous literature (e.g., Wosinska 2002) has documented that direct to consumer advertising (DTCA) could impact prescriptions dispensed by physicians. As these effects are likely to be idiosyncratic, they are captured in $\in_{ijt}$.

A firm's detailing decisions (dtl_{ijt}) could depend on its knowledge of the unobservable (to the researcher) factors described above, leading to correlations between the observed dtl_{ijt} and unobserved factors $\in_{ijt}$ in Eq. (2). This endogeneity in the dtl_{ijt} variable needs to be accounted for in order to obtain unbiased parameters for the prescription model. Industry practice and previous research has documented that firms collect and use data on each physician before deciding on the amount of detailing targeted at each physician (Manchanda et al. 2004; Dong et al. 2009 and Matta and Mehta 2001). Hence we include a detailing model to account for the endogeneity of detailing.

2.3 The firm behavior (detailing) model

Given the multi-category nature of our analysis, the underlying behavioral mechanism driving the level of details for a firm across therapeutic categories is unclear. Rather than impose a specific behavioral assumption on firm behavior, we use a limited information approach to accounting for endogeneity (in contrast with the full information approach as in Dong et al. 2009). Specifically, we assume that the number of detailing visits to physician i for a drug j in time period t, dtl_{ijt} follows a Poisson distribution (cf. Manchanda et al. 2004). The parameter of the

[2] About 25% of the prescriptions in the data are 0, and about 31% of the detailing visits are 0 in the data.

conditional Poisson distribution is represented by γ_{ijt}, which is specified as a log-link function below.[3]

$$\begin{aligned} log(\gamma_{ijt}) = {} & \widetilde{\alpha_{j0}} + \alpha_{j1}\beta_{ij0} + \alpha_{j2}\beta_{ijj} \\ & + \alpha_{jn} log(1 + dtl_{ij,t-1}) + \sum_{k \neq j} \alpha_{jk} log(1 + dtl_{ik,t-1}) + \alpha_{jl} log(1 + rx_{ij,t-1}) \\ & + \alpha_{jw} W_{it} + \sum_{g \neq j, g \in F} \alpha_{jg} log(\gamma_{igt}) + \widetilde{\eta_{ijt}} \end{aligned} \tag{3}$$

$\widetilde{\alpha_{j0}}$ represents a brand specific average detailing intensity level. There are four sets of explanatory variables. These are described in turn. The first set is based on heuristics used by managers resulting in non-random marketing mix settings and is incorporated as in Manchanda et al. (2004). It includes the effects of β_{ij0} and β_{ijj}, which are the physician specific intercept and own-detailing response parameters from the prescription model (see Eq. (2)). A potential drawback with including only these factors as drivers of detailing levels is that while they are all physician-brand specific, they are time invariant–thus they cannot fully explain observed detailing levels that vary over time.

The second set of variables that we incorporate are time varying. They are lagged own and competitive detailing decisions targeted to the same physician as well as the prescriptions written by that physician in the previous time period (via a logarithmic transformation). These variables ($log(1+dtl_{ij,\ t-1})$, $log(1+dtl_{ik,\ t-1})$, and $log(1+rx_{ij,\ t-1})$) reflect the dependence of current period detailing levels on previous own and competitive decisions (Leeflang and Wittink 1992, 1996). Including all these time-varying factors is consistent with the "limited information" approach to accounting for endogeneity (e.g., Nair 2007). Inclusion of these variables can also be construed loosely as an attempt to approximate, via "reduced form," a more structural model of detailing based on firms' objective functions for these categories. The third set of explanatory variables consists of exogenous variables, i.e., variables that could shift the detailing levels without directly affecting physicians' prescription levels. These variables are required for identification purposes and are denoted as W_{it} as in the detailing model.

When a company markets drugs across categories, we explicitly incorporate the simultaneity in detailing decisions across those categories. Thus another factor that could impact the firm's detailing decision for a specific brand at the physician level is the firm's detailing efforts in marketing other categories to the same physician. There are two ways of including this information–using the realized number of details, ($dtl_{igt,\ g\neq j}$), or the expected number of details ($\gamma_{igt,\ g\neq j}$). For our empirical application, the latter specification is more appropriate since at the time of setting their details in one category, managers may not be able to condition on the realized details in the other categories. This can come about because of organizational issues (managers in different categories set details in

[3] The function we use is based on exploratory analyses of several alternative specifications. We tried 3 functional forms—a linear specification, a diminishing returns specification (via the log-transform) and an S-shaped functional form (via a log-reciprocal transform). We also tried some combinations of these functional forms across the various terms in Eq. (4). The final specification was chosen based on the model that provided the best fit to the data.

their own categories with these numbers adjusted subsequently by the sales force coordinator) or implementation issues (due to unobserved within detail factors that require the salesperson to make adjustments to the calling plan). So the effect of detailing in the other categories is represented by the effects of the Poisson parameters of all the other drugs (γ_{igt}) marketed by the firm marketing the drug j, $\Sigma_{g\neq j, g\in F}\alpha_{jg}log(\gamma_{igt})$, where g represents each of the other brands (other than brand j) marketed by the same firm F.

Finally, $\widetilde{\eta_{ijt}}$ in Eq. (3) denotes the idiosyncratic random shocks, representing other unobserved factors. An example of such a factor is a situation where a physician learns about a drug through sources not in our data (e.g., scientific journals or peers) and therefore would be more welcoming to detailers–in order to obtain more information about the drug—for a certain time period.

2.3.1 Accommodating information from categories for which we do not have data

In most empirical applications the analyst usually does not have data on all categories in which the companies market products and will also not have data on all the companies that compete in the various categories. To accommodate this feature of the data, let us first divide the set of products marketed by firm F into 2 groups–*F1* for which we have data and *F2* for which data are not available. Then the term $\Sigma_{g\neq j, g\in F}\alpha_{jg}log(\gamma_{igt})$ in Eq. (4) can be written as:

$$\sum_{g\neq j,\ g\in F}\alpha_{jg}log\left(\gamma_{igt}\right)=\sum_{g1\neq j, g1\in F1}\alpha_{jg1}log\left(\gamma_{ig1t}\right)+\sum_{g2\neq j, g2\in F2}\alpha_{jg2}log\left(\gamma_{ig2t}\right)$$

Let $\Sigma_{g2\neq j, g2\in F2}\alpha_{jg2}log(\gamma_{ig2t}) = m_j+n_{ijt}$; where m_j is the expected value of this factor across time and physicians and n_{ijt} is the residual or deviation from the mean. Then, the detailing model is:

$$\begin{aligned}log\left(\gamma_{ijt}\right) &= \alpha_{j0}+\alpha_{j1}\beta_{ij0}+\alpha_{j2}\beta_{ijj}\\ &+\alpha_{jn}log\left(1+dtl_{ij,t-1}\right)+\sum_{k\neq j}\alpha_{jk}log\left(1+dtl_{ik,t-1}\right)+\alpha_{jl}log\left(1+rx_{ij,t-1}\right)\\ &+a_{jw}W_{it}+\sum_{g1\neq j, g1\in F1}\alpha_{jg1}log\left(\gamma_{ig1t}\right)+\eta_{ijt}\end{aligned} \tag{4}$$

Where:

$$\alpha_{j0}=\widetilde{\alpha_{j0}}+m_j \tag{5}$$

$$\eta_{ijt}=\widetilde{\eta_{ijt}}+n_{ijt} \tag{6}$$

If firms are setting detailing levels across categories and if we have data only on a subset of categories, then γ_{ig1t}, g1∈*F*1 will be correlated with n_{ijt} and hence η_{ijt}. By

allowing for correlations across all η_{ijt} as described below, we account for this source of endogeneity as well.

2.4 Heterogeneity distributions & correlation structures across categories

To complete the specification of the demand (prescription) and the firm behavior (detailing) models, we need to describe our assumptions on β_{ij0}, β_{ijj}, β_{ijk}, β_{ijl}, $\in_{ijt}$, and η_{ijt}. We assume a hierarchical structure for β_{ij0} as:

$$\beta_{ij0} = \theta_{j0} + Z_i\theta_{j1} + \delta_{iF0}I_{j\in F1} + v_{ij}^0, \forall j \tag{7}$$

where θ_{j0} and θ_{j1} are parameters to be estimated, Z_i's represent cross-sectional differences across physicians that affect the mean level of prescriptions; $I_{j\in F1}$ is an indicator variable that takes the value 1 when j belongs to the set *F1* of the observed brands whose firm F also market products in other categories and takes the value 0 otherwise, with δ_{iF0}[4] representing an additional error component that "switches on" only for such brands; $\delta_{iF0} \sim N(0, S_{\delta F0})$; and v_{ij}^0 is a random variable following a Multivariate Normal distribution with zero mean and unknown covariance matrix $\sum_v^0$. The dimension of $\sum_v^0$ is the same as the number of brands within *each* category.[5] With a non-diagonal $\sum_v^0$, we allow for intrinsic brand preferences to be correlated across brands in the same category. Further, with non-zero δ_{i0} we allow for intrinsic brand preferences of products marketed by the same firm *across* categories to also be correlated. However if a firm only markets products in 1 category, this effect would drop out.

The parameter β_{ijj} is specified, using an error component structure, as

$$\beta_{ijj} = \kappa_i + \delta_{iF}I_{j\in F1} + \widetilde{\beta}_{ijj} + v_{ij}^1, \forall j \tag{8}$$

In the above equation, κ_i refers to an error component that is common across all brands j and categories c such that $\kappa_i \sim N(0,s_k)$ and s_κ is an unknown parameter to be estimated. This is one of the two components that capture possible cross-category correlations in the detailing response parameter, which is due to the same individual physician prescribing multiple categories. $I_{j\in F1}$ is the indicator variable as defined previously with δ_{iF} representing an additional error component that switches on only for brands in the set *F1,* $\delta_{iF} \sim N(0,s_{\delta F})$. This is the other component that captures cross-category correlations, due to the same firm marketing across categories. $\widetilde{\beta}_{ijj}$ is a brand-specific component such that for all brands j and k within category c, we assume

$$\left\{\widetilde{\beta}_{ijj}, \beta_{ijk}, \beta_{ijl}\right\} \sim N\left(\overline{\beta}, \Sigma_\beta\right) \tag{9}$$

where $\overline{\beta} = \left\{\beta_{jj}, \beta_{jk}, \beta_{jl}\right\}$ denotes the mean vector and Σ_β represents the non-diagonal covariance matrix. Finally, we allow the random variable in this formulation v_{ij}^1, and the one in the decomposition of the intercepts v_{ij}^0, to jointly follow a Multivariate Normal distribution $\left\{v_{ij}^0, v_{ij}^1\right\} \sim N(0, \Sigma_v)$. With a non-diagonal Σ_v, we allow for intrinsic brand preferences to be correlated with

[4] Here we use the index for the firm and the set of products marketed by the firm interchangeably.

[5] Note that as we use δ_{iF0} to explicitly capture the cross-category correlations in the intercepts, all the other components in equation (7) are allowed to be correlated within each category, but not across categories.

responsiveness to detailing across brands in the same category. As the cross-category effects are explicitly captured by the other two components (κ_i and $\delta_{iF}I_{j\in F1}$), Σ_β and Σ_ν only allow correlations across brands in the same category, but not across-categories.

With such a formulation, we are able to incorporate the following features into our model specification.

