Heterogeneous Learning and the
Targeting of Marketing Communication for New Products

Sridhar Narayanan¹

August 15, 2004

¹ Sridhar Narayanan is a Doctoral Student at the University of Chicago. Address for correspondence: University of Chicago, Graduate School of Business, 1101 East 58th Street, Chicago, IL 60637. Email: sridhar.narayanan@gsb.uchicago.edu.
Abstract

New product launches are often accompanied by extensive marketing communication campaigns. Firms’ allocation decisions for these marketing communication expenditures have two dimensions – across consumers and over time. What makes this problem hard in the case of new products is that consumers are uncertain about the quality of new products and learn about them through marketing communication. Further, different consumers may have different rates of learning about product quality, i.e. there may be heterogeneous learning. Thus, consumer responsiveness to marketing communication would vary along two dimensions. For each consumer, this responsiveness would vary over time, as she learns about product quality. Across consumers, there would be differences in responsiveness in each time period. For optimal allocation of marketing communication across both consumers and time, firms would need estimates of how responsiveness to marketing communication varies across consumers and over time.

Past studies in this area have typically studied one of these two dimensions in which responsiveness varies. They have either looked at heterogeneity in responsiveness across agents or the variation in responsiveness over time. In the context of new products, past research has looked at how consumer learning about product quality causes responsiveness to vary over time. However, there is no study that we are aware of that allows for heterogeneous learning rates, i.e heterogeneity in how consumers learn over time. In this study, we develop the methodology for estimating individual-level parameters of learning for consumers that differ on their learning processes and use a rich panel dataset that allows us to estimate these parameters of the model.

To obtain individual-level estimates of learning, we add a hierarchical Bayesian structure to the Bayesian learning model. We exploit the natural hierarchy in the Bayesian learning process to incorporate the learning model within the hierarchical Bayesian model. We use data augmentation, coupled with the Metropolis Hastings algorithm to make inferences about individual-level parameters of learning. We conduct this analysis on a unique panel dataset of physicians, where we observe prescription decisions and detailing (salesforce efforts) at the individual physician-level for a new prescription drug category.

Our results show that there is significant heterogeneity across physicians in their rates of learning about the quality of new drugs. We also find that there are asymmetries in the temporal evolution of responsiveness of physicians to detailing – physicians who are more responsive to detailing in early periods are less responsive later on and vice versa. These finding have interesting implications for targeting of detailing across physicians and over time. We find that firms could increase their revenues if they took these temporal and cross-sectional differences in responsiveness into account while deciding their allocations of detailing.
1. Introduction

Marketing of new products is a very important part of firms’ problem. A study of 700 firms by Booz, Allen and Hamilton (cited in Urban and Hauser 1993) reported that new products accounted for 28% of the growth. Another cross-industry study by Wind, Mahajan and Bayless (cited in Urban and Hauser 1993) found that 25% of sales were accounted for by new products. A recent unpublished McKinsey study also found that firms spend about 50% of their marketing budgets on average promoting new products. Various research questions related to new products have been extensively studied in the marketing literature. There has been a particularly large and rich literature looking at various aspects of diffusion of new products (see Mahajan, Muller and Bass 1990 for a review).

A distinguishing feature of new products, compared to mature products, is the presence of considerable uncertainty about their quality. Firms use marketing communication to disseminate information about their new products and consumers use this communication to learn about the new products and reduce their uncertainty about their quality. As a result of this learning, the role of marketing communication is itself likely to be changing over time and firms would need to decide how to allocate their launch communication over time. Further, consumers may differ in how they learn about these new products through the same communication. Some consumers may learn faster than others about the quality of these new products than other consumers. In the presence of this heterogeneity in learning, firms would need to decide how to allocate their resources across consumers. In sum, firms would need to decide how to allocate their resources both over time and across consumers in the presence of heterogeneous learning.

Prescription drug categories are particularly well suited to study the problem of allocation of resources for launch communication in the presence of heterogeneous learning. First, uncertainty about drug quality is particularly relevant in the case of prescription drugs. In spite of an extensive process of clinical trials before the launch of new drugs, there is still considerable uncertainty about their efficacy, their side effects and any risks associated with administering the drug. For instance, the Food and Drug
Administration, without whose approval prescription drugs cannot be marketed, has the following to say about new drugs:

“The practical size of pre-marketing clinical trials means that we cannot learn everything about the safety of a drug before we approve it. Therefore, a degree of uncertainty always exists about the risks of drugs.”

Second, pharmaceutical firms spend a large amount of money on communication directed towards physicians. In the year 2000, close to 8.5 billion dollars was spent by the pharmaceutical industry in the United States on communication directed to physicians (Wittink 2002, Neslin 2001). Most of this is spent on detailing, which involves personal sales calls made by salespersons of the pharmaceutical firm on physicians. Pharmaceutical firms spend particularly large sums of money when launching new drugs. For instance, in the first three months after launching the block bustor Erectile Dysfunction drug Viagra, Pfizer had conducted about 270,000 detailing contacts for the drug (IMS 1998) amongst the estimated 100,000 active physicians for the category in the US. Thus, allocation of these launch communication is an important decision for firms.

Third, since detailing is a personal interaction between a physician and the firm’s representatives, it is allocated at the individual physician level. Firms need to decide how many calls to make to each individual physician and when to make these calls. The problem of resource allocation is a very complicated one and hence interesting from a research point of view.

Discussions with industry experts and pharmaceutical consulting firms suggest that pharmaceutical firms typically allocate detailing efforts across physicians using simple rules based on physicians’ prescription volumes (see Manchanda, Rossi and Chintagunta 2003 for a discussion of this). Physicians are typically classified into deciles based on their total prescription volumes in the category and detailing targets are set for each of these deciles. In general, higher prescribing physicians receive a larger number

of detailing calls. However, such simple volume-based rules are unlikely to be very useful. For instance, a typical industry practice is to continue to detail to high volume physicians till access is denied to the detailers. Since the response to detailing is likely to show diminishing returns, such a strategy is likely to be suboptimal, since there might be other low-volume physicians for whom detailing might give greater bang for the buck. Manchanda, Rossi and Chintagunta (2003) find specifically that physicians are not detailed optimally. However, firms can optimize their allocations by setting detailing efforts so as to equalize the marginal effects of detailing across physicians. This would involve estimation of the response to detailing at the individual physician level.

Past research has looked at the issue of targeted promotions in general (c.f. Rossi, McCulloch and Allenby 1996) and in the specific context of pharmaceuticals (Manchanda, Rossi and Chintagunta 2003). New product introductions, on the other hand involve significant differences from mature product categories, which have been the focus of most of the previous research on targeted marketing. New products involve significant uncertainty about product quality in the minds of consumers, or in the case of pharmaceuticals, in the minds of physicians. Marketing activities play a significant role in reducing this uncertainty (Narayanan, Manchanda and Chintagunta 2003). This uncertainty reduction (or in other words learning) role of marketing communication induces significant amount of dynamics in the way these promotions affect sales (or prescriptions in the case of pharmaceuticals). Therefore, optimizing the allocations of detailing efforts for new drug introductions would involve estimating individual physician level response to detailing in the presence of these dynamic effects.

Learning is likely to differ significantly by physician. Different physicians have different levels of training, ability and different environments (for instance, some are in research hospitals, while others have rural practices). Thus, heterogeneity in learning across physicians is an important issue to incorporate, particularly when looking at targeting of detailing for new products. However, prior research on learning has assumed homogenous learning effects across consumers or physicians. This has been led both by data limitations, particularly in the case of studies that have used aggregate data as well
as limitations in the available estimation methodologies. Past research on learning in pharmaceutical categories has been limited by the absence of panel data of physicians with detailing efforts recorded for all the drugs in a category. Data has been available at the aggregate level or in the case of individual data, detailing data is available for only one firm in the category. Further, past studies using learning models have used frequentist methods for estimation, making the estimation of individual physician level parameters infeasible. A Bayesian approach is a natural way to estimate individual physician-level parameters, but there have been no studies that have extended models of learning to Bayesian methods of estimation.

In this study, we use a unique panel dataset of physicians, which contains prescription and detailing information for all drugs in a category to estimate individual physician-level effects in the presence of learning. We allow for both informative effects of detailing (i.e. the effect that comes about through the process by which physicians learn about new drugs) as well as the persuasive effects of detailing (all other effects of detailing apart from the informative effect). We develop a methodology to estimate individual physician level effects using Markov Chain Monte Carlo methods. We use this methodology to estimate the model parameters and use these estimates to demonstrate how firms can optimize their detailing allocations across physicians.

2. Related Literature
This study is broadly related to three streams of research in the literature – pharmaceutical promotions, learning and targeted marketing. There have been numerous studies on pharmaceutical promotions and recent ones that look at learning in the context of pharmaceutical promotions, but none of these look at targeting of these pharmaceutical promotions in the presence of learning. Our contribution would be to study the heterogeneity in physician learning about new drugs and use this to analyze the targeting of detailing expenditures for new drug introductions.
There has been a large amount of research in the area of pharmaceutical demand in marketing, economics as well as medical sciences. Early studies in the marketing literature (Parsons and Vanden Abeele, 1981; Lilien, Rao and Kalish 1981) studied the effect of sales force effort on sales using aggregate data. Recent research (Kamakura and Kossar, 1998; Manchanda and Chintagunta 2000; Gonul et al. 2001; Wosinska 2002 and Manchanda, Rossi and Chintagunta 2003) has used panel data to investigate the effect of detailing and/or DTC advertising on pharmaceutical demand. There has also been research that has specifically investigated informative and persuasive effects of pharmaceutical promotions (Leffler 1981; Hurwitz and Caves 1988; Rizzo 1999; Currie and Park 2002; Narayanan, Manchanda and Chintagunta 2003). The broad consensus in this literature is that detailing positively affects prescriptions by physicians. Some of these studies find evidence for significant amount of heterogeneity in physician response to detailing and find that accounting for this heterogeneity and using it to optimize their detailing allocations is potentially beneficial to pharmaceutical firms.