(a) The presence of κ_i allows for some physicians to be more responsive than other physicians for all brands (across all categories), because of their experience, training and/or some other unobserved individual specific characteristics (as in Ainslie and Rossi 1998).
(b) Next, $\widetilde{\beta}_{ijj}$ allows for a physician's responsiveness across brands *within* a category to be correlated over and above the common component across categories, κ_i. Further, it allows for correlation with all the other response parameters within the category.
(c) The inclusion of δ_{iF} provides the basis for additional correlation across physician detailing parameters for brands *marketed by the same company*. Let *b* and *b*' (that belong to set F1) index the two brands marketed by a firm in two different product categories (as in our empirical application). We allow the intrinsic brand preferences and the detailing sensitivity parameters to be correlated across these two brands via this error component.

Next, we turn to the specification of the random error terms from the prescription to the detailing models $\in_{ijt}$ and η_{ijt}. Each of these error terms is decomposed into two components: a category specific component and a brand specific component as follows:

$$\in_{ijt} = e_{ict} + f_{ijt}, \text{where } j \in c \tag{10}$$

Thus, c indexes the category that brand j belongs to, and e_{ict} is the category specific random component, which is common to all brands in the same category c, and f_{ijt} indicates the brand specific random component, which captures the rest of the randomness in e_{ijt}. The same framework is adopted for the random factors in the detailing model, that is

$$\eta_{ijt} = \nu_{ict} + \gamma_{ijt}, \text{where } j \in c \tag{11}$$

Further, we assume that the category-specific random factors from both models (e_{ict} and u_{ict}) jointly follow a Multivariate Normal distribution, with dimension $\overline{C}$: the number of categories times 2 (for both prescription and detailing models).

$$\left\{ \begin{pmatrix} e_{ict} \\ u_{ict} \end{pmatrix} | c = 1, 2 \right\} \sim MVN\left(0, \Sigma^{\overline{C}}\right) \tag{12}$$

In our empirical application, as there are two categories in both prescription and the detailing models, $\Sigma^{\overline{C}}$ has dimension of 4×4, to capture the correlations across categories and across the two models.

Second, the brand-specific random factors within each category, across both models (prescription and detailing models) also jointly follow a Multivariate

Normal distribution, with dimension equal to the number of brands in each category times 2.

$$\left\{ \begin{pmatrix} f_{ijt} \\ v_{ijt} \end{pmatrix} | \, j = 1,2,3,4 \right\} \sim MVN\left(0, \Sigma^{\overline{JC}}\right) \tag{13}$$

In our case as there are four brands in each category and for each of the two (prescription and detailing) models, $\Sigma^{\overline{JC}}$ has dimension 8×8, to capture the correlations across brands within category as well as across the prescription and detailing models. We obtain two covariance matrices $\Sigma^{\overline{JC}}$, one for each category. Finally, there is no covariance between the two error components: the category-specific and the brand-specific components.

$$cov\left(e_{ict}, f_{ijt}\right) = 0 \; and \; cov\left(u_{ict}, v_{ijt}\right) = 0, \forall c \; and \; j \tag{14}$$

To demonstrate the variance-covariance structure of this setup, we create a simple example with two categories and two brands in each category. To distinguish indexes of brands and categories, we use numbers to denote brands j =1,2,3,4, and letters *A and B* to denote the two categories. Let the first two brands j=1,2 belong to category A; and j=3,4 belong to category B. Let subscript r indicate the prescription model, and subscript d indicate the detailing model. The covariance matrix of the category specific error components for both prescription and detailing models $\Sigma^{\overline{C}}$ can be written as

$$\sum^{\overline{C}} = \begin{array}{c} \begin{matrix} \text{Prescription Model} & & \text{Detailing Model} & \end{matrix} \\ \begin{matrix} \text{CategoryA} & \text{CategoryB} & \text{CategoryA} & \text{CategoryB} \end{matrix} \\ \begin{bmatrix} \sigma A_r A_r & \sigma A_r B_r & \sigma A_r A_d & \sigma A_r B_d \\ \sigma A_r B_r & \sigma B_r B_r & \sigma B_r A_d & \sigma B_r B_d \\ \sigma A_r A_d & \sigma B_r A_d & \sigma A_d A_d & \sigma A_d B_d \\ \sigma A_r B_d & \sigma B_r B_d & \sigma A_d B_d & \sigma B_d B_d \end{bmatrix} \end{array}$$

The first two rows and two columns are for the two categories in the prescription model, and the last two rows and two columns are for the two categories in the detailing model.

We use two covariance matrices to represent the covariances among brand-specific random error terms $\Sigma^{\overline{JA}}$ for category A and $\Sigma^{\overline{JB}}$ for category B, which are specified in the following.

$$\sum^{\overline{JA}} = \begin{bmatrix} \sigma^A_{1_r 1_r} & \sigma^A_{1_r 2_r} & \sigma^A_{1_r 1_d} & \sigma^A_{1_r 2_d} \\ \sigma^A_{1_r 2_r} & \sigma^A_{2_r 2_r} & \sigma^A_{2_r 1_d} & \sigma^A_{2_r 2_d} \\ \sigma^A_{1_r 1_d} & \sigma^A_{2_r 1_d} & \sigma^A_{1_d 1_d} & \sigma^A_{1_d 2_d} \\ \sigma^A_{1_r 2_d} & \sigma^A_{2_r 2_d} & \sigma^A_{1_d 2_d} & \sigma^A_{2_d 2_d} \end{bmatrix} \qquad \sum^{\overline{JB}} = \begin{bmatrix} \sigma^B_{3_r 3_r} & \sigma^B_{3_r 4_r} & \sigma^B_{3_r 3_d} & \sigma^B_{3_r 4_d} \\ \sigma^B_{3_r 4_r} & \sigma^B_{4_r 4_r} & \sigma^B_{4_r 3_d} & \sigma^B_{4_r 4_d} \\ \sigma^B_{3_r 3_d} & \sigma^B_{4_r 3_d} & \sigma^B_{3_d 3_d} & \sigma^B_{3_d 4_d} \\ \sigma^B_{3_r 4_d} & \sigma^B_{4_r 4_d} & \sigma^B_{3_d 4_d} & \sigma^B_{4_d 4_d} \end{bmatrix}$$

Where the first two rows and two columns are for the prescription model (2 brands), and the last two rows and columns are for the detailing model for the same 2 brands.

With these three matrices, $\Sigma^{\overline{C}}$, $\Sigma^{\overline{JA}}$ and $\Sigma^{\overline{JB}}$, we can obtain the covariance matrix for the random error terms in both prescription and detailing models, which has dimension of 8×8. The covariance matrix is presented in Table 1.

In this covariance matrix, the first 4 rows and columns are for the prescription model, and the last 4 rows and columns are for the detailing model. In the prescription model, the within category covariances are calculated as the sum of entries in the covariance matrix for each category ($\Sigma^{\overline{JA}}$ or $\Sigma^{\overline{JB}}$) and the entries in $\Sigma^{\overline{C}}$. The cross category covariances are represented by the entries in $\Sigma^{\overline{C}}$.Using this error component structure dramatically reduces the number of parameters to estimate in this covariance matrix, and more importantly, it can explicitly measure the cross-category covariances, which is of particular interest in this study.

2.5 Key features of our model specification

1) *Multivariate Poisson-Lognormal model:* Our model formulation represents a multivariate system of Poisson variables with the random error terms in the log-link specification of the Poisson parameters assumed to follow a multivariate Normal distribution. Taken together, our formulation falls into the class of Multivariate Poisson-Lognormal models (Aitchison and Ho 1989; Munkin and Trivedi 1999; Riphahn et al. 2003). This model has the following three important features, as compared to a standard Poisson model. It allows for (1) over dispersion; (2) correlations across Poisson variables; and (3) larger proportion of zeros than a standard Poisson model.
2) *Simultaneity:* Different from previous Multivariate Poisson-Lognormal specifications, our model explicitly accounts for simultaneity across the different Poisson variants since the Poisson parameter in one (detailing) equation is an explicit function of the Poisson parameters in other (detailing) equations. What results is a simultaneous system of Poisson equations that generalizes the multivariate Poisson models presented in the literature so far. As we discuss below, the estimation of the proposed Poisson simultaneous equation model is a nontrivial task since it requires a model that can accommodate both the count feature and the interdependencies of the dependent variables within (between prescriptions and detailing) and across categories (for prescriptions and detailing).
3) *Endogeneity*: When accounting for endogeneity of marketing instruments, there are two approaches that have typically been used in the literature. The first is the "full information" approach which requires the researcher to specify the data generating process (DGP) for the endogenous variable. Here, researchers have typically resorted to the assumptions of profit maximizing by firms and equilibrium outcomes across firms to specify the DGP. Such an approach, if correctly specified, results in estimates that are consistent and efficient. However, if the DGP is incorrectly specified, then not only will the estimates in the marketing instrument equations be inconsistent, but this inconsistency will spill over into the demand parameter estimates as well. In our analysis, one reason for not following that "optimizing" approach is that it is not clear what the firms' objective functions are in a multi-category context that are generating the observed data in the marketplace, and using that approach would require

Table 1 The full covariance matrix in the example

Prescription model				Detailing model			
Category1		Category2		Category1		Category2	
Brand1	Brand2	Brand3	Brand4	Brand1	Brand2	Brand3	Brand4
$\sigma^{A}_{1_r1_r} + \sigma_{A_rA_r}$	$\sigma^{A}_{1_r2_r} + \sigma_{A_rA_r}$	$\sigma_{A_rB_r}$	$\sigma_{A_rB_r}$	$\sigma^{A}_{1_r1_d} + \sigma_{A_rA_d}$	$\sigma^{A}_{1_r2_d} + \sigma_{A_rA_d}$	$\sigma_{A_rB_d}$	$\sigma_{A_rB_d}$
$\sigma^{A}_{2_r1_r} + \sigma_{A_rA_r}$	$\sigma^{A}_{2_r2_r} + \sigma_{A_rA_r}$	$\sigma_{A_rB_r}$	$\sigma_{A_rB_r}$	$\sigma^{A}_{2_r1_d} + \sigma_{A_rA_d}$	$\sigma^{A}_{2_r2_d} + \sigma_{A_rA_d}$	$\sigma_{A_rB_d}$	$\sigma_{A_rB_d}$
$\sigma_{A_rB_r}$	$\sigma_{A_rB_r}$	$\sigma^{B}_{3_r3_r} + \sigma_{B_rB_r}$	$\sigma^{B}_{3_r4_r} + \sigma_{B_rB_r}$	$\sigma_{B_rA_d}$	$\sigma_{B_rA_d}$	$\sigma^{B}_{3_r3_d} + \sigma_{B_rB_d}$	$\sigma^{B}_{3_r4_d} + \sigma_{B_rB_d}$
$\sigma_{A_rB_r}$	$\sigma_{A_rB_r}$	$\sigma^{B}_{4_r3_r} + \sigma_{B_rB_r}$	$\sigma^{A}_{4_r4_r} + \sigma_{A_rA_r}$	$\sigma_{B_rA_d}$	$\sigma_{B_rA_d}$	$\sigma^{B}_{4_r3_d} + \sigma_{B_rB_d}$	$\sigma^{B}_{4_r4_d} + \sigma_{B_rB_d}$
$\sigma^{A}_{1_r1_d} + \sigma_{A_rA_r}$	$\sigma^{A}_{1_d2_r} + \sigma_{A_rA_d}$	$\sigma_{B_rA_d}$	$\sigma_{B_rA_d}$	$\sigma^{A}_{1_d1_d} + \sigma_{A_dA_d}$	$\sigma^{A}_{1_d2_d} + \sigma_{A_dA_d}$	$\sigma_{A_dB_d}$	$\sigma_{A_dB_d}$
$\sigma^{A}_{1_r2_d} + \sigma_{A_rA_d}$	$\sigma^{A}_{2_d2_r} + \sigma_{A_rA_d}$	$\sigma_{B_rA_d}$	$\sigma_{B_rA_d}$	$\sigma^{A}_{2_d1_d} + \sigma_{A_dA_d}$	$\sigma^{A}_{2_d2_d} + \sigma_{A_dA_d}$	$\sigma_{A_dB_d}$	$\sigma_{A_dB_d}$
$\sigma_{A_rB_d}$	$\sigma_{A_rB_d}$	$\sigma^{B}_{3_d3_r} + \sigma_{B_rB_d}$	$\sigma^{B}_{3_d4_r} + \sigma_{B_rB_d}$	$\sigma_{A_dB_d}$	$\sigma_{A_dB_d}$	$\sigma^{B}_{3_d3_d} + \sigma_{B_dB_d}$	$\sigma^{B}_{3_d4_d} + \sigma_{B_dB_d}$
$\sigma_{A_rB_d}$	$\sigma_{A_rB_d}$	$\sigma^{B}_{4_d3_r} + \sigma_{B_rB_d}$	$\sigma^{B}_{4_d4_r} + \sigma_{B_rB_d}$	$\sigma_{A_dB_d}$	$\sigma_{A_dB_d}$	$\sigma^{B}_{4_d3_d} + \sigma_{B_dB_d}$	$\sigma^{B}_{4_d4_d} + \sigma_{B_dB_d}$

information on all categories marketed by all firms (in order to set up the detailing equations)—an impractical requirement in most empirical contexts. A consequence of our chosen "limited information" approach is that we will not be able to say anything about the DGP for the marketing instruments of interest. This also results in the estimates for the demand parameters being less efficient. The major benefit of this approach however, is that we are guaranteed to obtain consistent estimates for the demand parameters which dealing with the endogeneity of marketing instruments. Chintagunta et al. (2006) discuss this issue in greater detail.

2.6 Bayesian estimation

The set of unknown parameters that need to be estimated can be represented as $\Theta = \left\{\{\theta_{j0}, \theta j\}, s_{\delta F0}, \Sigma v, s_k, s_{\delta F}, \{\beta_{jj}.\beta_{jk}, \beta_{jl}\}, \Sigma_\beta, \{\alpha_j\}, \Sigma^{\overline{C}}, \Sigma^{\overline{JC}}\right\}$. In order to obtain individual physician level parameters in the prescription model, we cast our model in a Hierarchical Bayesian framework. All the model parameters[6] are estimated using the joint likelihood of all the model parameters being given by

$$
\begin{aligned}
&L\left(\{\theta_{j0}, \theta_j\}, s_{\delta F0}, \textstyle\sum_v, s_k, s_{\delta F}, \sum \beta, \{\beta_{jj}, \beta_{jk}, \beta_{jl}\} \sum \beta, \{\alpha_j\}, \sum^{\overline{C}}, \sum^{\overline{JC}} | rx_{ijt}, dtl_{ijt}, Z_i\right) = \\
&\prod_{i,j,t} prob\left(rx_{ijt} | rx_{ij,t-1}, \{dtl_{it}\}, \{\beta_{ij}\}, \in_{ijt}\right) \\
&\times prob\left(dtl_{ijt} | rx_{ij,t-1}, \{dtl_{i,t-1}\}, \beta_{ij0}, \beta_{ijj}, \{\alpha_j\}, \eta_{ijt}\right) \\
&\times prob\left(\{\beta_{ij0}, \beta_{ijj}\} | \{\theta_0, \theta_j\}, s_{\delta F0}, s_k, s_{\delta F}, \textstyle\sum_v\right) \\
&\times prob\left(\left\{\widetilde{\beta}_{ijj}, \beta_{ijk}, \beta_{ijl}\right\} | \{\beta_{jj}, \beta_{jk}, \beta_{jl}\}, \sum \beta\right) \\
&\times prob\left(\in_{ijt}, \eta_{ijt} \middle| \sum^{\overline{C}}, \sum^{\overline{JC}}\right)
\end{aligned}
\tag{15}
$$

where $prob\left(rx_{ijt} | rx_{ij,t-1}, \{dtl_{it}\}, \{\beta_{ij}\}, \in_{ijt}\right)$ is the Poisson probability for the number of prescriptions rx_{ijt} dispensed by physician i for brand j in time period t, which is specified by the prescription model. $prob\left(dtl_{ijt} | rx_{ij,t-1}, \{dtl_{i,t-1}\}, \beta_{ij0}, \beta_{ijj}, \{\alpha_j\}, \eta_{ijt}\right)$ is the Poisson probability for the number of detailing visits targeted to physician i for brand j in time period t, specified by the detailing model. Details of deriving this probability will be discussed next. $prob\left(\{\beta_{ij0}, \beta_{ijj}\} | \{\theta_0, \theta_j\}, s_{\delta F0}, s_k, s_{\delta F}, \Sigma_v\right)$ is the multivariate normal probability function for both the hierarchical regression on the individual-brand specific intercepts and the error component structure of the own detailing response parameters. $prob\left(\left\{\widetilde{\beta}_{ijj}, \beta_{ijk}, \beta_{ijl}\right\} | \overline{\beta}, \Sigma_\beta\right)$ is the multivariate normal probability function for all the other prescription parameters within a category. $prob\left(\in_{ijt}, \eta_{ijt} \middle| \Sigma^{\overline{C}}, \Sigma^{\overline{JC}}\right)$ is the probability function resulted from the error component structure specified in Eqs. (12) and (13).

Markov Chain Monte Carlo method (MCMC), coupled with data augmentation techniques (Tanner and Wong 1987) are used to facilitate the parameter estimations. Details of the full conditional distributions in the MCMC procedure are presented in Appendix A.

[6] For easier reference, we list all the parameters we estimate in Table 2.

Table 2 List of parameters to estimate in the model

Prescription model	$\{\beta_{ij}\}$	Individual level parameters in the prescription model for individual i brand j
	$\{\theta_{j0},\theta_j\}$	Parameters for the hierarchical regression on the intercept β_{ij0}
	Σ_v	Covariance matrix for the intercepts and own detailing parameters
	Σ_β	Covariance matrix for the parameters in the prescription model other than the intercepts and the cross-category components of the own detailing parameters.
Detailing model	$\{\alpha_j\}$	Parameters for the detailing model
Parameters for cross-category correlations	s_k	Additional covariance for own detailing response parameters across brands and categories
	$s_{\delta F}$	Additional covariance of the own detailing response parameters for the two brands marketed by the same company F
	$s_{\delta F0}$	Covariance of the intercepts for the brands marketed by the same company F
	$\sum^{\overline{C}}$	Covariance matrix for the category-specific shocks for both prescription and detailing models, in the Multivariate Normal distribution
	$\sum^{\overline{JC}}$	Two covariance matrices for the brand-specific shocks for both prescription and detailing models, in the Multivariate Normal distribution, one for each category

For the detailing model, according to Eq. (4), a brand's Poisson parameter for its detailing distribution, γ_{ijt} is a function of other brands' Poisson parameters γ_{ijt}, $g \in F$, $g \neq j$. In order to derive the joint distribution of the detailing visits, we need to derive the joint distribution of the multivariate Poisson variables. A key contribution of our analysis is to propose an approach that facilitates estimation of the model parameters in the presence of such simultaneity. Essentially, we solve the linear system of simultaneous equations associated with the Poisson parameters, i.e., the system of Eq. in (4), to obtain the system's "reduced form." This system has a unique and closed form solution and expresses the (logarithm of the) Poisson parameters as functions of the original model parameters. We refer to these as the "transformed" Poisson parameters. We then use these transformed parameters in the specification of the likelihood function.

As the likelihood function is an explicit function of the original model parameters (transformed), identification does not require the rank condition, (the detailed discussion can be seen in Greene (2008), p. 368). However, it does require the order condition to ensure the existence of the solution to the simultaneous equation model. As long as there is at least one exogenous variable in each equation, the order condition is satisfied. This requirement can be easily achieved by Eq. (4), as multiple variables (other than the other categories' expected detailing) are included that are category-specific.

In the MCMC algorithm, we still simulate the original model parameters, as specified in Eq. (4). Given the simultaneity of the Poisson parameters, we obtain the transformed versions of these parameters by solving the system in Eq. (4) across brands. We then use these transformed parameters in the likelihood function.[7] This

[7] We also contacted a simulation study, which confirmed the validity of this method. The details of the simulation study are in Appendix B.

method is very easy to implement. More importantly, it allows us to estimate a simultaneous equation model with a Multivariate Poisson-Lognormal specification while still maintaining the cross-category correlation structure.

3 Data and results

Our data are collected and made available to us by ImpactRx, a marketing research company. The data represent: (1) a detailed record of (Primary Care) physicians' prescription behavior for drugs in the Proton Pump Inhibitor (PPI) and Anti-Depressant (AD) categories over a 3 year period from 2001 to 2004; (2) the detailing effort at the individual physician level from all the firms in these two categories. The data are available at the bi-monthly level. Both therapeutic categories are large in that they treat symptoms that affect a large number of patients—about 60 million for PPI and about 34 million for AD. This makes both categories among the top three prescription categories in the US. Total sales are $12.5 billion for PPI (ranked second) and $11 billion for AD (ranked third), according to IMS Health (2005).

The data are obtained from a representative panel of primary care physicians who report to the market research firm on their prescription writing activity and marketing activity directed at them (detailing). The physicians are compensated for their time. The main "sales" activity recorded by the physician is the number of new prescriptions (NRx)[8] written in each time-period. The physicians record this for all the major brands in each category. They also report the number of detailing calls from each brand in each category they received.

The quality of the information received from the panel is ensured via constant monitoring. Based on almost 9 years of experience collecting these data, the company is confident of the integrity and representativeness of the information collected. Indeed, we were able to verify this claim by correlating the prescription data from our aggregated sample with those from the actual aggregated data made available to us by IMS—a leading market research provider in the pharmaceutical industry. We find that these correlations range from 0.86 to 0.97. We view this as reasonable corroboration for the market representativeness of our panel data.

3.1 Data patterns

For our analysis, we use the four top brands in each category as they account for 99.1% and 74.4% of all prescriptions in the PPI and AD categories respectively. Protonix in the PPI category and Effexor in the AD category are marketed by the same company—Wyeth Pharmaceuticals. These two brands are both among the top four brands in each category, and they are marketed to the same physicians, but face different competitors across the two categories. Six other firms market the remaining brands. Note that while these seven firms (in total) compete in other categories as

[8] The industry distinguishes between total prescriptions (TRx) and new prescriptions (NRx). TRx contain both new prescriptions and renewals (which are mostly automatic). NRx are more important from our point of view as they capture the outcome of an actual decision process.

well, it is quite difficult to obtain data across all these categories for each individual physician.