The second stream of research that is relevant to this study is the literature on learning. Early studies that used a Bayesian learning process include Stoneman (1981), Jensen (1982), Meyer and Sathi (1985) and Roberts and Urban (1988). Erdem and Keane (1996) was a pioneering paper in marketing that used a model of Bayesian learning to incorporate informative effects of advertising. Since then, there has been a growing interest in problems involving learning. Crawford and Shum (2003), Coscelli and Shum (2003), Ching (2002), Anand and Shachar (2001), Currie and Park (2002), Ackerberg (2003), Narayanan, Manchanda and Chintagunta (2003), Byzalov and Shachar (2003) and Mukherjee (2002) all apply Bayesian learning models to a variety of contexts. These studies find evidence that there is significant amount of learning and uncertainty reduction through advertising and product experience. Ackerberg (2003), Currie and Park (2002), Narayanan, Manchanda and Chintagunta (2003) and Byzalov and Shachar (2003) specifically address the issue of informative and non-informative roles of advertising or promotional activity and attempt to estimate both of them from the data. Narayanan, Manchanda and Chintagunta (2003) find evidence for the presence of both these roles in the case of a new pharmaceutical category, while the other studies find
evidence for only the informative effect. Akçura, Gönül and Petrova (2004) adopt an alternative reduced form approach to learning, using a Kalman-filter based model instead of Bayesian learning model to study patients’ learning about over-the-counter drugs. They find variation in learning rates across patients and drugs.

Targeted promotions have interested researchers in marketing in recent years. Targeting promotions to segments of consumers has long been an industry practice as well as a topic for research. Numerous studies have looked at price discrimination (cf. Villas Boas 1999; Fudenberg and Tirole 2000), targeted coupons (Shaffer and Zhang 1995) and the targeting of advertising (cf. Iyer, Soberman and Villas-Boas 2002) to different segments of consumers. Rossi, McCulloch and Allenby (1996), Manchanda, Rossi and Chintagunta (2003) are instances of research that study the targeting of promotional activities towards individual consumers. These involve the estimation of individual-level response parameters using Bayesian methods. This stream of research demonstrates that firms can obtain significant benefits by targeting their promotions.

There has been a significant amount of study about the allocation of salesforce efforts in the salesforce literature. This study is also related to that literature, since we focus on the targeting of detailing and detailing is a kind of salesforce activity. Early studies like Lodish (1971); and Lodish (1980) used decision calculus based methods to arrive at efficient allocation of salesforce efforts across customer accounts. Lodish (1975); Zoltners (1976); Zoltners and Sinha (1983); Rangaswamy, Zoltners and Sinha (1990) have studied how to design optimal salesforce territories. Horsky and Nelson (1996) study the problem of optimal salesforce size and productivity. Mantrala, Sinha and Zoltners (1992) study the effect of resource allocation rules on profitability.

Thus, there is a relatively large stream of literature that studies allocation of resources, in particular of salesforce efforts, another stream that studies learning and a third stream that studies the responsiveness of physician prescription behavior to marketing communication. However, the study of new products and specifically new drugs, on which firms spend a significant part of their promotional budget is still an open
area for research in the literature. This study attempts to fill this gap by studying the targeting of detailing to individual physicians in the case of new drugs.

3. Data

The dataset used for the empirical analysis in this study is from the category of prescription drugs known as ‘Erectile Dysfunction Drugs’. The drugs in this category are prescribed to treat ‘Erectile Dysfunction’ (ED) amongst adult men. About 15 to 30 million men in the United States are believed to be affected by the condition. There is only one category of oral drugs that can treat this condition and currently, there are three drugs that have been approved by the Food and Drug Administration (FDA). The first drug to be approved in the category was Viagra, marketing by Pfizer, and was approved by FDA in March 1998. Levitra, marketed jointly by GSK Pharmaceuticals and Bayer was approved in August 2003 and Cialis, marketed by Eli Lilly, was approved in November 2003. No further drugs have been approved after Cialis.

The data are at the physician level and consists of a panel of 900 physicians in the United States. The panel is a representative sample of the universe of physicians, balanced across geographic regions, physician specialties and prescription volume. For these physicians, we have observations of prescriptions written by them for their patients and also of detailing calls made by representatives of pharmaceutical firms (detailers). These data are collected directly from the physician, using a Personal Digital Assistant based diary method. All detailing calls made to the physician are recorded. Prescription decisions are recorded on two days of each week, with the specific days of the week distributed randomly across physicians and rotated every week. Totally, we have 15320 prescription observations and 16700 observations of detailing calls in the dataset.

For each physician, we have information about their specialty (General Practitioner or specialist and the specific specialty). In each prescription observation, we observe the drug that was prescribed to the patient. Unlike most existing data sources, which collect prescription data from pharmacies or insurance companies, our dataset actually captures
the physician’s decision. This is because pharmacy data, for instance, may not be able to capture the fact that the drug actually filled out in the pharmacy may sometimes be different from that prescribed by the physician. A unique feature of the prescription data in our dataset is that we observe if a patient requested a drug. This is recorded by the physician at the time of the consultation by the patient. Every time a detailer walks into the physician’s office, the physician records the drug that is detailed in that call.

4. Model Development
In this section, we discuss the model used in this study. We first describe the model structure and lay out the underlying assumptions. We then describe the model specification in detail.

4.1 Prescription Decision
We assume that the physician is the sole decision maker for which drug is to be prescribed to a patient. However, we shall allow for the patient to influence the prescription decision of the physician. The data allows us to incorporate this influence of the patient and we shall describe this in more detail in the discussion on the model specification. In general, the prescription decision also involved influence by intermediaries like insurance firms, HMOs, Medicare, Medicaid etc through the inclusion or exclusion of specific drugs in their formularies. However, in the specific category of Erectile Dysfunction drugs, this is not as big a concern as in many other categories since these drugs are not on the formularies of most HMOs and are not supported by Medicare. Also, there are no generic substitutes to the brand name drugs. Hence, it is reasonable to assume that the physician is the decision maker.

The physician is assumed to value the health of the patient and her preferences are assumed to map into a utility function over the space of treatment options. The physician may value the health of the patient both out of a sense of professional integrity and also to avoid malpractice litigation in the future. A desire to build and maintain a reputation may also motivate a physician to desire the best treatment for her patients. The physician
is assumed to choose the option that provides the greatest utility. The drugs in this Erectile Dysfunction category are seen as substitutes and there are no instances when multiple drugs within the category are prescribed to the same patient. Further, the condition is not usually treated using non-drug options. Thus, the physician makes a discrete choice amongst the drugs within the category.

The physician is assumed to be a utility maximizer, i.e. her preferences over the space of drugs are assumed to map into a utility function defined on these drugs. Further, she is assumed to be uncertain about the quality of a new drug. The term ‘quality’ refers to a scalar that summarizes aspects of the drug like its efficacy (how well it treats the condition for which it is prescribed), side effects, etc. This quality enters the utility function of the physician, as described in the section on model specification. Since the physician is uncertain about this quality, she is in turn uncertain about the utility function also. Thus, we make the assumption that she is an expected utility maximizer, i.e. she chooses the alternative that provides her the greatest expected utility.

We also make the assumption that the physician is an expected utility maximizer, i.e. he maximizes the utility at each prescription occasion and does not consider future prescription occasions. Potentially, one may argue that physicians may be forward looking. However, we assume this away for analytical tractability. Conceptually, it should be possible to incorporate forward looking behavior into this model, but that is left for future research.

4.2 Learning
As described earlier, physicians are assumed to be uncertain about the quality of drugs. At this stage, it is important to clarify that there could be two levels of uncertainty about drug quality. The first level of uncertainty is regarding the mean quality of the drug across the patient base of the physician. A second level of uncertainty could be regarding the match between a specific patient and a specific drug. In this study, we focus on the uncertainty about the mean quality of the drug and learning about this mean quality. There are two reasons for this focus. First, uncertainty about the mean quality of the drug
is likely to be important in the case of new drugs. Second, we do not have data that would allow us to model the learning about the patient-drug match for a specific patient, since we do not have repeated observations of prescription decisions for the same patient. We only observe the first prescription occasion for a particular patient. Henceforth, when we refer to quality, we would be referring to the mean quality of the drug across the patient base of the physician.

Physicians are assumed to learn about the quality of a new drug through two main sources – the feedback they receive from their own patients who were prescribed the drug in the past (which we shall henceforth refer to as patient feedback), and marketing communication directed towards them by pharmaceutical firms. Marketing communication directed towards physicians is primarily in the form of detailing.

Physicians are assumed to be Bayesian updaters. Thus, at any given period of time, they have some prior belief about the efficacy of the drug, which they update with information from patient feedback and marketing communication received at that time using Bayes Rule to form their posterior belief. This posterior belief in turn becomes the prior belief in the next time period and the process is repeated then. Whenever a patient walks in and a prescription decision is to be made, the physician uses the most updated belief about the quality of the drug in the decision.

We make the following assumptions regarding the learning process. In our dataset, we do not observe when a particular patient gives feedback to the physician. We assume therefore, that every prescription results in feedback and that the physician receives this feedback one month after the prescription is written. The fact that every prescription results in a feedback is not a critical assumption. The results of the analysis would not be affected as long as we are willing to believe that a fixed proportion of past prescriptions would result in feedback. If that were the case, the parameters related to this feedback would be scaled by this proportion of past prescriptions. However, marginal effects and elasticities would remain unaffected, as we shall see in subsequent sections.
Patient feedback and marketing communication are assumed to provide unbiased but noisy signals about the true quality of a drug. In our operationalization, we assume that these signals are normally distributed, with the mean being the true quality. Physicians are assumed to see the realizations of these signals and know the variances of these signals but do not know the true mean. The assumption that detailing provides an unbiased signal needs some explanation. Firstly, the signal is what the physician perceives from the interaction with the detailer. Given that a physician has repeated contact with detailers, an underlying assumption that physicians are able to discount any bias in detailers’ messages would make the assumption of unbiasedness a valid one. Secondly, the information given in a detailing call is regulated by the Food and Drug Administration (FDA). Any visual aids and brochures used during a detailing call have to be pre-approved by the FDA and have to include negative as well as positive aspects of the drug. Thus, the information has to be factual and to that extent, the detailing call has to be unbiased.