We restrict our analysis to physicians who receive at least one detail in any bi-monthly period in both categories. This is done to ensure that we have enough information to model individual physician's response to detailing in both categories. Our sample consists of 223 physicians—we randomly selected 200 physicians for our estimation sample and 23 for our holdout sample. We thus have a total of 3600 observations–18 bi-months of data for 200 physicians.[9] The summary statistics of the data for all eight brands across all observations are given in Table 3.

From Table 3, the average number of prescriptions written by each physician ("Prescriptions") shows that Nexium is the market leader in the PPI category, and Celexa is the leader in the AD category. In the table, "Detailing" refers to the mean number of details received in each bi-monthly period. Next, we examine the correlations across brands in prescriptions and detailing in different ways. First, Tables 4 and 5 list the raw correlations across physicians and over time. They show that correlations of both prescriptions and detailing are comparable within and across categories,[10] which indicates the importance of accounting for the cross category effects in modeling both prescription and detailing decisions. In addition, among all the detailing correlations in Table 5, the highest at 0.56 is between Protonix and Effexor, two brands marketed by the same company.

We then carry out a few simple analyses in order to better understand the phenomenon that leads to the high correlations between the detailing from the two brands that are marketed by the same company. To do this, we estimated three simple regressions for each of the two brands, and then computed the correlations of the residuals between the two brands. In the first regression, the dependent variable is one brand's detailing, and the explanatory variables include the intercept, lagged detailing (same brand) and lagged prescriptions (same brand). In the second regression, the other brand's detailing is added as an explanatory variable. The correlation between the two brands' residuals changes dramatically, from 0.40 to −0.28. In the third regression, physician specific dummy variables are added to the variables from the second regression. The correlation between the two brands' residuals barely changes–it goes from −0.28 to −0.32. These results indicate that the high correlation between the two brands by Wyeth, as shown in Table 5, could mainly be attributed to the simultaneity of the detailing decisions for these brands. The same pattern cannot be replicated for other brand pairs.

Overall, these correlations suggest that detailing efforts targeted at individual physicians are not set randomly, even across categories. Finally, the correlations among the detailing visits within and across categories emphasize the importance of considering the detailing endogeneity and simultaneity issues in understanding the marketing variables' effects on prescription behavior *across* categories.

[9] This sample size is typical among studies using physician panel data—Gönül et al. (2001) use 157 physicians over 4 years, Janakiraman et al. (2008) use a panel of 108 physicians and Dong et al. (2009) use a panel of 330 doctors.

[10] Note that on average, the cross-category correlation for prescriptions is slightly lower than the within-category correlations.

Table 3 Summary statistics

Brand	Marketed by	Prescriptions		Detailing	
		Mean	Sd. Dev.	Mean	Sd. Dev.
Aciphex	Janssens & Eisai	4.4	(6.3)	2.0	(2.0)
Nexium	Astra Zeneca	8.1	(9.3)	2.7	(2.3)
Prevacid	TAP	5.9	(8.5)	2.5	(2.2)
Protonix	Wyeth	4.4	(5.9)	2.0	(2.3)
Celexa	Forest	9.4	(10.8)	2.6	(2.1)
Effexor	Wyeth	4.4	(5.9)	1.9	(2.0)
Paxil	GlaxoSmithKline	5.8	(6.4)	1.7	(1.6)
Zoloft	Pfizer	5.5	(6.5)	1.2	(1.4)

3.2 Other variables included in the model

We need two additional sets of variables before we proceed to the estimation of our model. The first set consists of variables for the hierarchy in the prescription model in Eq. (4)–represented by Z. Physicians are likely to differ in their average propensities to prescribe drugs in these two categories based on the size of their practices. We do not have data on the size of the practice for each physician, but we proxy for it using two variables. The first one is the initial total annual category prescriptions across three large therapeutic categories–Antihistamines, PPI and Anti-Infectives for the AD prescription model and Antihistamines, AD and Anti-Infectives for PPI prescription model. The second one is the population density of the zip code where the physician is located.

Another challenge in this analysis is to find exogenous variables that impact the detailing decisions by each firm for each physician (the W_{it} variable in Eq. (4)). We use two variables, one varying across physicians and the other one over time. For the first variable, we used the zip code location information for each physician (the only demographic data available to us at the physician level) and obtained data on number of physicians for each zip code from the 2004 physician directory published by the

Table 4 Correlation across prescriptions

	Aciphex	Nexium	Prevacid	Protonix	Celexa	Effexor	Paxil	Zoloft
Aciphex	1	0.30	0.26	0.22	0.26	0.09	0.20	0.19
Nexium		1	0.31	0.22	0.29	0.18	0.25	0.28
Prevacid			1	0.25	0.22	0.13	0.20	0.18
Protonix				1	0.23	0.20	0.12	0.22
Celexa					1	0.30	0.21	0.26
Effexor						1	0.23	0.27
Paxil							1	0.24
Zoloft								1

Table 5 Correlation across detailing

	Aciphex	Nexium	Prevacid	Protonix	Celexa	Effexor	Paxil	Zoloft
Aciphex	1	0.33	0.38	0.35	0.34	0.30	0.24	0.30
Nexium		1	0.41	0.33	0.35	0.26	0.30	0.28
Prevacid			1	0.36	0.37	0.30	0.35	0.30
Protonix				1	0.34	0.56	0.35	0.36
Celexa					1	0.38	0.31	0.31
Effexor						1	0.32	0.34
Paxil							1	0.36
Zoloft								1

American Medical Association. Given that most physicians are located at different zip codes in our estimation data set, this variable captures cross-sectional variation across physicians. This variable could have two possible effects on detailing decisions. On one hand, a lower number of physicians in the same zip code could lead to higher cost of detailing visits, and therefore a lower level of detailing visits to each physician. On the other hand, a lower number of physicians in the zip code could mean that physicians in this area are more dependent on detailing visits for their information about drugs, and therefore pharmaceutical companies tend to allocate more visits to these doctors.[11] The second variable included in W_{it} is a holiday dummy that takes the value 1 if it is during the Christmas-New Year holiday season. A data summary shows that during the holiday season, the number of detailing visits is greatly reduced, while the number of prescriptions written is not impacted much. One possible explanation is that although during holiday season, patients continue to get sick and need prescriptions, additional physicians might be on vacation during this time period leading to a higher load on the remaining physicians, which in turn leads to fewer detailing visits to physicians' offices. The holiday dummy variable varies over time, but not across physicians. These two variables play two roles in the model–first, they vary across physicians (in most cases) and time and, second, they help to capture some information on factors that could *drive detailing levels without directly affecting prescriptions at the individual physician level*.

3.3 Estimation results

Using Markov Chain Monte Carlo estimation, we obtain the parameter estimates for both the prescription and the detailing models simultaneously. Appropriate but diffuse priors are used for this estimation, and the details of each prior are provided in Appendix A. The MCMC chain was simulated for 200,000 iterations, and the last 100,000 iterations are used to obtain the parameter estimates. To ensure convergence, different starting values are used and the time series plots are inspected.

The population level parameter estimates of the prescription model are in Table 6, with the parentheses listing the (2.5%, 97.5%) percentiles of the marginal posterior

[11] To check whether this variable affects prescription decisions of doctors, we checked the effects of this variable in explaining the variation in physicians' individual level intercepts. Our regression results show that the effects are not statistically significant.

Table 6 Population level mean estimates for the prescription model

	β_{ijj}	$\beta_{ijk}, k \neq j$				β_{ijl}
	Own detailing	Competitive detailing				Lagged prescription
PPI	$\frac{1}{1+own\_dtl}$	$dtl_{Aciphex}$	dtl_{Nexium}	$dtl_{Prevacid}$	$dtl_{Protonix}$	$ln(1+rx_{t-1})$
Aciphex	−1.0416 (−1.21, −0.88)		−0.0179 (−0.09, 0.06)	−0.0203 (−0.10, 0.06)	−0.0421 (−0.11, 0.03)	0.1089 (0.02, 0.20)
Nexium	−0.6437 (−0.78, −0.51)	−0.0722 (−0.13, −0.01)		0.0104 (−0.06, 0.08)	0.0235 (−0.04, 0.09)	0.1307 (0.05, 0.20)
Prevacid	−0.5355 (−0.70, −0.36)	−0.0328 (−0.10, 0.03)	−0.0179 (−0.08, 0.05)		0.0168 (−0.06, 0.10)	0.0476 (−0.03, 0.13)
Protonix	−0.6499 (−0.81, −0.49)	0.0114 (−0.06, 0.09)	−0.0263 (−0.09, 0.04)	−0.0768 (−0.15, −0.00)		0.1804 (0.09, 0.27)
Anti-Depressant		dtl_{Celexa}	$dtl_{Effexor}$	dtl_{Paxil}	dtl_{Zoloft}	
Celexa	−0.5659 (−0.70, −0.44)		−0.0001 (−0.07, 0.07)	−0.0729 (−0.14, −0.00)	−0.084 (−0.15, −0.01)	0.2115 (0.14, 0.28)
Effexor	−0.7196 (−0.86, −0.58)	0.0116 (−0.06, 0.09)		−0.0975 (−0.18, −0.01)	−0.0078 (−0.10, 0.09)	0.1588 (0.07, 0.25)
Paxil	−0.5112 (−0.75, −0.25)	−0.0332 (−0.10, 0.03)	0.0099 (−0.07, 0.09)		−0.0028 (−0.09, 0.08)	0.1466 (0.07, 0.22)
Zoloft	−0.3034 (−0.54, −0.07)	−0.068 (−0.12, −0.01)	−0.0131 (−0.08, 0.06)	−0.0119 (−0.08, 0.06)		0.0439 (−0.03, 0.12)

distribution for each parameter. The upper part of the table shows the results for the PPI category, and the lower part for the AD category.

The first column lists the population level estimates for the own detailing effects. Recall that own detailing is incorporated using the log-reciprocal transformation, therefore the own detailing parameters are expected to be negative. The results show that all the estimates are negative and the 95% intervals are all negative as well. As we will show later, these estimates yield elasticities that are consistent with those obtained by previous researchers. The next four columns list the estimates for the competitive detailing parameters within each category. In the PPI category, 2 of 12 cross-effects are negative and statistically significant; in the AD category, 4 are negative and significant. All the others contain 0 in their 95% intervals. These results imply that (1) there is considerable heterogeneity in competitive effects of detailing both within as well as across categories; (2) the AD category has more firms influencing their competitors than the PPI category. The last column shows the estimates of the lagged prescription variable, and most of them are positive, indicating the prevalence of state dependence in physicians' prescription behavior.

Recall from Eq. (4) that we allow for cross-sectional differences in physicians' intrinsic preferences for drugs via a hierarchical regression with an intercept (θ_{j0}) and coefficients on Z_i, the proxies for practice size (θ_{j1} and θ_{j2}). The estimates are in Table 7. Column 1 lists the estimates for the intercepts, θ_{j0} in Eq. (4). The second and third columns list the estimates for θ_{j1} and θ_{j2}. All estimated θ_{j0} are positive and the 95% confidence intervals are positive as well. Among the estimated θ_{j1}, four are positive with positive 95% confidence intervals. This shows that physicians with larger practice sizes in general tend to write more prescriptions.