Finally, we make the assumption that at the initial period (i.e. before they prescribe any drug or receive any detailing calls), physicians have a normally distributed prior belief about the drug, summarized by its mean and variance. This belief is assumed to be the same for all physicians for identification purposes. In general, this assumption that all physicians have the same belief can be relaxed to some extent. The underlying assumption is that there is some public knowledge set that all physicians draw from in this initial phase.

4.3 Informative and Persuasive effects of marketing communication
We allow for marketing communication to affect physician utility in two ways. We shall refer to the effect on utility through the learning process as the informative effect. Additionally, marketing communication can directly affect the utility of the physician. This effect, which we shall refer to as the persuasive effect may represent any prestige or image effects (cf. Becker and Murphy 1993, Ackerberg 2003) or reminder effects. We shall capture these effects in a reduced form manner by allowing a linear stock of
detailing calls to enter the utility function. In our empirical specification, we assume that
the count of detailing calls in the preceding month constitutes this stock for the direct
effect.

4.4 Heterogeneity
Physician-level heterogeneity is a critical aspect of our study. We are primarily
interested in estimating individual physician-specific parameters, both for the informative
and persuasive effects of marketing communication. Heterogeneity in response to
marketing variables like price and advertising has been studied extensively in the
literature. It has been well established that accounting for this heterogeneity is important
(cf. Rossi and Allenby 1999). The approach in the past was that heterogeneity was a
nuisance parameter that had to be dealt with, but was not the focus of the study.
However, recent studies in marketing have focused on estimating individual-level
parameters and have demonstrated how decisions of firms can be improved by focusing
on these differences between consumers. In the case of targeted marketing activities,
such an understanding of individual differences is critical to making optimal decisions,
for instance, optimal allocation of marketing effort. It is important in our specific case of
pharmaceuticals, since unlike many consumer categories, where marketing
communication takes the form of television advertising, which is less targetable, detailing
calls are physician-salesperson interactions and hence are entirely targeted at individual
physicians. Thus, modeling heterogeneity in physician response to detailing is critical.

In our setup, detailing (or more generally marketing communication) has two
effects – the informative effect and persuasive effect. Heterogeneity in the informative
effect, in our specification, is incorporated by allowing the detailing signals to have
different variances for different physicians. This implies that different physicians see
these signals with different levels of noise. This is a plausible assumption, considering
that different physicians have different levels of ability and training and pay different
levels of attention to information they receive from detailing calls. It is important to note
that a signal is the physician’s perception of the information given in a detailing call.
Thus, the same signal may be perceived by different physicians with different levels of
noise. Some physicians may be able to perceive the information received in the detailing call precisely while others may not be able to perceive it as precisely. The implication of this is that the former would learn through fewer detailing calls than the latter. Heterogeneity in the persuasive effect implies that different physicians have different levels of response to prestige or reminder effects for instance, and in our specification, this would manifest itself in heterogeneity in the coefficient for the detailing stock variable.

Similar to the heterogeneity in the informative effect of detailing, we also allow heterogeneity in the learning through patient feedback by allowing the variance on the feedback signals to differ by physician. Again, the underlying assumption is that physicians differ in their abilities and training levels and therefore, the degree of noise in patient feedback differs by physicians. The physicians who receive more precise signals of patient feedback learn faster than physicians who receive less precise signals.

Further, the true quality is also allowed to be heterogeneous. The reason why different physicians may have different true quality levels is that the patient base may be different for different physicians. For instance, some physicians may have typically older patients than other physicians. Since drugs differ on their efficacy and side effects for different types of patients, the true quality levels may also differ by patient base of physician.

5. Model Specification

5.1 Utility Function

When physician $i$ has to make a decision on which drug to prescribe at occasion $t$, she chooses the alternative $j$ that provides the greatest utility, with the utility function defined as

$$
\tilde{U}_{ijt} = f(\tilde{Q}_{ijt}) + X_{ijt} \beta_i + \epsilon_{ijt}
$$

where
\( \tilde{Q}_{ijt} \) is physician \( i \)'s belief about the mean efficacy of drug \( j \) at time \( t \) and is stochastic from the point of view of the physician

\( f(\cdot) \) is some function through which quality enters the utility

\( X_{ijt} \) is a row vector \((I \times K)\) of physician, drug and time (patient) specific variables, including variables that define patient-drug match

\( \beta_i \) is a column vector \((K \times 1)\) of physician specific sensitivity to these variables

\( \epsilon_{ijt} \) is an i.i.d. physician, drug and time specific shock and could include a patient and drug specific match value

Note that since we have an observation every time a patient walks into the office and since we do not have repeated observations for the same patient, the prescription occasion \( t \) is identical to patient \( t \) for the physician.

This utility is stochastic from the physician’s perspective because the belief about the quality \( \tilde{Q}_{ijt} \) is stochastic. But it must be remembered that \( \epsilon_{ijt} \) is not stochastic from the point of view of the physician. The physician is assumed to be an expected utility maximizer. This expected utility is

\[
U_{ijt} = E[\tilde{U}_{ijt}] = E[f(\tilde{Q}_{ijt})] + X_{ijt}\beta_i + \epsilon_{ijt}
\]

(2)

Whether the physician is risk averse or risk neutral depends on the exact specification of the function \( f(\tilde{Q}_{ijt}) \). In particular, if the physician is assumed to be risk-neutral, the function would be linear in \( \tilde{Q}_{ijt} \). If \( f(\tilde{Q}_{ijt}) \) is specified to be non-linear, the physician would not be risk-neutral, for instance, this function could be of the Constant Absolute Risk Aversion (CARA) form to allow for risk aversion.

5.2 Quality Evolution
Physicians are assumed to update their quality belief in each period based on signals they receive through detailing and through patient feedback they receive from past prescriptions. These signals are assumed to be normally distributed around the physician-specific true mean quality of the drug.

Assume that there are \( nd_{ijt} \) detailing signals at time \( t \) and the \( m^{th} \) signal is assumed to be given as

\[
\tilde{D}_{ijm} \sim N \left( Q_{ij}, \sigma^2_{Dij} \right)
\]

and that there are \( nf_{ijt} \) patient feedback signals feedback signal, and the \( m^{th} \) signal is given by

\[
\tilde{F}_{ijm} \sim N \left( Q_{ij}, \sigma^2_{Fij} \right)
\]

A series of unobserved signals that are normally distributed can be summarized by their sample mean, which is also normally distributed. We define these sample means as follows

\[
\tilde{D}_{ijt} = \frac{\sum_{m} \tilde{D}_{ijm}}{nd_{ijt}} \sim N \left( Q_{ij}, \frac{\sigma^2_{Dij}}{nd_{ijt}} \right)
\]

\[
\tilde{F}_{ijt} = \frac{\sum_{m} \tilde{F}_{ijm}}{nf_{ijt}} \sim N \left( Q_{ij}, \frac{\sigma^2_{Fij}}{nf_{ijt}} \right)
\]

The quality belief at time \( t=0 \) is assumed to be normally distributed, as given below

\[
\tilde{Q}_{ij0} \sim N \left( Q_{ij0}, \sigma^2_{ij0} \right)
\]
The physician is assumed to update his beliefs in a Bayesian manner, i.e. at any given period of time, he combines his prior belief about the quality of the drug with the information obtained through both detailing and feedback signals and applies Bayes Rule to form his posterior belief. Since the prior belief at time t=0 and all signals are assumed to be normally distributed, it turns out that the posterior belief at every time period is also a normal distribution. This posterior belief is given by

\[ \tilde{Q}_{ijt} \sim N\left(Q_{ijt}, \sigma^2_{ijt}\right) \]  

where

\[ Q_{ijt} = \frac{\sigma^2_{Qij(t-1)}}{\sigma^2_{Qij(t-1)}} Q_{ij(t-1)} + nf_{ijt} \frac{\sigma^2_{Qijt}}{\hat{F}_{ijt}} + nd_{ijt} \frac{\sigma^2_{Qijt}}{\sigma^2_{D_{ijt}}} \tilde{D}_{ijt} \]  

and

\[ \sigma^2_{Qij} = \frac{1}{\frac{1}{\sigma^2_{Qij(t-1)}} + nd_{ijt} + \frac{nf_{ijt}}{\sigma^2_{D_{ijt}}} + \frac{nf_{ijt}}{\sigma^2_{F_{ijt}}}} \]  

It is important at this stage to point out how heterogeneity in learning manifests itself in the model. Note that the variances of the detailing and feedback signals (\( \sigma^2_{D_{ijt}} \) and \( \sigma^2_{F_{ijt}} \) respectively), are physician specific. It can be seen from equation (9) that for a physician for whom \( \sigma^2_{D_{ijt}} \) is a large value, relatively lower weight is placed on the detailing signal \( \tilde{D}_{ijt} \) than for a physician with a low value of \( \sigma^2_{D_{ijt}} \), everything else remaining the same. The former physician is a slow learner from detailing than the latter. Similarly, a physician with a higher value of \( \sigma^2_{F_{ijt}} \) would be a slower learner from feedback than a physician with a lower value of this parameter. Thus, the variances of the detailing and feedback signals summarize the heterogeneity in learning across physicians.

5.3 Informative Effect of Detailing

The physician’s belief at the initial period (t=0) is given by equation (7). As he receives detailing and feedback signals, the belief changes. In particular, if we see equations (9)
and (10), it will be clear that the mean of the quality belief, $Q_{ij}$, evolves to the true mean quality $Q_j$ and the variance of the belief $\sigma_{Qj}^2$ evolves to 0. Focusing on the detailing signal, every detailing signal thus has an effect on the quality belief. This in turn affects the expected utility in any given period, which in turn influences the probability of prescribing the drug. Thus, detailing has an effect on prescription behavior through this process of learning. We shall refer to this effect as the informative effect of detailing.