To understand the own detailing effects, we also compute the posterior mean of the own elasticity. In the PPI category, these values range from 0.13 for Prevacidto 0.28 for Aciphex. In the AD category, the values range from 0.09 for Zoloft to 0.20 for Effexor. These results are consistent with current research (Albers et al. 2010). The parameter estimates for the detailing model are listed in Tables 8 and 9. As mentioned earlier, besides the intercepts in the detailing model (estimated results listed in the first column in Table 8), the specification contains three sets of

Table 7 Estimation results for the hierarchical regression in the prescription model

	Intercept θ_{j0}	Total Rxs of other categories θ_{j1}	Population density θ_{j2}
Aciphex	0.2883 (−0.09, 0.62)	1.2957 (0.71, 1.87)	0.0580 (−0.05, 0.17)
Nexium	0.7567 (0.49, 1.02)	1.3072 (0.82, 1.80)	−0.0052 (−0.08, 0.06)
Prevacid	0.5705 (0.26, 0.89)	1.0949 (0.63, 1.57)	0.0580 (−0.04, 0.16)
Protonix	0.0459 (−0.30, 0.40)	1.0308 (0.52, 1.56)	0.1486 (0.04, 0.25)
Celexa	1.0803 (0.81, 1.35)	0.4738 (0.06, 0.90)	0.0153 (−0.06, 0.09)
Effexor	0.2533 (−0.15, 0.64)	0.6003 (0.02, 1.02)	0.0913 (0.01, 0.17)
Paxil	0.4647 (0.17, 0.78)	1.2117 (0.76, 1.66)	0.0794 (0.00, 0.16)
Zoloft	0.2472 (−0.10, 0.57)	1.4644 (1.00, 1.95)	0.0985 (0.02, 0.19)

Table 8 Estimates of the first set of parameters in the detailing model

	Intercept of the detailing model	α_{j1} Effect of prescription model intercept β_{ij0}	α_{j2} Effect of prescription model detailing response β_{iji}
Aciphex	−0.0118 (−0.06, 0.04)	0.267 (0.24, 0.30)	−0.1369 (−0.21,−0.06)
Nexium	0.023 (−0.02, 0.06)	0.2566 (0.20, 0.30)	−0.3099 (−0.35,−0.27)
Prevacid	0.018 (−0.02, 0.05)	0.2733 (0.24, 0.34)	−0.4169 (−0.45,−0.36)
Protonix	−0.0391 (−0.08,−0.01)	0.2086 (0.17, 0.26)	−0.2589 (−0.32,−0.20)
Celexa	−0.0487 (−0.07,−0.02)	0.1504 (0.11, 0.20)	−0.3562 (−0.40,−0.32)
Effexor	0.0485 (0.01, 0.07)	0.3003 (0.23, 0.34)	−0.2734 (−0.31,−0.23)
Paxil	0.0835 (0.06, 0.11)	0.2713 (0.24, 0.33)	−0.2868 (−0.36,−0.24)
Zoloft	−0.1778 (−0.21,−0.14)	0.2119 (0.19, 0.25)	−0.2512 (−0.28,−0.21)

explanatory variables. The first set relates to two parameters that are consistent with the heuristic approach by Manchanda et al. (2004). They vary across physicians and brands, but are time invariant. These explanatory variables include the (i) parameters for the individual-brand specific intercepts in the prescription model and (ii) parameters for the response to own detailing in the prescription model. The results (α_{j1} and α_{j2}) of the coefficients for these two variables are listed in Table 8, together with the intercept in the detailing model α_{j0}.

The second numeric column of Table 8 shows that all of the parameters α_{j1} for the prescription model intercepts are positive and statistically different from zero, this is true across all brands. It indicates that firms tend to detail physicians who have a higher intrinsic preference towards their own brands. The third column shows that all of the parameters α_{j2} for the prescription model's own detailing responses are negative and most of them are statistically different from zero. Given that in the prescription model, the own detailing responses are higher when the estimates of the response parameters are more negative, the negative estimates for α_{j2} indicate that firms allocate more detailing effort to those physicians who are more responsive to the firm's detailing. These results for both α_{j1} and α_{j2} reflect our intuition that firms set their detailing efforts at the individual physician level based on their understanding of how physicians respond to these efforts. The results also emphasize the importance of incorporating the detailing model to account for the endogeneity of the detailing decisions at the physician level (Manchanda et al. 2004).

The second set of variables relates to the time varying variables represented by the lagged own detailing and lagged competitive detailing, as well as lagged prescriptions. As with the heuristic model, we could let the effects of these factors be invariant across physicians. However, we were able to allow for additional heterogeneity in the effects of lagged own and competitive detailing levels.

The population levelestimates for the first two lagged variables are listed in the first two numeric columns in Table 9—for own and competitive detailing respectively. They are all positive and statistically different from zero, which

Table 9 Estimates for the other parameters in the detailing model

	α_{jn} Lagged own detailing $ln(1+owndtl_{t-1})$	α_{jk} Lagged competitive detailing $ln(1+compdtl_{t-1})$	α_{jl} Lagged prescriptions $ln(1+rx_{t-1})$	α_{jw1} Number of physicians in the same zip code ln(number of physicians)	α_{jw2} Holiday dummies
Aciphex	0.3491	0.0476	0.0004	−0.3042	−0.111
	(0.25, 0.45)	(−0.03, 0.12)	(−0.03, 0.03)	(−0.38,−0.24)	(−0.14,−0.07)
Nexium	0.2623	0.026	0.0456	−0.3549	−0.0638
	(0.18, 0.34)	(−0.05, 0.10)	(0.02, 0.07)	(−0.40,−0.31)	(−0.09,−0.03)
Prevacid	0.266	0.0051	0.0324	−0.4042	−0.0787
	(0.18, 0.35)	(−0.07, 0.08)	(0.00, 0.06)	(−0.43,−0.37)	(−0.10,−0.05)
Protonix	0.2681	0.0049	0.0229	−0.1919	−0.0898
	(0.14, 0.38)	(−0.08, 0.09)	(0.00, 0.04)	(−0.23,−0.16)	(−0.13,−0.06)
Celexa	0.3037	0.0153	0.0457	−0.1966	−0.0316
	(0.22, 0.38)	(−0.06, 0.09)	(0.02, 0.08)	(−0.24,−0.15)	(−0.05,−0.01)
Effexor	0.2502	−0.0336	0.0532	−0.2893	−0.1119
	(0.15, 0.35)	(−0.12, 0.05)	(0.02, 0.10)	(−0.31,−0.26)	(−0.13,−0.09)
Paxil	0.2485	0.0082	−0.0079	−0.2363	−0.1474
	(0.14, 0.35)	(−0.07, 0.09)	(−0.03, 0.01)	(−0.27,−0.20)	(−0.18,−0.12)
Zoloft	0.1844	0.055	−0.0312	−0.2513	−0.1415
	(0.09, 0.28)	(−0.06, 0.17)	(−0.06,0.00)	(−0.29,−0.22)	(−0.16,−0.11)

indicates carryover in detailing effort over time at the individual physician level. The second column lists the estimated coefficients for the lagged competitive detailing variable. Note that the detailing model is flexible enough to allow for different competitive effects (see Eq. (4)) for the different competitors. However, our empirical results show that estimating a single parameter across all competitors is adequate for our data. The results show that none of these parameters is statistically significant. The third column in Table 9 lists the estimates for the lagged prescriptions. Among the 8 estimates, 5 of them are positive and statistically different from zero. This indicates that firms adjust their detailing efforts based on information about lagged prescriptions written by each physician and is especially true for the four firms in the AD category.

The third set of variables in the detailing model includes two exogenous variables, which are (1) number of physicians in the same zip code incorporated in the log-transformation form; and (2) a dummy variable indicating the Christmas-New Year holiday season. The estimates for the first one are negative and statistically significant, for all eight brands. As discussed earlier, a negative coefficient indicates that the pharmaceutical companies tend to visit physicians that are isolated more, as these physicians have fewer opportunities to learn from other fellow physicians, compared to those in areas of greater physician density. Finally, the last column in Table 9 shows that the coefficients for the holiday dummy variable are negative and statistically significant for all 8 brands. This implies that during the holiday time period, doctors are visited less by detailers.

The last set of parameters estimated in the detailing model captures the simultaneity of the detailing decisions across categories for a firm marketing drugs in those categories. As noted earlier, we have two brands that are marketed (one in each category) by the same company. In this case, we allow one brand's detailing to impact the other brand's detailing decisions for both detailing models. The cross-detailing parameter $\alpha_{jg}, g \neq j, g \in F$ indicates how the other brand's detailing in a different category impacts this brand's detailing decision by the same company for the same physician. This parameter could be positive or negative. On one hand, the firm has limited resources (such as the size of the sales force) in marketing all of its drugs to physicians. As a result, if the firm spends more money on detailing in one category, they might have to decrease their detailing in the other category. This will result in a negative cross-detailing relationship. On the other hand, if physicians differ in terms of their preference for the brands from a particular company or in their responsiveness to detailing in general, firms would want to leverage the information by finding physicians that like the company's products better or who are more responsive than other doctors, and target them with higher levels of detailing effort for all drugs. This will lead to a positive cross-detailing relationship.

The estimated cross-detailing parameters are 0.17 (0.08, 0.24) for the coefficient of Effexor's detailing on Protonix's detailing and 0.11 (0.04, 0.21) for the coefficient of Protonix's detailing on Effexor's detailing. The results show that, Wyeth's detailing decisions in one category positively influence the detailing decisions in the other category. This is consistent with the data

summary statistics, where the correlation between these two brands' detailing decisions across physicians is 0.56.

In addition to the estimates above, recall that we estimate three other sets of parameters capturing the cross-category correlations. The parameters are the variances and the covariance matrices for the error components. We find the following. 1) The covariance between the response parameters $\beta_{ijj} \forall j$, which is the covariance between any two brands across categories is 0.19 (0.12,0.26). Note that even though the covariance values are the same for any pair of brands across categories, the correlation values are different across pairs, as the variances are different. 2) The estimate for the correlation between the response parameters to own detailing effects for Protonix and Effexor (marketed by Wyeth) is 0.27 (0.17, 0.39), which is indicative of the spill-over effect of the detailing effort across categories. The estimate for the correlation between the intercepts of the two brands from Wyeth is 0.26 (0.16, 0.38). This indicates physicians' intrinsic preferences towards these two brands from the same company are correlated, after controlling for all other factors that drive prescriptions.

We now discuss the estimates of the covariance matrices for the random shocks in both prescription and detailing models. In our model, using the error component structure in the random error terms for both prescription and detailing models, we can obtain the cross-category correlations explicitly. The results are listed in Table 10. It shows that the cross-category correlations are non-trivial, especially for the prescription model. One possible reason that physicians' prescription decisions might be correlated across these two categories (PPIs and Anti-Depressants) through the random error terms is co-morbidity. The medical literature (e.g., Wiklund and Butler-Wheelhouse (1996), Avidan et al. (2001), Jansson et al. (2007)) shows that psychiatric diseases or depression (treated by the Anti-Depressant category drugs) could lead to severe reflux symptoms (treated by the PPI category drugs). We do not observe the co-morbidity directly as we do not have data on the prescription(s)for a given patient. Nevertheless, given that we have controlled for a variety of other factors, this appears to be a plausible explanation for our findings.