It is important to note that the informative effect of detailing could positively or negatively affect the probability of prescribing a drug. If the physician’s initial quality belief about the drug is pessimistic – i.e. if the physician believes that the drug is of poorer quality than it actually is, detailing would increase the probability of prescribing the drug by reducing this pessimism of the physician. On the other hand, if the physician starts off by being overly optimistic, i.e. if he initially believes that it is of better quality than it actually is, the probability of prescribing the drug would reduce with detailing due to the informative effect.

It can also be seen from equations (9) and (10) that the effect of detailing on the physician’s quality belief is highest initially and reduces with every subsequent updation. This can be seen from the fact that the variance of the detailing signal $\sigma_{Qj}^2$ asymptotes towards zero with every updation and thus, the coefficient of the detailing signal in equation (9) asymptotes to zero. Thus, the informative effect of detailing is highest at the introductory phase of a drug and after the physician has learnt about the drug and reduced his uncertainty, this effect is negligible.

5.4 Persuasive Effect of Detailing
In the previous section, we discussed the informative effect of detailing that influences the physician’s prescription decision through the learning process. In particular, we found that this is highest initially and is negligible in later phases of a new drug’s life cycle. However, it is well documented that detailing has an effect on physicians’ prescription behavior even in the case of mature products (cf. Gonul et al 2001;
Manchanda et al. 2003). It has been suggested that this effect in the case of mature products may be because of an image or prestige role of detailing, or perhaps due to reminder effects. We shall refer to all effects of detailing except the informative effect defined in the previous section as the persuasive effect of detailing. It is important to note that these terms – informative effect and persuasive effect - are just labels to differentiate between the two effects and do not refer to the underlying constructs of information and persuasion.

The persuasive effect of detailing is captured in a reduced form manner by including a stock of detailing counts in the linear $X_{ijt}$ variable in the utility function (equation 1). The coefficient of this variable measures the persuasive effect of detailing. This effect would capture any role of detailing that remains unchanged over the product life cycle of the drug.

5.5 Patient Influence
The data allows us to account for patient influence on the prescription decision of the physician. We observe in the data if the patient requested a specific drug or not. We allow a dummy variable indicating whether the patient requested the drug or not in the linear $X_{ijt}$ variable in the utility function in equation (1). The coefficient of this variable captures the influence of the patient’s request on the prescription decision of the physician. In an indirect way, this also captures the effect of direct to consumer advertising (DTC), since DTC often asks a patient to talk to the doctor about the drug.

5.6 Hierarchical Bayesian Model
The objective of our empirical analysis is to suggest revenue-enhancing allocation plans for detailing. Since detailing has to be allocated at the individual physician level, the estimates of the effect of detailing also have to be obtained at the individual physician level. We thus specify our model as a Hierarchical Model in order to use Markov Chain Monte Carlo (MCMC) methods to estimate the individual-level parameters.
However, the challenge in specifying the model as a Hierarchical Bayesian Model is that the quality beliefs are unobserved. In a standard frequentist estimation of learning models (Erdem and Keane 1996), one could integrate out these unobserved quality beliefs by simulation methods. However, for using MCMC methods, we make use of a simple observation. From equation (9), it is clear that not just is the quality belief \( \tilde{Q}_{ijt} \) a stochastic variable, but even its mean \( Q_{ijt} \) is stochastic. This is because, from equation (9), \( Q_{ijt} \) is a function of two stochastic variables – the realizations of the detailing and feedback signals, \( \tilde{D}_{ijt} \) and \( \tilde{F}_{ijt} \) respectively. Further, since these two variables are assumed to have normal distributions, \( Q_{ijt} \) is a linear combination of normal variables, and therefore is also a normal variable.

In particular, we can derive the distribution of \( Q_{ijt} \), conditional on \( Q_{ij(t-1)} \) as

\[
Q_{ijt} | Q_{ij(t-1)} \sim N(\tilde{Q}_{ijt}, \nu_{ijt}^2)
\]

where

\[
\tilde{Q}_{ijt} = \frac{\sigma_{Q_{ijt}}^2}{\sigma_{Q_{ij(t-1)}}^2} Q_{ij(t-1)} + \left( \frac{n_f}{n_f^2} \frac{\sigma_{Q_{ijt}}^2}{\sigma_{F_i}^2} + \frac{n_d}{n_d^2} \frac{\sigma_{Q_{ijt}}^2}{\sigma_{D_i}^2} \right) Q_{ijt}
\]

and

\[
\nu_{ijt}^2 = \frac{n_f}{n_f^2} \frac{\sigma_{Q_{ijt}}^4}{\sigma_{F_i}^2} + \frac{n_d}{n_d^2} \frac{\sigma_{Q_{ijt}}^4}{\sigma_{D_i}^2}
\]

Note that from equation (10), the variance of the quality belief \( \sigma_{Q_{ijt}}^2 \) is not a stochastic variable, conditional on the parameters of the model. Given the parameters of the model, it is known deterministically. The main difference between the mean \( Q_{ijt} \) and variance of the quality belief \( \sigma_{Q_{ijt}}^2 \) is that while the former depends on the (unobserved) realizations of the detailing signal \( \tilde{D}_{ijt} \) and the feedback signal \( \tilde{F}_{ijt} \), the latter does not. Hence, in a frequentist estimation of a learning model, we integrate out the mean but not
the variance of the quality belief. Analogously, in a hierarchical model, we specify the mean of the quality belief as a level of the hierarchy, but not the variance.

Given that we can write the unobserved mean of the quality belief in any period as a random variable, conditional on the mean of the quality in the previous period, we thus have a natural hierarchy of quality beliefs

\[
Q_{ijt} \mid Q_{ij(t-1)} \sim N(\bar{Q}_{ijt}, \nu_{ijt}^2) \\
Q_{ij(t-1)} \mid Q_{ij(t-2)} \sim N(\bar{Q}_{ij(t-1)}, \nu_{ij(t-1)}^2) \\
\ldots \\
Q_{ij1} \mid Q_{ij0} \sim N(\bar{Q}_{ij1}, \nu_{ij1}^2)
\]  

(14)

We can then specify the rest of the hierarchical model as follows. We assume that the prescription choice follows a probit process. Thus, the random errors \(\epsilon_{ijt}\) of the utility function in equation (1) follow a multivariate normal distribution.

\[
\begin{bmatrix}
\epsilon_{i1t} \\
\vdots \\
\epsilon_{iht}
\end{bmatrix} \sim MVN\left(0, \Sigma\right)
\]  

(15)

The alternative that provides the greatest utility is chosen. Hence, \(U_{ijt}\) follows a truncated multivariate distribution, conditional on choice. If choice is given by the indicator variable \(I_{ijt}\) (which is 1 if brand \(j\) is chosen and 0 otherwise), the truncation is such that

\[
U_{ijt} > U_{iks}, I_{ijt} = 1, I_{iks} = 0 \forall k \neq j
\]  

(16)
Let the vector $\gamma_i$ (dimension K x 1) denote the individual level parameters of the model. These parameters are specified as a function of physician-level characteristics, as follows:

$$\gamma_i = \left[ \beta_i^T \ln \left( \sigma_{i,1}^2 \right) \ln \left( \sigma_{i,F}^2 \right) Q_{i1} \ldots Q_{ij} \right]^T \sim MVN \left( \Lambda Z_i, V_{\gamma} \right) \quad (17)$$

where

- $Z_i$ is a M x 1 column vector of physician characteristics, including a first element which has the value 1; and
- $\Lambda$ (K x M matrix) and $V_{\gamma}$ (K x K matrix) are parameters.

Thus, the hierarchical model can be specified as follows:

$$
\begin{align*}
U_{ij} & \mid I_{ij}, X_{ij}, Q_{ij}, \beta_i, \Sigma \\
Q_{ij} & \mid Q_{ij(t-1)}, nd_{ij}, I_{ij(t-1)}, \sigma_{D_i}^2, \sigma_{F_i}^2, Q_{ij}, Q_{j0} \\
\gamma_i & \mid \Lambda, Z_i, V_{\gamma} 
\end{align*}
$$

In order to complete the model, the priors for the parameters are specified as follows:

$$\Sigma = \begin{pmatrix}
\sigma_i^2 & \ldots & 0 \\
\ldots & \sigma_i^2 & \ldots \\
0 & \ldots & 1
\end{pmatrix}, \quad \sigma_i^2 \sim IG \left( s_{ij}, s_{ij} \right)$$

$$\lambda = vec \left( \Lambda \right) \sim N \left( \bar{\Lambda}, V_{\lambda} \right)$$

$$V_{\gamma} \sim \text{Wishart} \left( g, G \right)$$

$$Q_{j0} \sim N \left( \bar{Q}_0, \theta_0^2 \right)$$
In the current version of the model, the individual-level parameters are not related to the physician demographics, i.e. $Z_i$ is a scalar with the value 1. Thus, $M=1$. However, we shall conduct some post-estimation analysis to find relationships between the individual-level parameters and physician demographics.

### 6. Estimation

As described earlier, the purpose of this study is to be able to estimate individual physician-level parameters of the model. One approach that may be considered is to specify a model of learning and then estimate such a model for each individual physician, given the data for that physician. However, this approach is not feasible since the data required for such estimation is not available for each of the physicians in the dataset. Hence, a Hierarchical Bayes approach is a natural alternative, which utilizes shrinkage to get usable estimates for each of the physicians in the data, irrespective of the amount of data available for these physicians.

The Bayesian method of inference for the parameters of the model, including the individual-physician level parameters, involves obtaining draws from the joint posterior distribution of these parameters. However, the joint posterior distribution does not correspond to any known distribution family. Hence, we use the Gibbs Sampler to obtain draws from this joint posterior distribution of the parameters by sequentially drawing from the full conditional distributions of sub-vectors of the full parameter vector. These sub-vectors are chosen such that it is easy to draw from their respective full conditional distributions. We iterate this process of making draws from the respective full-conditional distributions until we attain convergence to the true joint posterior distribution of the parameters (cf. Gelfand and Smith 1990). Some of these full conditional distributions are known distributions and hence drawing from them is trivial. However, the full conditional distributions for some of the parameters are in turn not from any of the known distribution families. Hence, we use the Metropolis Hastings Algorithm (Chib and Greenberg 1995) to make draws from these full conditional distributions. Additionally, using the Data Augmentation approach (Tanner and Wong
(1989), we treat the unknown utilities $U_{ijt}$ and mean efficacies $Q_{ij}$ as parameters and make draws for them from their own full conditional distributions. The full details of the likelihood, the full conditional distributions and the details on the algorithm are given in Appendix A.