Table 10 Correlation estimates for the category-specific random terms

		Prescription		Detailing	
		Category 1	Category 2	Category 1	Category 2
Prescription	Category 1	1	0.5356 (0.48, 0.59)	0.0083 (−0.10, 0.12)	0.0804 (−0.01, 0.17)
	Category 2		1	0.0564 (−0.04, 0.15)	0.0641 (−0.04, 0.19)
Detailing	Category 1			1	0.3141 (0.22, 0.40)
	Category 2				1

3.4 Model comparisons

Our model explicitly accounts for cross-category effects in two ways. First, in the detailing model, we allow detailing decisions in one category to be a function of detailing decisions in other categories marketed by the same company. Second, in both the prescription and detailing models, we allow response parameters and unobserved factors to be correlated across categories using three error-component structures. To examine the importance of incorporating cross-category effects using the above-mentioned two ways, we compare our model with two alternative models by changing each of the above two features. The comparisons among these models are conducted using both Log Marginal Likelihood values for in-sample model fit and Mean Absolute Percentage Errors (MAPE) for holdout sample predictions. These values are listed in Table 11. Following Dubé et al. (2010), we calculated the Log Marginal Likelihood values using a Newton-Raftery style estimator, after trimming the top and bottom 1% of the likelihood values. They show that our model fits the data the best. The MAPE values also demonstrate that our model performs the best in holdout sample test.

4 Model implications

We now demonstrate the importance of a multicategory analysis when studying the responsiveness of physicians to detailing behavior of a firm that participates in several product categories that are prescribed by the same set of physicians. Via simulation, we demonstrated earlier that ignoring the cross-category effects in the detailing model would lead to biased estimates in the detailing model. This would lead to biased inferences in the prescription model as well. In the last section, we showed that our model fits the data the best for both in-sample and out-of-sample tests. In this section, we quantify the consequences of applying the alternative models using our data where the cross- category effects are ignored. Specifically, we focus on Wyeth's brands to examine two aspects—the implications for segmentation and insights into the nature of the data generating process.

4.1 Consequences for segmentation

To illustrate the "upside" of using our model estimates, we carry out a counterfactual exercise with respect to segmentation. We compare the segmentation results across

Table 11 Comparison of model fits

	Log marginal likelihood	MAPE
The base model	−92337	1.94
Ignoring detailing efforts in other categories	−92352	2.03
Ignoring all cross-category effects	−92451	1.98

measures from four setups: M1 (the proposed model)- where cross-category effects are incorporated explicitly, M2–based on a comparison model where the other brands' detailing cross-category effects are ignored, M3–based on a comparison model where all of the cross-category effects are ignored and M4–based on the total number of prescriptions by each physician; here no model is necessary for segmenting the doctors.

First, using our model, within each category, we allocate physicians into 10 groups/segments/cells based on their individual detailingelasticities. This is in accordance with industry practice where physicians are segmented into deciles. As a result, this gives us a total of 100 groups across the two categories. Each of the 200 physicians is then classified into one of these 100 groups.

Second, we then implement the same segmentation strategy using the parameter estimates from the other three setups as mentioned above. In M4 we use the prescription-based decile rule that is common industry practice. In Table 12, the first numerical column lists the number of physicians assigned to the same segmentation as in the base model (M1), for each of the three comparison models–M2, M3 and M4. In addition, as a less strict rule, we also compute the number of physicians assigned to either the same "cell" or to one of the neighboring cell. The neighboring cell is defined as a cell that 1 cell away from the base cell. In other words, for an "interior" cell (2–9 for both categories), each cell has 8 neighboring cells. If the model ignores the other categories' detailing in the detailing model (M2), 44 physicians (22% of all the physicians) will still be assigned to the same segments as the base model indicates. If we consider the neighboring groups, it is 148 physicians (74%). If all the cross-category effects are ignored, including the other categories' detailing and the cross-category correlations captured with the three error-component structures (M3), only 20 physicians (10% of all the physicians) will be assigned to the same segments as the base model indicates. If we consider the neighboring cell, it is 93 physicians (46.5%). Finally, if no model is estimated, and the segmentation of the physicians is based on the total number of prescriptions written, none of the physicians will be assigned to the same segments. If we consider the neighboring cells, it is 9 physicians (4.5%). Hence there appear to be significant consequences for segmentation if cross-category effects are ignored.

In the above analysis, we demonstrate the impact on segmentation when cross-category effects are ignored using the short-term elasticities. Next, we study the impact on segmentation using long-term elasticities. The long-term

Table 12 Number of physicians assigned to the same or Neighboring segments as the base model

	Number and percentage of physicians in the same segments		Number and percentage of physicians in the same or neighboring segments	
M2 (the model without the other brand's detailing)	44	(22%)	148	(74%)
M3 (the model without any cross-category factors)	20	(10%)	93	(46.5%)
M4 (based on the total prescriptions)	0	(0%)	9	(4.5%)

elasticities are calculated by scaling the short-term detailing elasticities by $\frac{1}{1-\beta_{ijl}}$ where β_{ijl} is the parameter on the lagged prescription variable in the prescription model (Eq. (2)). While not reported here, we find that the implications on the segmentation using long-term elasticities are very similar to those using the short-term elasticities.

4.2 Consequences for profits

Finally, we compute the impact of alternative models on a firm's profit when marketing across these two categories.[12] As before, we first classify the physicians into deciles using the proposed model. Next we compute the mean numbers of details received by the physicians in each of these segments. With these detailing numbers, we compute the firm's profits assuming that all the physicians in a segment receive the mean number of details for that segment. The corresponding profit level for the firm is our base case. Next, we take the classification from the incorrectly specified models. We then assign the detailing level for each physician using the number of details from the base case for physicians in that segment and compute the firm's profits. In other words, suppose a physician is assigned to the third decile using the incorrectly specified model, we then assign to this physician the mean number of details for physicians from the third decile but using the decile classification from our proposed model. We then compute the loss in profits to the firm from an incorrect classification. We find that if the impact of the other brand's detailing decisions is ignored in the detailing model (M2), the profit drops by 6.4%; and if the correlation across unobserved random factors is ignored (M3), the drop in profit becomes 9.3%. This implies that there are sizable economic consequences of incorrect segmentation of physicians.

5 Conclusion

In this paper, we develop a new multivariate count data model based on the Multivariate Poisson-Lognormal model that accounts for simultaneity across the Poisson variants. We also propose a very simple estimation procedure based on an MCMC algorithm. We apply this model to study multi-category physician prescription behavior, while accounting for the endogeneity simultaneity of detailing decisions across categories. The empirical study demonstrates the importance of a multiple category analysis to estimate physician-level detailing sensitivities for segmentation purposes when firms that market drugs across several therapeutic categories detail these physicians. Specifically, we find

[12] In order to calculate revenue, we make the following additional assumptions: (1) prices of the two drugs Protonix and Effexor are the retail prices obtained from http://www.drugstore.com (accessed on April 14, 2009), and they are assumed to be fixed over time. (2) Cost of detailing is set to be $106, based on a Global Business Insight Report by Seget 2004. (3) We assume competitors do not respond to changes in the company's segmentation. This would give us the short-term effects of the strategy change in segmentation, before the competitors adjust their strategies.

statistically significant cross-category effects from multiple sources. A firm's detailing level in a category appears to be influenced by how responsive the physician is to that firm's detailing; by how much that physician was detailed by that and the rival firms in the previous time period; and by the firm's detailing in other categories in which it markets drugs. Importantly, we demonstrate the need to incorporate these cross category effects, in terms of accurately segmenting physicians based on their responsiveness to detailing efforts, and further, the impact on profits.

Our contributions are therefore, methodological and substantive. On the methods front, our approach is general enough to be applied to other situations with count dependent variables where researchers encounter the simultaneity issue. On the substantive front, quantifying the implications of accounting for cross-category behavior on segmentation and firm profits extends the current literature on cross-category analysis.

While our model is general, our analysis is limited by our data to two therapeutic categories (although we explicitly allow for unobservables that could include information on categories beyond those in our data). Also, our estimates are consistent even though we do not have access to data from other categories. Finally, our data preclude us from modeling the effects of other physician marketing instruments such as meetings and events. We hope that future research will be able to address these limitations. As noted in the introduction, our modeling approach is also suitable in other situations. One example would be in the context of direct marketing where the dependent variable of interest is e.g., the number of items ordered by consumers in different clothing categories. Driving this would be the number of catalogs mailed out for each of these categories where the firm has to decide the number of catalogs in one category (e.g., lingerie) relative to those in another category (e.g., winter wear). Similarly, the model would apply in business-to-business settings where firms are marketing a number of categories to their customers with the numbers of marketing contacts in each of these categories driving purchases in those categories. Future applications of our modeling framework could look at these contexts.

Appendix A: The MCMC Algorithm for Obtaining the Model Parameters

In the following equations, we use * to represent all other parameters

1. Draw the parameters in the prescription model.

Define β_i as the vector of all the parameters in the prescription model for individual i and for all the brands. In order to show clearly the construction of the likelihood, we divide these parameters into two sub-vectors of β_i. The first one, define $\beta_i^1 = \{\beta_{ij0}, \beta_{ijj}, \forall j\}$ as a sub-vector of β_i, which contains the intercept and the own detailing response parameters for all brands. In other words, β_i^1 refers to the parameters in the prescription model that enter the heuristic part of the detailing model. The second sub-vector β_i^2 represents all the other parameters, namely $\{\beta_{jj},\beta_{jk},\beta_{jl},\forall j\}$. Draw β_i for each physician across all brands.

$$
\begin{aligned}
[\beta_i|^*] \propto & \prod_{j,t} Poisson\left(\lambda_{ijt}\left(\beta_{ij}\right)\right) && \text{Likelihood from the prescription model}\\
& \times \prod_{j,t} Poisson\left(\gamma'_{ijt}\left(\beta^1_{ij}\right)\right) && \text{Likelihood from the detailing model}\\
& \times \prod_{j} \left[\left\{ \begin{array}{c} \beta_{ij0} \\ \beta_{ijj} \end{array} \right\} \sim N\left(\left\{ \begin{array}{c} \theta_{j0} + Z_i\theta_j + \delta_{iF0} I_{j\in F1} \\ \kappa_i + \delta_{iF} I_{j\in F1} + \widetilde{\beta}_{ijj} \end{array} \right\}, \sum\nolimits_{v} \right) \right] && \text{Prior for subvector } \beta^1_i\\
& \times \prod_{j} \left[\left\{ \begin{array}{c} \widetilde{\beta}_{ijj} \\ \beta_{ijk} \\ \beta_{ijl} \end{array} \right\} \sim N\left(\overline{\beta}, \sum\nolimits_{\beta}\right) \right] && \text{Prior for subvector } \beta^2_i
\end{aligned}
$$

Where λ_{ijt} is specified as in Eqs. (1) and (2), which is a function of β_{ij} and others; and γ_{ijt} is derived from Eq. (4), which is a function of $\beta^1_{ij} = \{\beta_{ij0}, \beta_{ijj}\}$ and others. The prior for sub-vector β^1_i is derived from Eqs. (7) to (8), with Σ_v as the joint covariance matrix of the random variables in both the hierarchical regression on the intercepts (Eq. (7)) and the error component equation for the own detailing response parameters (Eq. (8)). The prior for sub-vector β^1_i is derived from Eq. (9).

This step involves solving the simultaneous Poisson model. We use γ'_{ijt} to represent the functional form in this conditional marginal distribution for β_i. It is different from that in Eq. (4), for the two brands' detailing parameters that are simultaneously decided. For these two brands, the Poisson parameters in the likelihood function are the "reduced form" solution to these two simultaneous equations. For the other brands with no simultaneity $\gamma'_{ijt}(\,) = \gamma_{ijt}(\,)$.

This step also involves augmenting the elements in the two error components structures (Eqs. (6) and (7)), and their covariance matrices. The augmentation happens in three steps in the sampler. First, the base parameters without the additional errors are drawn, conditional on κ_i and δ_{iF}. Then, the error component κ_i, which is common across all brands and represents the correlations across own detailing response parameters for each individual, is drawn. When drawing these κ_i, the likelihood function is the same as the part for simulating β_i. The only change here is the prior, which is changed to be $\kappa_i \sim N(0,\Sigma_\kappa)$. Finally, the error component δ_{iF}, which is additive to the own detailing parameter for only the two brands marketed by the same company, is drawn. Again, the only change here is the prior, which is changed to be $\delta_{iF} \sim N(0,\Sigma_{\delta F})$. Additionally, two standard Inverted Wishart draws are performed to draw Σ_κ and Σ_δ.

2. Draw the parameters in the detailing model

Define $\alpha = \{\alpha_j, \forall_j\}$ as the vector for all the parameters in the detailing model, for all brands. Draw $\alpha = \{\alpha_j, \forall_j\}$ using the detailing model likelihood function.

$$
\begin{aligned}
[a|^*] \propto & \prod_{i,j,t} Poisson\left(\gamma'_{ijt}\left(\alpha_{ij}\right)\right) Likelihood\ from\ the\ detailing\ model\\
& \times [\alpha \sim N(\alpha_0, \Sigma_{\alpha 0})] \qquad Prior
\end{aligned}
$$

The prior distribution is specified as $\alpha_0 = \{0\}$, and $\Sigma_{a0} = 0.01I$.

3. Draw the category-specific and brand-specific random factors in both the prescription and detailing models (see Eqs. (11) and (12)), and the covariance matrices.

First, draw the category-specific random factors (e_{ict} in the prescription model and u_{ict} in the detailing model). Define $e_{it} = \{e_{ict}, \forall_C\}$ and $u_{ict} = \{u_{ict}, \forall_C\}$

$$\begin{aligned}[e_{it}, u_{it}|^*] \propto & \prod_{j \in c} Poisson\left(\lambda_{ijt}(e_{ict})\right) && \textit{Likelihood from the prescription model} \\ & \times \prod_{j \in c} Poisson\left(\gamma'_{ijt}(u_{ict})\right) && \textit{Likelihood from the detialing model} \\ & \times \left[\{e_{it}, u_{it}\} \sim N\left(0, \sum \overline{C}\right)\right] && \textit{Prior}\end{aligned}$$

Note that, at this step we allow the category-specific random factors in the prescription and detailing models to be correlated across the two models as well as across categories.

Second, draw the brand-specific random factors (f_{ijt} in the prescription model and v_{ijt} in the detailing model). Define $f_{ijt} = \{f_{ijt}, \forall_j\}$ and $v_{ijt} = \{v_{ijt}, \forall_j\}$

$$\begin{aligned}[f_{it} v_{it}|^*] \propto & \prod_{j} Poisson\left(\lambda_{ijt}\left(f_{ijt}\right)\right) && \textit{Likelihood from the prescription model} \\ & \times \prod_{j} Poisson\left(\gamma'_{ijt}\left(v_{ijt}\right)\right) && \textit{Likelihood from the detailing model} \\ & \times \left[\{f_{it}, v_{it}\} \sim N\left(0, \sum \overline{JC}\right)\right] && \textit{Prior}\end{aligned}$$

Similar to drawing the category-specific random factors, at this step we allow the brand-specific random vectors in the prescription and detailing models to be correlated across the two models as well as across brands.

Now, conditional on the random factors from these two steps, draw the covariance matrices for each group of them, using Inverted Wishart conjugate priors. Details are omitted here, but can be found in any Bayesian text book, such as Rossi et al. (2006).

Third, draw $\sum \overline{C}$, conditional on $\{e_{ict}, u_{ict}, \forall c\}$.

Forth, draw $\sum \overline{JC}$, conditional on $\{f_{ijt}, v_{ijt}, \forall c\}$

4. Draw $\overline{\beta}$, Σ_β with normal and Wishart conjugate prior. The conditional posterior distribution of $\overline{\beta}$ *is* $N(\mu_\beta, \Delta_\beta)$, where

$$\Delta_\beta = \left(P \times \sum\nolimits_{\beta}^{-1} + \Delta_0\right)^{-1}$$

$$\mu_\beta = \Delta_\beta \times \left(\sum\nolimits_{\beta}^{-1} \times \sum_i \beta_{i1} + \Delta_0^{-1} \times \mu_0\right)$$

Where P is the number of physicians in the analysis. $\Delta_0^{-1} = 100I$, and $\mu_0 = \{0\}$. The conditional posterior distribution of $\Sigma_\beta \propto$ Inverted Wishart $\left(\sum_i \left(\beta_{i1} - \overline{\beta}\right)' \left(\beta_{i1} - \overline{\beta}\right) + V_0, P + n_0\right)$, where $n_0 = 42$, and $V_0 = n_0 I$.

5. Draw $\theta = \{\theta_{j0}, \theta_j, \forall j\}$ in the hierarchical regression for the intercepts in the prescription model, as specified in Eq. (7). Conditional on all the other parameters, including the intercepts in the prescription model β_{ij0} drawn from 1, this is a multivariate normal regression, as it assumes that $\{v_{ij}, \forall j\} \sim \text{MVN}(0, \Sigma_v)$. For more details of normal regression, please refer to the book by Rossi et al. (2006).

Appendix B: Estimation of the Simultaneous Poisson Model

When a firm that is marketing a product in one therapeutic class is also marketing a product in other therapeutic classes, it is likely that its detailing decisions in one category depend on the detailing decisions in the other categories since all categories are being detailed to the same set of physicians. To account for this inter-dependence in the firm's detailing decisions, in the detailing model, we allow the Poisson parameter of one brand's detailing (in one therapeutic class) to depend on the firm's other brands' (in other therapeutic classes) Poisson parameter. This creates simultaneity among the equations. In order to estimate the structural parameters we need to write out the joint distribution (the multivariate Poisson) .

Unfortunately, such a distribution function is difficult to obtain (see Cameron and Trivedi 1998, Chapter 8). We develop the following method that allows us to estimate the structural parameters without explicitly writing out the multivariate Poisson distributions. Specifically, in the estimation process, we use MCMC draws to simulate the posterior distribution of the structural parameters. Based on the simulated draws of the structural parameters, we then transform them to the parameters corresponding to the reduced form of the simultaneous equations. Using these reduced form parameters, we then write out the joint likelihood of Poisson parameters across the therapeutic classes. Here we assume the marginal distributions of the Poisson variables are independent conditional on the normal error terms, i.e. the η_{ijt} in Eq. (4) (cf. Aitchison and Ho (1989)).

To verify whether the above approach allows us to recover the structural parameters, we construct a stylized example with only three brands A, B and C. Each follows a Poisson-Lognormal distribution,[13] with parameters λ_A, λ_B and λ_C. That is $d_A \sim \text{Poisson}(\lambda_A)$, $d_B \sim \text{Poisson}(\lambda_B)$ and $d_C \sim \text{Poisson}(\lambda_C)$. Denote the log of Poisson distribution parameters as $l\lambda_A = \log(\lambda_A)$, $l\lambda_B = \log(\lambda_B)$ and $l\lambda_C = \log(\lambda_C)$, that are specified as follows.

$$\begin{aligned} l\lambda_A &= \alpha_{A0} + \alpha_{A1}Z_1 + \alpha_{A2}Z_2 + && \beta_{AB}l\lambda_B + \beta_{AC}l\lambda_C + \in_A \\ l\lambda_B &= \alpha_{B0} + \alpha_{B1}Z_3 + \alpha_{B2}Z_4 + \beta_{BA}l\lambda_A && + \beta_{BC}l\lambda_C + \in_B \\ l\lambda_C &= \alpha_{C0} + \alpha_{C1}Z_5 + \alpha_{C2}Z_6 + \beta_{CA}l\lambda_A + \beta_{CB}l\lambda_B && + \in_C \end{aligned} \tag{16}$$

[13] For the advantages of using Poisson-Lognormal distribution relative to regular Poisson distribution, please refer back to the paper. The method discussed in the following applies to the regular Poisson model as well.

Where $\{\alpha A\}$, $\{\alpha B\}$ and $\{\alpha_C\}$ are the structural parameters for the exogenous variables,[14] which are noted as Z_1,Z_2,...Z_5,Z_6. $\{\beta_A\}$, $\{\beta_B\}$ and $\{\beta_C\}$ are the parameters for the endogenous variables, two in each equation. $\{\in_A,\in_B,\in_C\}$ are assumed to follow a Multivariate Normal distribution. Because in Eq. (16), the log of each Poisson parameter is a function of the other Poisson parameters, we cannot write the likelihood function of the observed data d_A,d_B and d_C directly using these three equations, unless we can write out a multivariate Poisson distribution function. Given the difficulty of doing that, we develop the following method, by writing out the likelihood function using the reduced form.

Before solving the three simultaneous Eq. in (16), we can rewrite them into Matrix format, which is

$$\alpha Z + \beta \Lambda + \in = 0 \tag{17}$$

α is a 3×7 matrix, containing all the parameters for the exogenous variables, which is

$$\alpha = \begin{bmatrix} \alpha_{A0} & \alpha_{A1} & \alpha_{A2} & 0 & 0 & 0 & 0 \\ \alpha_{B0} & 0 & 0 & \alpha_{B1} & \alpha_{B2} & 0 & 0 \\ \alpha_{C0} & 0 & 0 & 0 & 0 & \alpha_{C1} & \alpha_{C2} \end{bmatrix}$$

Z is a 7×1 column vector, consisting of all the exogenous variables, in addition to 1 for intercept,

$$Z = \begin{bmatrix} 1 \\ Z_1 \\ \cdots \\ Z_6 \end{bmatrix}$$

β is a 3×3 matrix, with −1 in the diagonal, and the rest filled by the parameters for the endogenous variables

$$\beta = \begin{bmatrix} -1 & \beta_{AB} & \beta_{AC} \\ \beta_{BA} & -1 & \beta_{BC} \\ \beta_{CA} & \beta_{CB} & -1 \end{bmatrix}$$

Λ is a 3×1 column vector, consisting of all the endogenous variables

$$\Lambda = \begin{bmatrix} l\lambda_A \\ l\lambda_B \\ l\lambda_C \end{bmatrix}$$

[14] For identification purposes, at least one exogenous variable is needed in each equation.