7. Identification

The identification of the parameters of the model is discussed in this section. In particular, it is important to point out how we are able to identify individual-specific parameters of learning and also how we are able to separate the informative and persuasive effects of detailing.

The individual level parameters in the model are – the variances of the detailing and feedback signals ($\sigma_{D_i}^2$ and $\sigma_{F_i}^2$ respectively), the true mean quality of each drug $Q_{ij}$ and the linear coefficients $\beta_i$. The variances of the detailing and feedback signals are identified from the evolution patterns of physician prescription behavior and how they are related with detailing and feedback signals. We have seen earlier that with every signal, the physician’s quality belief is updated. As a result, the mean of this quality belief evolves from the mean of the initial quality belief towards the mean of the final quality belief. Equation (9) summarizes the process of updation of the mean of the quality belief. As the quality belief for a drug evolves, the probability of prescribing that drug also evolves over time. We have already discussed, when discussing the model specification, how a greater variance of the detailing signal implies a slower learning rate. Thus, the rate of learning helps identify the variance parameter for the detailing signal. A similar argument holds in the case of the variance of the feedback signal.

It must also be noted that we need multiple observations of new drugs in order to infer that the learning rates are systematic to the physician. It would otherwise be hard to separate from random events like specific realizations of the random error $e_{ijt}$, which would give us similar prescription patterns. However, if we observe the patterns of
evolution of prescription behavior for the same physician for multiple drugs, we would be able to make inferences about the learning rate and thus about the parameters that summarize this learning rate ($\sigma_{D_i}^2$ and $\sigma_{F_i}^2$).

The true mean quality $Q_{ij}$ of the drug is identified out of the steady state prescription behavior of the physician. We have already seen that as the physician learns, the mean quality belief $Q_{ijt}$ evolves to the true mean quality of the drug $Q_{ij}$. At the extreme, at steady state, the mean quality belief is indeed the true mean quality. Thus, we need to observe a long enough time series of prescriptions for the physicians to be able to correctly identify the true mean quality. In the data we use for our empirical analysis, the data is observed for about 9 months after launch. Hence, the prescription behavior of the physician towards the end of the dataset tells us about the true mean quality of the drug.

The linear coefficients $\beta_i$ are identified from the covariance of the prescription behavior of the physician with the linear variables (detailing stock variable and dummy variable for patient requests). An important concern could be about the separate identification of the informative and persuasive effects of detailing. The identification of the persuasive effect, which is the coefficient of the detailing stock variable in the utility equation, is aided by the fact that in our dataset, the incumbent drug – Viagra – has existed for about 5 years before the first observation in the data. Hence, we make the a priori assumption that physicians have learnt fully about Viagra and hence the informative effect for this drug is zero. Hence, any effect of detailing of Viagra on the prescription behavior of physicians is entirely due to the persuasive effect. For the new entrants – Levitra and Cialis - both the informative and persuasive effects are present. Hence, we are able to identify both these effects.

The initial shares of the drugs before any detailing or feedback was received identifies the initial mean efficacy $Q_{ij0}$. Note that this parameter is pooled across physicians. Potentially, one could estimate this parameter at the individual level, provided we had observations of prescriptions written by the physician after the
introduction of the new drug and before any detailing or feedback signals were received for that drug. However, since we have such observations for only a subset of the physicians, we are constrained to make the assumption that this parameter is common across physicians.

A potential concern could be about endogeneity. Specifically, if firms optimally decide on their detailing levels to individual physicians based on a complete knowledge of their behavior, then the detailing would be endogenous and hence parameter estimates would be biased. There are two reasons why this may not be a big concern for the specific problem we are studying. First, the data that is used in this study was generally not available to firms at the time of collection of the data. This is a new data source from a new data company, which is only now being subscribed to by pharmaceutical firms. Further, even the firms that did have access to the data did not have access to the statistical tools necessary to estimate sophisticated models of physician learning and their dynamic behavior. Second, the typical rule that firms use to decide on their detailing levels to physicians is a volumetric decile-based rule. Physicians are typically classified into deciles based on their total category volumes and detailing allocations are made at the decile levels. Manchanda, Rossi and Chintagunta (2004) have a discussion on the rules used by firms and also show empirically that firms are not optimal in their decision. If concerns on endogeneity were to persist, it might be possible to specify a model in which the detailing decision is also modeled and parameters of the detailing equation and the physician prescription equation are estimated simultaneously, like in Manchanda, Rossi and Chintagunta (2004).

8. Results
The empirical estimation of the model was done using a program written in ‘C’. The chain for the Gibbs Sampler was run for a total of 100,000 observations. The first 50,000 draws were discarded as ‘burn-in’ before convergence was attained. The subsequent draws were used for inference.
We first present the parameter estimates of the model. In Table 1, we report the parameter estimates for the individual-level parameters for the model. The individual-level parameters are the detailing signal variance ($\sigma_{Di}^2$), feedback signal variance ($\sigma_{Fi}^2$), the coefficients for the detailing stock ($\beta_{Di}$) and patient requests ($\beta_{Fi}$); and the true mean qualities for Cialis ($Q_{Di}$) and Levitra ($Q_{Fi}$). The natural logarithms of the detailing and feedback signal variance parameters are reported since these parameters are constrained to be positive and hence they were reparametrized. In our Bayesian inference, we obtain a distribution for each individual-level parameter for each physician. Since it is neither feasible nor interesting to report these distributions for each individual physician, we compute the mean parameter value for each physician and then report the mean and standard deviation across physicians of this individual-level mean parameter value in Table 1.

The estimate for logarithm of the detailing signal variance is found to be -1.2537 on average. This parameter value implies that it takes about 2.5690 detailing calls for the uncertainty of a physician to reduce to one-tenth of its initial value. This is an estimate of the informative effect of detailing. Similarly, the parameter estimate of the logarithm of the feedback signal variance of -0.8932 corresponds to 3.6841 feedback signals to reduce the uncertainty to 10% of its initial value. This also suggests that an average detailing call is more informative than an average feedback signal since it requires a smaller number of detailing calls than feedback signals to reduce the physician’s uncertainty about drug quality. This is consistent with the findings of prior research (cf. Narayanan, Manchanda and Chintagunta 2003). The standard deviations for these estimates suggests that there is considerable heterogeneity across physicians in these parameters. We shall return to a discussion on heterogeneity in these parameters later in this section.

---

3 This is computed by solving for $nd$ in the equation $\frac{1}{0.1} = 1 + \frac{nd}{\sigma_{Di}^2}$, which is itself derived by recursive application of the variance updation expression in equation (10) and setting the initial variance to be 1.
The parameter estimates for the coefficient of the linear detailing stock \( \beta_{ib} \) suggest that there is a positive persuasive effect of detailing. Thus, even after a physician has reduced his uncertainty about the drug, there is still a positive effect of detailing. There is also a strong positive effect of patient feedback, as indicated by the parameter estimates for the coefficient for patient requests \( \beta_{2i} \). This is an interesting finding in itself, since absence of data on patient requests for specific drugs has made it impossible to document an effect of patient requests on the physician’s prescription behavior, even though it has been speculated that this would have a positive effect.

The parameter estimates for the true mean qualities for Cialis and Levitra are both greater than zero. This suggests that physicians have a higher quality perception for both these drugs than Viagra on average (since Viagra is indexed at zero). Thus, after they have learnt about the drug, i.e. in steady state, the physician should have a higher probability of prescribing these two drugs than Viagra, everything else (detailing stock and patient request) remaining constant. Specifically, the highest mean quality belief is for Cialis, followed by Levitra and Viagra in that order.

In Table 2, we report the parameter estimates for the pooled parameters. The pooled parameters that we report include the initial quality means for Cialis \( Q_{1,0} \) and Levitra \( Q_{2,0} \); and the variances of the normal errors in the utility function \( \sigma^2_i \) and \( \sigma^2_2 \). The additional pooled parameters include the parameters that relate the individual-level parameters to physician demographics, i.e. \( \lambda \) and \( V nightly. However, we do not report these parameters since in our current estimation, we have not included physician demographics in the model. Thus, the vector of means of the individual-level parameters is the same as \( \lambda \) and the variance-covariance matrix is the same as \( V nightly.

The parameter estimates for the initial quality mean suggests that physicians started out with an initial belief that the mean quality of these new drugs was lower than that of Viagra. Further, their initial belief in the case of Cialis was more pessimistic than in the case of Levitra. The value of the initial quality belief affects the probability of
prescribing the new drug before the physician has been exposed to any detailing signals or feedback signals. Thus, the parameter estimates suggest that the initial probability of prescribing Cialis is lower than that for Levitra.

Next, we move to the heterogeneity in the individual-level parameter estimates. Figure 1 depicts the histograms of the individual-level parameters across physicians. Figures 1(a) and 1(b) respectively show the histograms of the logarithms of the detailing and feedback signal variances. These figures suggest that there is considerable heterogeneity in these parameters across physicians. There is much greater heterogeneity in the detailing signal variance than the feedback signal variance.

We find from observing Figure 1(c) that for an overwhelming majority of physicians, the detailing stock coefficient is positive. Thus, for most physicians, the persuasive effect of detailing is positive. There is a very small segment of physicians for whom the persuasive effect is negative. However, further analysis reveals that the 95% credible interval for these physicians (not shown in the Figure) includes the value zero, thus suggesting that it may be appropriate to characterize them as having no persuasive effect. However, for a majority of physicians, the persuasive effect of detailing is positive and the 95% credible interval (not shown) does not include zero. Figure 1(d) shows the distribution across physicians of the patient request coefficient. It is positive for all physicians and the 95% credible interval does not include the value zero for any physician.

In Figures 1(e) and 1(f), we show the distributions across physicians of the true mean qualities of Cialis and Levitra respectively. Once again, both these new drugs are perceived to have a higher true mean quality than the incumbent, Viagra, by a majority of physicians. It must be said however, that the difference between the qualities of these drugs is not very high.

When talking about heterogeneity, a valid concern could be whether the individual-level parameters are significantly different from each other. Hence, in Table
3, we compare the across-physician and within-physician standard deviations of these parameters. If the across-physician standard deviation for a parameter is smaller than or similar to the within-physician standard deviation, it would suggest that the 95% credible intervals for the physicians overlap and hence they are not significantly different from each other in terms of that parameter. On the other hand, if the across-physician standard deviation is larger than the within physician standard deviation, it would provide greater support to the claim that the individual-level parameters are significantly different for different physicians.

We find that the across-physician variation is much higher than the within-physician variance in the case of the two signal variances. Since the heterogeneity in the detailing signal variance summarizes the heterogeneity in learning as discussed in the model specification section, we can conclude that there is significant amount of heterogeneity in learning across physicians.

The coefficient of the detailing stock also has a higher across-physician standard deviation than the within-physician standard deviation. Thus, there is a significant amount of heterogeneity in this parameter as well. In the case of the other individual-level parameters – the coefficient for patient requests and the true mean qualities for Cialis and Levitra – the within-physician standard deviation is larger than the across-physician standard deviation, suggesting that physicians do not significantly differ in these parameters.

We might also be concerned whether we are able to estimate tight distributions of these individual-level parameters for all physicians. In particular, we have some physicians with a lot of observations and others with few observations. The question is whether the parameters are estimated tightly for physicians for whom we have few observations versus those for whom we have a large number of observations. Figure 2 depicts the relationship between the within-physician standard deviation of the individual-level parameters and the number of observations for the physician. We can see from figures 2(a) through 2(f) that the standard deviation is negatively correlated with
the number of observations. Thus, physicians for whom we have a larger number of observations are also likely to be those for whom the parameters are estimated more tightly. For physicians for whom we have few observations, the individual-level parameter distributions look similar to the population distributions.

In order to explore the heterogeneity in learning across physicians even further, we plot a histogram of the number of detailing calls required to reduce the physician’s uncertainty about a new drug to one-tenth its initial value. This is computed using the parameter estimates of the detailing signal variance for each physician\textsuperscript{4}. Figure 3, which shows this plot, suggests that heterogeneity in the detailing signal variance parameter indeed manifests itself in significant heterogeneity in the number of calls required to reduce the physicians’ uncertainty.

Another way to look at heterogeneity in learning would be to plot the evolution patterns for the means of the quality beliefs for physicians at two ends of the heterogeneity distribution. In Figure 4(a), we plot the evolution of the mean of the quality belief for Cialis for a physician who is in the 90th percentile in terms of learning rate, i.e. whose detailing variance is lower than 90% of the physicians. In Figure 4(b), we have a similar plot for a physician in the 10th percentile, i.e. whose detailing variance is lower than only 10% of the physicians. The plots show that while the physician who is a fast learner (90th percentile) takes just one or two detailing calls for the mean quality to evolve from the initial value to the true mean quality, the slow learner takes a large number of repeated detailing calls in order to converge to the true mean quality.

An interesting fact about the parameter estimates is the negative correlation between the informative and persuasive effects of detailing for physicians. What this means is that physicians who have a high informative effect of detailing are likely to have a low persuasive effect and vice versa. The informative effect of detailing is summarized by the detailing signal variance. We have already seen that a high informative effect is

\footnote{See footnote 3 for how to compute the number of detailing calls required to reduce the uncertainty to one-tenth its initial value.}
equivalent to saying that the learning rate for the physician is high. And this manifests itself in the parameter estimates in terms of a low value of the detailing signal variance. Similarly, a low informative effect manifests itself in a high detailing variance. The persuasive effect is summarized by the coefficient of the detailing stock variable. Thus, a physician for whom the value of this coefficient is low would have a low persuasive effect and vice versa. Therefore, a negative correlation between the informative and persuasive effects of detailing implies a positive correlation between the detailing signal variance and the detailing stock coefficient. In Figure 5, we plot the detailing signal variance for each physician against the detailing stock coefficient to show this positive correlation between the parameters and consequently the negative correlation between the informative and persuasive effects.

In order to understand what explains the heterogeneity in the informative and persuasive effects across physicians, we regress the individual level parameters on physician demographic characteristics. We have two sets of physician demographic characteristics, one of which is the specialty of the physician (whether a General Physician or a Urologist or some other specialty) and the volume-based decile of the physician. The volume-based decile is obtained by computing the total category volume of the physician in the 3-month period immediately before the first month in our dataset and then categorizing all physicians into deciles based on this volume. Thus, we have sets of dummy variables for specialty and decile. We conduct this regression for two sets of individual-level parameters – the detailing signal variance ($\sigma^2_{Di}$) and the coefficient of the linear detailing stock variable ($\beta_{Di}$), which respectively summarize the informative and persuasive effects of detailing. The regression estimates are reported in Table 4.

We find that the variation in both the individual level parameters is explained to some extent by the decile of the physician. The detailing signal variance is approximately constant for the physicians in the 1st through the 7th decile. However, from the 7th decile to the 10th decile, this variance declines monotonically with decile (with decile 10 being at 0). Since a lower detailing signal variance indicates a faster learning rate, these regression estimates suggest that physicians in higher deciles are
faster learner than those in lower deciles in the decile 7 to 10 range. Physicians in deciles 1 through 7 are approximately equal in their learning rates. One of the two specialty variables – the dummy variable for urologists - is significant. The parameter estimates suggest that urologists are faster learners than all other physicians. These estimates make intuitive sense. Physicians who are in high deciles, i.e. who are heavy prescribers in the category, are likely to have greater ability to process information from the detailing call, ask the right questions to detailers and show greater interest in the information that detailers provide than those in the lower deciles. Similarly, urologists, who are the relevant specialists for this category, are likely to be faster learners.

Table 4 also has estimates of the regression for the coefficient for the detailing stock, which summarizes the persuasive effect of detailing. In this case, the coefficients for specialty dummy variables are not significant. We find that decile does explain the variation in the detailing stock coefficient. From decile 1 through 4, there is no monotonic pattern for this parameter, but from decile 4 through 10, it reduces monotonically. Thus, physicians in higher deciles are likely to have the least magnitude of the persuasive effect than those in lower deciles. While there is no intuitive reason to expect this effect to go one way or another, one potential explanation for this monotonically decreasing persuasive effect is the following. Some underlying phenomena by which the persuasive effect is thought to influence physician prescriptions are reminder effects and top-of-mind effects. If this is the case, then physicians who are heavy prescribers are likely to be less require reminders in the first place. This could be a potential explanation for why we see the monotonically declining relationship between the persuasive effect and decile.

A caveat in this analysis is that the R-squared in both the regressions is not very high. It is higher in the case of the detailing signal variance than in the case of the detailing stock coefficient, where it is actually quite low. This suggests that there are other unobservables that drive the heterogeneity in learning (i.e. the informative effects) and persuasive effects of detailing.
9. Managerial Implications

The negative correlation between the informative and persuasive effects has implications for the allocation of detailing efforts across physicians and over time. To illustrate this, let us consider a situation with two physicians, one with a high informative effect and a low persuasive effect, and the other with the opposite. Let us, for the sake of convenient refer to the former as the fast learner and the latter as the slow learner. For both these physicians, the informative and persuasive effects are present in the introductory phases of the new drug’s life cycle. In later stages, only the persuasive effect plays a role. For the fast learner, the total effect of detailing starts off at a very high level, but rapidly reduces till it converges to the persuasive effect and then remains constant at that level. For the slow learner, everything remaining the same, the total effect again starts at a high level (perhaps not as high as for the fast learner) but falls more slowly. It converges to a persuasive effect that is higher than for the fast learner. Since this total effect denotes the responsiveness of the physician to detailing and the optimal detailing level depends on responsiveness, it would be optimal for firms to allocate higher amounts of detailing initially to the fast learner but then rapidly reduce this allocation. On the other hand, for the slow learner, we might expect to see a corresponding increase in allocation over time.

We have also seen that the informative and persuasive effects are related to the decile of the physician. Specifically, physicians in higher deciles have a greater informative effect and a lower persuasive effect and those in lower deciles have a lower informative effect but a higher persuasive effect. Thus, if firms were to optimally allocate their resources to physicians, one would expect to see them allocating a high proportion of resources to physicians in high deciles in the early stages after the introduction of the drug and then reducing this proportion over time. However, discussions with practitioners in the industry reveal that firms do not allocate their detailing efforts in this manner. The most commonly used rule for allocating resources to physicians is a decile-based rule, which remains constant over time. Figure 6 shows the allocation of detailing by decile for Levitra for the first 7 months after its launch. We find that detailing allocations remain largely constant for each decile.
We conduct three counterfactual experiments to see how firms could increase their revenues if they took into account learning by physicians and heterogeneity in this learning. In each of these experiments, we conduct separate experiments for each of the two new drugs – Levitra and Cialis – one by one. We keep the total amount of detailing by firms in the first three months after the launch of the respective drug as fixed. We only alter the allocation of detailing across physicians or over time and then compute the predicted revenues. In all these experiments, detailing calls are rounded off, i.e. a physician can only get an integer number of calls. All these calls are assumed to be made in the beginning of the month, so as to abstract away from the problem of timing of detailing calls within a month. Competitors’ detailing is kept unchanged in all these experiments. The ‘optimal’ allocations are obtained using a numerical optimization routine. We conduct all these experiments for a subset of 100 physicians randomly chosen from the sample. This is done in order to keep the optimization feasible. We compare the predicted revenues in the counterfactual case with those using the actual allocation plans.