Table 13 Estimation results from the generated data

		Intercept α_0	First exogenous variable α_1	Second exogenous variable α_2	Endogenous variable A β_A	Endogenous variable B β_B	Endogenous variable C β_C
Brand A	True Value	0.2	0.2	−0.4		0.3	−0.2
	Full model	0.2074	0.2048	−0.4129		0.3298	−0.2765
		(0.13,0.28)	(0.19,0.23)	(−0.46,−0.37)		(0.23,0.44)	(−0.49,−0.01)
	Ignoring cross-category detailing	0.3424	0.1557	−0.3151			
		(0.30,0.38)	(0.14,0.17)	(−0.33,−0.30)			
Brand B	True Value	0.3	0.3	−0.2	0.5		0.1
	Full model	0.3039	0.2988	−0.2002	0.4943		0.104
		(0.25,0.38)	(0.27,0.34)	(−0.23,−0.17)	(0.40,0.62)		(−0.14,0.27)
	Ignoring cross-category detailing	0.5981	0.2556	−0.1711			
		(0.55,0.64)	(0.24,0.29)	(−0.19,−0.15)			
Brand C	True Value	0.1	−0.2	0.2	0.4	0.3	
	Full model	0.0997	−0.1857	0.2038	0.3974	0.3086	
		(0.03,0.16)	(−0.28,−0.07)	(0.10,0.32)	(0.35,0.46)	(0.23,0.39)	
	Ignoring cross-category detailing	0.4929	−0.1837	0.2042			
		(0.46,0.53)	(−0.29,−0.09)	(0.10,0.31)			

Table 14 For the covariance matrix

		Brand A	Brand B	Brand C
Brand A	True Value	0.4	0.1	0.1
	Full model	0.4498	0.1139	0.1299
		(0.33,0.57)	(0.05,0.20)	(0.03,0.21)
	Ignoring cross-category detailing	0.536	0.543	0.532
		(0.51,0.58)	(0.50,0.57)	(0.49,0.56)
Brand B	True Value	0.1	0.3	0.2
	Full model	0.1139	0.3173	0.1912
		(0.05,0.20)	(0.22,0.49)	(0.10,0.31)
	Ignoring cross-category detailing	0.5430	0.9723	0.9525
		(0.50,0.57)	(0.92,1.03)	(0.91,1.00)
Brand C	True Value	0.1	0.2	0.4
	Full model	0.1299	0.1912	0.4050
		(0.03,0.21)	(0.10,0.31)	(0.36,0.45)
	Ignoring cross-category detailing	0.5320	0.9525	1.1964
		(0.49,0.56)	(0.91,1.00)	(1.14,1.27)

$\in$ is a 3×1 column vector, consisting of all the random terms

$$\in = \begin{bmatrix} \in_A \\ \in_B \\ \in_C \end{bmatrix}$$

Solving these three simultaneous equations in Eq. (17), we can get the reduced form solution of each Poisson parameter, which is

$$\Lambda = -\beta^{-1}(\alpha Z + e) \tag{18}$$

In Eq. (18), the right hand side of each equation has only the exogenous variables, therefore we can write out the joint likelihood function of the three Poisson variables as $f(d_A,d_B,d_C) \sim \text{Poisson}(\lambda_A)-\text{Poisson}(\lambda_B)$, where $\lambda_A = \exp(l\lambda_A)$, $\lambda_B = \exp(l\lambda_B)$ and $\lambda_C = \exp(l\lambda_C)$.

In estimating the structural parameters using the Metropolis-Hastings method, we just need to (1) draw the structural parameters from the proposal distribution, (2) conduct the transformation as specified in Eq. (18), then (3) plug in the Poisson parameters λ_A,λ_B and λ_C into the Poisson likelihood.

To summarize, we generate the MCMC draws for the structural parameters, but using the likelihood function derived with the reduced form parameters. In implementation, the only thing different for this method is that we need to add the transformation in Eq. (18) before using the likelihood function derived from the Poisson distribution.

We conducted a simulation study using this example with three brands. We simulate 50 sets of data, each with 5000 observations. Tables 13 and 14 presents the

mean of the estimation results across the 50 sets of data for both brands. We also provide the true values for both model parameters and the covariance matrix of the Normal error term. The first five rows are for the parameters related with the first brand, the next five rows for the second brand, and the last five rows for the third brand. For each brand, we list the true values, followed by the estimation results across the 50 data sets. Finally we list the (2.5%,97.5%) percentiles in parentheses. The same format applies to both tables. The results show that our method is able to recover all the parameters, including the model parameters, and the covariance matrix of the Normal error terms. This gives us the confidence to proceed with this approach using our data.

In addition, to demonstrate the importance of incorporating the other brand's effect into the model, we also estimate a comparison model by ignoring the other brand's detailing. The estimation results are listed in the last two rows for each brand. It shows that all the parameter estimates are biased when the simultaneity of the decisions are ignored in the model. This is true for both the parameter estimates and the covariance matrix estimates of the error terms.

This method can be easily extended to the general case of M ($M>3$) companies marketing in N ($N>3$)categories. Here, a set of M simultaneous equations needs to be solved where each set comprises of a number of simultaneous equations equal to number of categories marketed by each company

References

Ainslie, A., & Rossi, P. E. (1998). Similarities in choice behavior across product categories. *Marketing Science, 17*(2), 91.

Aitchison, J., & Ho, C. H. (1989). The multivariate poisson-lognormal distribution. *Biometrika, 76*(4), 643–653.

Albers, S., Mantrala, M. K., & Sridhar, S. (2010). Personal selling elasticities: A meta-analysis. *Journal of Marketing Research, 47*(5), 840–853.

Avidan, B., Sonnenberg, A., Giblovich, H., & Sontag, S. J. (2001). Reflux symptoms are associated with psychiatric disease. *Alimentary Pharmacology & Therapeutics, 15*(12), 1907–1912.

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

Brijs, T, Swinnen, G., Vanhoof, K., Wets, G. (1999). "Using association rules for product assortment decisions: A case study." *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 254–60

Cameron, C., & Trivedi, P. K. (1998). *Regression analysis of count data*: Cambridge University Press

Chib, S., Seetharaman, P. B., & Strijnev, A. (2002). Analysis of multi-category purchase incidence decisions using IRI market basket data. *Advances in Econometrics, 16*, 57–92.

Chintagunta, P. K., Erdem, T., Rossi, P. E., & Wedel, M. (2006). Structural modeling in marketing: Review and assessment. *Marketing Science, 25*(6), 604–616.

Dong, X., Manchanda, P., & Chintagunta, P. K. (2009). Quantifying the benefits of individual level targeting in the presence of firm strategic behavior. *Journal of Marketing Research, 46*(2), 207–221.

Dubé, J.-P., Hitsch, G. J., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The Rand Journal of Economics, 41*(3), 417–445.

Elrod, T., Russell, G. J., Shocker, A. D., Andrews, R. L., Bacon, L., Bayus, B. L., et al. (2002). Inferring market structure from customer response to competing to complementary products. *Marketing Letters, 13*(3), 221–232.

Erdem, T. (1998). An empirical analysis of umbrella branding. *Journal of Marketing Research, 35*(3), 339–351.

Erdem, T., & Sun, B. (2002). An empirical investigation of the spillover effects of advertising and sales promotions in umbrella branding. *Journal of Marketing Research, 39*(4), 408–420.

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association, 85*, 972–985.
Gönül, F., Carter, F., Petrova, E., & Srinivasan, K. (2001). Promotion of prescription drugs and its impact on physician choice behavior. *Journal of Marketing, 65*(3), 79–90.
Greene, W. H. (2008). *Econometric analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
Gupta, S., & Steenburgh, T. J. (2008). "Allocating marketing resources." *HBS Working paper*
Hansen, K., Singh, V., & Chintagunta, P. (2006). Understanding store-brand purchase behavior across categories. *Marketing Science, 25*(1), 75–90.
IMS Health. (2005). "2004 year-end U.S. Prescription and sales information and commentary," http://www.imshealth.com/ims/portal/front/articleC/0,2777,6599_18731_69890098,00.html, Accessed on 3/25/2005
Janakiraman, R., Dutta, S., Sismeiro, C., & Stern, P. (2008). Physicians' persistence and its implications for their response to promotion of prescription drugs. *Management Science, 54*(6), 1080–1093.
Jansson, C., Nordenstedt, H., Wallander, M.-A., Johansson, S., Johnsen, R., Hveem, K., et al. (2007). Severe gastro-oesophageal reflux symptoms in relation to anxiety, depression and coping in a population-based study. *Alimentary Pharmacology & Therapeutics, 26*(5), 683–691.
Jones, D. S. (2002). *Pharmaceutical statistics* (1st ed.). London, UK: Pharmaceutical Press.
King, G. (1989). A seemingly unrelated poisson regression model. *Sociological Methods & Research, 17*(3), 235–255.
Leeflang, P. S. H., & Wittink, D. R. (1992). Diagnosing competitive reactions using (aggregated) scanner data. *International Journal of Research in Marketing, 9*, 39–57.
Leeflang, P. S. H., & Wittink, D. R. (1996). Competitive reaction versus consumer response: Do managers overact? *International Journal of Research in Marketing, 13*, 103–119.
Li, S., Liechty, J. C., Montgomery, A. L. (2009). "Modeling category viewership of web users with multivariate count models." *Working paper*
Lilien, G., Kotler, P., Moorthy, K. S. (1995). *Marketing models* (1st ed.): Prentice Hall
Ma, Y., Seetharaman, P. B., Narasimhan, C. (2009). "Modeling dependencies in brand choices across complementary categories." *Working paper*
Manchanda, P., Ansari, A., & Gupta, S. (1999). The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing Science, 18*(2), 95.
Manchanda, P., Rossi, P. E., & Chintagunta, P. K. (2004). Response modeling with nonrandom marketing-mix variables. *Journal of Marketing Research, 41*, 467–478.
Matta, E., & Mehta, N. (2001). "Using the internet to maximize the effectiveness of pharmaceutical marketing-direct to physician marketing." CCG. XM Strategic Services Team
Munkin, M. K., & Trivedi, P. K. (1999). Simulated maximum likelihood estimation of multivariate mixed-poisson regression models, with application. *Econometrics Journal, 2*(1), 29.
Nair, H. (2007). Intertemporal price discrimination with forward-looking consumers: Application to the US market for console video-games. *Quantitative Marketing and Economics, 5*(3), 239–292.
Riphahn, R. T., Wambach, A., & Million, A. (2003). Incentive effects in the demand for health care: a bivariate panel count data estimation. *Journal of Applied Econometrics, 18*(4), 387–405.
Rossi, P. E., Allenby, G. M., & McCulloch, R. (2006). *Bayesian statistics and marketing*. England: John Wiley & Sons Inc.
Russell, G., Bell, D., Bodapati, A., Brown, C., Chiang, J., Gaeth, G., et al. (1997). Perspectives on multiple category choice. *Marketing Letters, 8*(3), 297–305.
Russell, G. J., Ratneshwar, S., Shocker, A. D., Bell, D., Bodapati, A., Degeratu, A., et al. (1999). Multiple-category decision-making: Review and synthesis. *Marketing Letters, 10*(3), 319–332.
Seetharaman, P. B., Siddhartha, C., Andrew, A., Peter, B., Tat, C., Sachin, G., et al. (2005). Models of multi-category choice behavior. *Marketing Letters, 16*(3–4), 239.
Sinha, P., & Zoltners, A. A. (2001). Sales-force decision models: Insights from 25 years of implementation. *Interfaces, 31*(3), S8–S44.
Song, I., & Chintagunta, P. K. (2006). Measuring cross-category price effects with aggregate store data. *Management Science, 52*(10), 1594–1609.
Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association, 82*, 528–550.
Wiklund, I., & Butler-Wheelhouse, P. (1996). Psychosocial factors and their role in symptomatic gastroesophageal reflux disease and functional dyspepsia. *Scandinavian Journal of Gastroenterology, 31*(S220), 94–100.
Wosinska, M. (2002). "Just what the patient ordered? Direct-to-consumer advertising and the demand for pharmaceutical products." *Working Paper*