In Experiment 1, we change only the temporal allocation of detailing, but keep the cross-sectional allocation unchanged, i.e. we vary the allocation of detailing to each of the three months in the launch quarter, but do not vary the proportions of detailing within that month to individual physicians. Thus, this is a two-dimensional optimization exercise, with the unknown variables being the allocation for the first two months (and the allocation of the third month automatically known from these since the proportions add up to 1). We find that by varying the month to month allocations of detailing, Levitra can get revenue gains of 5.4% compared to the current allocation plan, while Cialis could obtain an even higher gain of 8.3%. This reflects the fact that firms could gain by frontloading their detailing to the period immediately after launch. This is because in this early period, all physicians have both an informative effect and a persuasive effect and are hence most responsive to detailing. As they learn about the drug, the informative effect asymptotes down to zero and hence only the persuasive effect is present. Thus, for all physicians, the responsiveness to detailing reduces over time. This result is similar to that reported in Narayanan, Manchanda and Chintagunta (2003).
In Experiment 2, we change the cross-sectional allocation of detailing, but keep the temporal allocation unchanged. Thus, we keep the total amount of detailing within each month fixed, but vary how much of that is allocated to each of the physicians. In each month, we have an independent optimization, with the dimension of the unknown vector being one less than the number of physicians (i.e. 99 since the experiment was conducted for 100 physicians). We find that there are relatively modest revenue increases that can result through this exercise, 3.8% for Levitra and 5.2% for Cialis. The reason for revenue increases being lower in this case is perhaps that the firms are using some decile-based rule for allocation and while this is not optimal, it does inherently take into account heterogeneity across physicians (since the informative and persuasive effects of detailing are explained to some extent by decile).

In Experiment 3, we allow both the temporal and cross-sectional allocation of detailing to change, i.e. we find the detailing level for each physician for each month that maximizes revenues. We find that this gives us a revenue increase of 10.1% for Levitra and 14.7% for Cialis. Note that the revenue gains for Experiments 1 and 2 do not add up exactly to the revenue gains in Experiment 3. This is due to the rounding off of detailing to individual physicians.

The results of these three counterfactual experiments are summarized in Table 5. To sum up, it appears that firms could significantly increase their revenues during the launch period by simply reallocating their existing expenditure on detailing. An interesting extension could be study what would be the optimal levels of detailing and not just the optimal allocations of detailing. Note that in all the three experiments we have conducted, the total amount of detailing in the first three months after launch was kept fixed. However, firms may be interested in knowing if they could reduce their total detailing expenditure during launch. This could also be addressed by future research.
10. Conclusion

We started off with a problem of optimal allocation of marketing communication for new products, with the specific focus on allocation of these resources over time and across consumers. Heterogeneous learning was recognized as a factor that would affect the temporal as well as cross-sectional allocation of these resources. We specified a structural model of heterogeneous learning and developed a methodology to estimate such a model at the individual physician level. We estimated this model using a unique panel dataset consisting of physician prescriptions and detailing calls. We allowed for detailing to have both an informative and a persuasive effect and estimated both these effects at the individual physician level. We then conducted a set of counterfactual experiments to find the implications of heterogeneous learning on optimal allocations of detailing over time and across agents (physicians in our case).

Our parameter estimates indicate that there is considerable heterogeneity across physicians in terms of their learning rates. Some physicians require only one detailing call to substantially reduce their uncertainty about a new drug, others require a large number of repeated detailing calls in order to reduce their uncertainty to the same extent. Physicians also differ significantly in their persuasive effect of detailing, which is the only effect of detailing after they have learnt about the drug. Because of both these effects, there is a significant amount of variation across physicians in terms of how their responsiveness to detailing varies over time.

We also find that volume based deciles explain the variation in both the informative and persuasive effects of detailing to some extent. Specifically, physicians who are heavy prescribers in the category are likely to be fast learners and thus, have a high informative effect in the very early stages after the launch of the drug. They also have a low persuasive effect. Thus, their responsiveness to detailing is high initially but rapidly drops to a relatively low level. Lighter prescribers have a low informative effect initially by comparison, but their persuasive effect is higher. Thus, they are likely to have a more modest responsiveness to detailing in early stages but this reduces more slowly and settles down at a relatively high responsiveness after they have learnt about the drug.
We conducted three counterfactual experiments to see if firms could increase their revenues in this category by changing their detailing allocation patterns. We find that if they change just their temporal allocation without changing the cross-sectional allocation of detailing, they get a 5.4% to 8.3% increase in revenues in the first three months after launch. This reflects gains in revenues by front-loading their detailing to early periods after the launch of the drug. If they change their cross-sectional allocation without altering the temporal allocation, firms obtain a more modest 3.8% to 5.2% increase in revenues. If they change both their temporal and cross-sectional allocations, they get a substantial 10.1% to 14.7% increase in revenues.

While our empirical analysis is conducted for a category of prescription drugs, the results are generalizable. They would be applicable to any situation where consumers learn about new products through some means of marketing communication. For instance, in the case of industrial goods, firms often send out salespersons to their customers not just to sell their products but also to inform them about new products or services. Firms need to decide how to allocate their salesforce efforts across consumers and therefore need to take into account any heterogeneity in learning rates that may be present across customer accounts. Even in the case of consumer goods, firms may need to allocate their resources for informative advertising across markets or segments of consumers even if they do not necessarily target individual consumers. Accounting for differences in learning rates across different segments of consumers or different markets may be important in arriving at optimal allocation plans for advertising.

We finally list some of the limitations of this study. In specifying this model of learning about new drugs, we have assumed away other potentially important sources of learning, for instance learning from other physicians. This assumption of no learning through other sources is due to the absence of appropriate data and might overstate the degree of learning through detailing that we infer. Another potentially important concern could be about endogeneity. Firms could be at least partially optimizing their detailing allocations already and the presence of the resulting endogeneity may bias our estimates.
While there are reasons to believe that this may not be a very big concern for the specific category we study, this could be addressed by including a detailing supply equation in the model and jointly estimating parameters of demand and supply. A complication that would arise in this case is that the detailing supply equation cannot be static due to the presence of learning. Learning causes detailing to have persistent effects over time and hence firms are likely to take into account the effect of their detailing on future prescription behavior of physicians. This would complicate the problem substantially. We also assume away any forward-looking behavior of physicians. Physicians, if they are aware that they learn through detailing, may have incentives to be more willing to see detailers early on in order to learn about the drug quickly. The presence of such behavior is also likely to bias our inferences about learning. These are challenging problems and the methodology to account for these phenomena are not fully developed yet. Future research could potentially address these questions.
References


Table 1: Individual Level Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(detailing signal variance) $\log\left(\sigma_D^2\right)$</td>
<td>-1.2537</td>
<td>1.4545</td>
</tr>
<tr>
<td>log(feedback signal variance) $\log\left(\sigma_F^2\right)$</td>
<td>-0.8932</td>
<td>0.9483</td>
</tr>
<tr>
<td>Coefficient – detailing stock $\beta_{ui}$</td>
<td>0.0529</td>
<td>0.0256</td>
</tr>
<tr>
<td>Coefficient – patient request $\beta_{2i}$</td>
<td>0.5292</td>
<td>0.0170</td>
</tr>
<tr>
<td>True Mean Quality – Cialis $Q_{i1}$</td>
<td>0.0160</td>
<td>0.0068</td>
</tr>
<tr>
<td>True Mean Quality – Levitra $Q_{i2}$</td>
<td>0.0076</td>
<td>0.0032</td>
</tr>
</tbody>
</table>

Note: Since these parameters are at the individual level, for each individual physician, the parameter has a mean and a standard deviation. The reported parameters are the mean and standard deviation of the parameter mean for each of the physicians.

Table 2: Pooled Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Mean Quality – Cialis $Q_{1,0}$</td>
<td>-0.1106</td>
<td>0.0167</td>
</tr>
<tr>
<td>Initial Mean Quality – Levitra $Q_{2,0}$</td>
<td>-0.0543</td>
<td>0.0146</td>
</tr>
<tr>
<td>Utility Error Variance $\sigma_1^2$</td>
<td>0.3267</td>
<td>0.0233</td>
</tr>
<tr>
<td>Utility Error Variance $\sigma_2^2$</td>
<td>0.3343</td>
<td>0.0204</td>
</tr>
</tbody>
</table>

Note: Other pooled parameters include the elements of the matrix $V_y$ and the vector $\lambda$. These are not reported here for the sake of brevity.
Table 3: Across-physician vs. within-physician standard deviations of the individual-level parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Across-physician standard deviation</th>
<th>Within-physician standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(detailing signal variance)</td>
<td>( \log(\sigma_{D_i}^2) )</td>
<td>1.4545</td>
</tr>
<tr>
<td>log(feedback signal variance)</td>
<td>( \log(\sigma_{F_i}^2) )</td>
<td>0.9483</td>
</tr>
<tr>
<td>Coefficient – detailing stock ( \beta_{D_i} )</td>
<td>0.0256</td>
<td>0.0123</td>
</tr>
<tr>
<td>Coefficient – patient request ( \beta_{F_i} )</td>
<td>0.0170</td>
<td>0.0210</td>
</tr>
<tr>
<td>True Mean Quality – Cialis ( Q_{i1} )</td>
<td>0.0068</td>
<td>0.0077</td>
</tr>
</tbody>
</table>
| True Mean Quality – Levitra \( Q_{i2} \)      | 0.0032                             | 0.0068                             | **Notes:**

1. The across-physician standard deviation: the mean parameter value for each physician is first computed and then the standard deviation of these mean values reported in this table.

2. The within-physician standard deviation: the within-physician standard deviation of the parameter is computed for each physician and then the mean of these standard deviations is reported in this table.

Table 4: Regression of Individual-level Parameters on Demographic Variables

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Detailing Signal Variance ( \sigma_{D_i}^2 )</th>
<th>Detailing Stock Coefficient ( \beta_{D_i} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.1330**</td>
<td>0.2262</td>
</tr>
<tr>
<td>Specialty – GP</td>
<td>-0.0980</td>
<td>0.1829</td>
</tr>
<tr>
<td>Specialty - Urologist</td>
<td>-1.3947**</td>
<td>0.2582</td>
</tr>
<tr>
<td>Decile 1</td>
<td>1.3706**</td>
<td>0.2127</td>
</tr>
<tr>
<td>Decile 2</td>
<td>1.4626**</td>
<td>0.2047</td>
</tr>
<tr>
<td>Decile 3</td>
<td>1.2308**</td>
<td>0.2127</td>
</tr>
<tr>
<td>Decile 4</td>
<td>1.4449**</td>
<td>0.1809</td>
</tr>
<tr>
<td>Decile 5</td>
<td>1.1763**</td>
<td>0.2156</td>
</tr>
<tr>
<td>Decile 6</td>
<td>1.3028**</td>
<td>0.1936</td>
</tr>
<tr>
<td>Decile 7</td>
<td>1.4239**</td>
<td>0.1800</td>
</tr>
<tr>
<td>Decile 8</td>
<td>1.0725**</td>
<td>0.2055</td>
</tr>
<tr>
<td>Decile 9</td>
<td>0.5275**</td>
<td>0.1770</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.2394</td>
<td>0.0718</td>
</tr>
</tbody>
</table>

** Significant at the 95% level
Table 5: Counterfactual Experiments
Revenue Gains through Reallocation of Detailing Compared to Current Situation

<table>
<thead>
<tr>
<th>Temporal Allocation</th>
<th>No Change</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Situation</td>
<td>No Change</td>
<td>Experiment 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cialis: 8.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Levitra: 5.4%</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>No Change</td>
<td>Experiment 3</td>
</tr>
<tr>
<td>Cialis: 5.2%</td>
<td></td>
<td>Cialis: 14.7%</td>
</tr>
<tr>
<td>Levitra: 3.8%</td>
<td></td>
<td>Levitra: 10.1%</td>
</tr>
<tr>
<td>Change</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Histograms of Means of Individual Level Parameters

1(a): detailing signal variance \( \log(\sigma^2_{D_i}) \)  
1(b): feedback signal variance \( \log(\sigma^2_{F_i}) \)

1(c): detailing stock coefficient \( \beta_{1i} \)  
1(d): patient request coefficient \( \beta_{2i} \)

1(e): true mean quality for Cialis \( Q_{i1} \)  
1(f): true mean quality for Levitra \( Q_{i2} \)
Figure 2: Standard Deviations of Individual Level Parameters
Relationship with Number of Observations for the Physician

2(a): detailing signal variance $\log (\sigma^2_{D_i})$

2(b): feedback signal variance $\log (\sigma^2_{F_i})$

2(c): detailing stock coefficient $\beta_{1i}$

2(d): patient request coefficient $\beta_{2i}$

2(e): true mean quality for Cialis $Q_{1i}$

2(f): true mean quality for Levitra $Q_{12}$
Figure 3: Histogram of number of calls required to reduce uncertainty (variance of quality belief) about a new drug to one-tenth of the initial value

Figure 4: Evolution of Mean of the Quality Belief for Cialis

4a: Physician in the 90\textsuperscript{th} Percentile

4b: Physician in the 10\textsuperscript{th} percentile
Figure 5: Plot of Informative Effect vs. Persuasive Effect
Detailing Stock Coefficient vs. log(Detailing Signal Variance)

Figure 6: Detailing Allocation by Decile – Levitra

Note: Each line corresponds to a decile
Appendix A : Full Conditional Distributions

Given the model described in the model section, the joint posterior distribution of all the parameters conditional on the data is given by is given by the following expression

\[
L \propto \prod_{i=1}^{N} \prod_{j=1}^{J} \left[ f \left( U_{ij} \mid \{ Y_{ij} \}, \{ Q_{ij} \}, \beta_i, \Sigma \right) \prod_{j=1}^{J} f \left( Q_{ij} \mid Q_{ij(t-1)}, nd_{ij}, I_{ij(t-1)}, \sigma_{D_i}^2, \sigma_{F_i}^2, Q_j, Q_{ij0} \right) \right] \\
\cdot f \left( \gamma_i \mid \Lambda, Z_i, V_r \right) \\
\cdot f \left( \lambda \mid \Lambda, V_k \right) f \left( V_r \mid G \right) \prod_{j=2}^{J} f \left( \sigma_{j}^2 \mid s_{1j}, s_{2j} \right)
\]

We shall now derive the full conditional distributions of all the parameters (and augmented parameters like \( U_{ij} \) and \( Q_{ij} \)), so that we can use a Gibbs sampling method.

For those parameters for which the full conditional distributions are not from known distribution families, we shall use the Metropolis Hastings algorithm to draw from the respective full conditional distributions.

The full expression for the joint posterior distribution can be written as follows

\[
L \propto \prod_{i=1}^{N} \prod_{j=1}^{J} \frac{1}{\sigma_{\gamma_i}^2 \sigma_{\lambda_j}^2} \exp \left( \frac{-1}{2} \left( \gamma_i - \Lambda Z_i \right) \right) \cdot \frac{1}{\sigma_{\gamma_i}^2} \exp \left( \frac{-1}{2} \left( Q_{ij} - \sigma_{\gamma_i}^2 Q_{ij(t-1)} - \sigma_{\gamma_i}^2 Q_j - \sigma_{\gamma_i}^2 \right) \right) \\
\cdot \frac{1}{\sigma_{\lambda_j}^2} \exp \left( \frac{-1}{2} \left( \lambda_j - \Lambda Z_j \right) \right) \\
\cdot \frac{1}{\sigma_{\gamma_i}^2} \exp \left( \frac{-1}{2} \left( \gamma_i - \Lambda Z_i \right) \right) \cdot \frac{1}{\sigma_{\lambda_j}^2} \exp \left( \frac{-1}{2} \left( \lambda_j - \Lambda Z_j \right) \right)
\]
where

\[ U_{it} = \begin{pmatrix} U_{i1t} \\ \vdots \\ U_{ih} \end{pmatrix}, \quad Q_{it} = \begin{pmatrix} Q_{i1t} \\ \vdots \\ Q_{ijt} \end{pmatrix}, \quad X_{it} = \begin{pmatrix} X_{i1t} \\ \vdots \\ X_{ijt} \end{pmatrix} \]

From this joint posterior, we can derive the full conditional distributions as follows:

1. \( Q_{ijt} \mid Q_{q(i-t)}, Q_{q(i+t)}, U_{ijt}, \sigma^2_i, I_{ij(t-1)}, I_{ij(t)}, nd_{ijt}, nd_{ij(t+1)}, \sigma^2_F, \sigma^2_D \sim N(m, v) \)

where

\[
\frac{1}{v} = \frac{1}{\sigma^2_j} + \frac{1}{\sigma^2_{Q_{ijt}}} \left( \frac{nf_{ijt} + nd_{ijt}}{\sigma^2_F} - \frac{1}{\sigma^2_{D_{ijt}}} \right) + \frac{1}{\sigma^2_{Q_{q(i+t)}}} \left( \frac{nf_{ij(t+1)} + nd_{ij(t+1)}}{\sigma^2_F} - \frac{1}{\sigma^2_{D_{ij(t+1)}}} \right)
\]

\[
m = \frac{1}{v} \left[ U_{ijt} \frac{\sigma^2_{Q_{ijt}}}{\sigma^2_{Q_{q(i-t)}}} Q_{ij(t-1)} + Q_{ij(t)} \left( \frac{nf_{ijt} + nd_{ijt}}{\sigma^2_F} - \frac{1}{\sigma^2_{D_{ijt}}} \right) + Q_{ij(t+1)} - Q_{ij} \sigma^2_{Q_{q(i+t)}} \left( \frac{nf_{ij(t+1)} + nd_{ij(t+1)}}{\sigma^2_F} - \frac{1}{\sigma^2_{D_{ij(t+1)}}} \right) \right]
\]

2. \( \sigma^2_j \mid \{U_{ijt}\}, \{X_{ijt}\}, \{Q_{ijt}\}, \{\beta_i\} \sim IG \left( \frac{\sum_{i=1}^N T_i}{2} + s_{1j} \right) \left( \frac{2s_{2j}}{2 + s_{2j} \sum_{i=1}^N (U_{ijt} - Q_{ijt} - X_{ijt} \beta_i)^2} \right)
\]

3. \( V \mid \gamma_i, \Lambda, Z_i, g, G \sim \text{Inverse Wishart} \left( g + N, \sum_{i=1}^N \left( (\gamma_i - \Lambda Z_i) (\gamma_i - \Lambda Z_i)^T + G^{-1} \right) \right) \]

4. \( \lambda \sim N \left( \left[ V^{-1} \otimes Z' Z + V \lambda^{-1} \right]^{-1} \left[ V^{-1} \otimes Z' Z \lambda^{-1} \lambda + V \lambda^{-1} \lambda \right], \left[ V^{-1} \otimes Z' Z + V \lambda^{-1} \right]^{-1} \right) \)

where

\[ \lambda = \text{vec} \left( (Z' Z)^{-1} Z \Gamma \right) \]
\[ Z = \begin{pmatrix} Z_1' \\ . \\ Z_N' \end{pmatrix}, \quad \Gamma = \begin{pmatrix} \gamma_1' \\ . \\ \gamma_N' \end{pmatrix} \]

5. Each physician is independent conditional on the \( X \) and \( Z \) matrices. And each observation for the physician is independent conditional on the vector of \( Q_{ijt} \). Thus, we can draw the latent utilities for a particular physician and a particular observation separately from the other observations. This involves sequentially drawing from a truncated multivariate normal distribution for each time period.

\[
\begin{bmatrix} U_{i1t} \\ . \\ U_{ikt} \end{bmatrix} \sim \text{Truncated MVN} \left( \begin{pmatrix} Q_{i1t} + X_{i1T} \beta_1 \\ . \\ Q_{ikt} + X_{i1T} \beta_1 \end{pmatrix}, \Sigma \right)
\]

with the truncation such that \( U_{ijt} > U_{ikt}, \forall k \neq j, I_{ijt} = 1\)

6. The full conditional distributions for the \( \gamma_i \) and \( Q_{j0} \) parameters are not from known families of distributions. Hence, draws from the distribution of these parameters for each individual physician are obtained using the Metropolis Hastings algorithm. We use a Random Walk Metropolis Hastings algorithm (Chib and Greenberg, 1995) to make these draws.