MANAGEMENT SCIENCE
Vol. 54, No. 2, February 2008, pp. 266–278
ISSN 0025-1909 | EISSN 1526-5501 | 08 | 5402 | 0266
DOI 10.1287/mnsc.1070.0776

Modeling and Analysis of Call Center Arrival Data: A Bayesian Approach

Refik Soyer, M. Murat Tarimcilar
Department of Decision Sciences, School of Business, The George Washington University, Washington, D.C. 20052
{soyer@gwu.edu, muratt@gwu.edu}

In this paper, we present a modulated Poisson process model to describe and analyze arrival data to a call center. The attractive feature of this model is that it takes into account both covariate and time effects on the call volume intensity, and in so doing, enables us to assess the effectiveness of different advertising strategies along with predicting the arrival patterns. A Bayesian analysis of the model is developed and an extension of the model is presented to describe potential heterogeneity in arrival patterns. The proposed model and the methodology are implemented using real call center arrival data.

Key words: call center; advertising strategy; modulated Poisson process; Bayesian analysis; heterogeneity
History: Accepted by Ger Koole, special issue editor; received October 29, 2004. This paper was with the authors 4 months for 2 revisions. Published online in *Articles in Advance* December 17, 2007.

1. Introduction and Background

Forecasting call volume data that is based on a specific advertisement or promotion plan poses a challenging problem for call centers. The need for optimal scheduling and staffing of telephone operators in call centers has made forecasting models an important component of decision making in many organizations (see, for example, Andrews and Parsons 1993, Gans et al. 2003). Some of the earlier forecasting models include autoregressive integrated moving average processes and transfer function models. These have been successfully used in forecasting call volumes (see, for example, Andrews and Cunningham 1995). More recent approaches involve the use of queuing models as in Jongbloed and Koole (2001), and doubly stochastic Poisson models as in Avramidis et al. (2004). The focus of this previous work was to model the call arrival (demand) process based on aggregate arrival data.

In addition to modeling the call center arrival process, it is desirable to evaluate the efficiency of and return on advertisement and promotion policies to develop marketing strategies. This requires an advertisement-specific analysis of arrival patterns rather than analysis of aggregate arrival data. In a recent project, we were faced with a need to develop models and methods for assessing the impact of individual advertising campaigns of print media on call arrivals. The predictions produced by such models have significant implications for marketing and advertising budget decisions. The models and the statistical methodology presented here are motivated by this project.

The primary objective of modeling and analysis of call center arrival data from a marketing point of view is to assess the effectiveness of different advertising strategies and promotion policies. The model should account for relevant covariate and time effects on the call arrival process. It should also be able to describe potential heterogeneity in advertisements beyond what is captured by the covariates. The ability of the model to predict call volume generated by a specific advertisement over any desired time interval would provide valuable information.

In this paper, we present a modulated Poisson process model (MPPM) that provides us with a framework to address the above issues and develop its Bayesian analysis. Unlike the previous models used in call arrival analysis, the proposed model allows for advertisement-specific analysis of call arrival data, and thus enables us to assess the effectiveness of various advertisement and promotion strategies. In modeling daily call arrival data, Avramidis et al. (2004) have noted that arrivals during different time partitions are correlated. They have proposed Poisson models with stochastic arrival intensities to deal with this problem. In our development, by taking a Bayesian approach, we describe our uncertainty about parameters of the MPPM probabilistically via specifying a prior distribution. This approach yields a model that can be considered as a doubly stochastic Poisson model as used in Avramidis et al. (2004), and this results in correlated call arrival counts. Thus, the Bayesian treatment of the MPPM provides an alternate modeling approach to describe correlated call arrival data. Furthermore, this introduces a methodology for statistical analysis of call center operations as Gans et al. (2003) alluded to as a potential research prospect.

In recent years, the Bayesian approach has been considered in many marketing-related problems; see Rossi and Allenby (2003) for a review. As pointed out by the authors, the Bayesian methods provide a more flexible framework in dealing with heterogeneity. As discussed in Allenby and Rossi (1999), Bayesian models have been used in dealing with consumer heterogeneity, but they have not been considered in the analysis of call center arrival data. In this paper, as an extension of the MPPM, we introduce a random-effects-type model to describe potential heterogeneity in advertisements and present Bayesian inference for the model.

In what follows, we present the call arrival data and introduce an MPPM to describe the arrival pattern of calls over time by accounting for covariate effects. This is done in §2. This model enables us to assess the effectiveness of different advertising venues and of various advertising strategies, and provides a framework for predicting call volume over time. In §3, we present a Bayesian analysis of the model and discuss how posterior and predictive distributions are obtained in such an analysis. To consider heterogeneity in different advertisements, a random-effects-type extension of the model is considered in §3.1 and a model comparison approach is discussed §3.2. An illustrative example is given in §4 using real call center arrival data. Conclusions are presented in §5.

2. A Model for Call Center Arrivals

2.1. Call Arrival Data

The data we used in our analysis comes from a consumer electronics producer who offers a limited variety of products. A significant portion of the sales goes through the call center. Products usually have long life cycles, and the life cycle is usually extended by periodic updates and upgrades. With the aging product line, the advertisement budget has been increasing drastically.

The data is more detailed than what is commonly available from a typical call center. The time arrival history is available for each advertisement separately along with advertisement-specific data such as cost, frequency, and type. The company issues advertising in most media venues, however, the majority is in print media. Each advertisement is targeted and aims at urging the customer to place the call. The company tracks the origin of the call to the specific advertisement that led to the call. Because each incoming call is stamped with an advertisement ID, there is 100% linkage between each call and the advertisements.

Table 1 **An Example of Call Arrival Data for an Advertisement**

Time interval (in days)	Number of calls
(0, 1]	6
(1, 2]	5
(2, 3]	1
(3, 4]	3
(4, 5]	2
(5, 6]	2
(6, 7]	2
(7, 8]	0
(8, 9]	2
(9, 10]	2
(10, 18]	0

The advertisement ID typically points out the cost of advertisement, its medium, the advertisement format, and the type of promotion being offered. The incoming number of calls for each advertisement is available as interval counts for different periods over the life cycle of the advertisement. An example of call arrival data for a typical advertisement is shown in Table 1.

2.2. Modulated Poisson Process Model

In modeling call arrivals generated by an advertisement venue such as in Table 1, it is important to note that the effectiveness of the advertisement decreases by time. In view of this fact, the number of calls generated by an advertisement in a specified time interval can be described by a nonhomogeneous Poisson process (NHPP). As pointed out earlier, it is of interest to assess the effectiveness of different advertising strategies as well as advertising venues on call volume. Thus, these factors, as well as the time effect, should be considered by any model describing the call volume. Modeling the effect of covariates on the call volume intensity requires a modulation of the Poisson process, and this can be achieved as proposed in Cox (1972a).

Let $N_i(t)$ denote the number of calls that arrived during a time interval of length t as response to the ith advertisement, and let $\mathbf{Z}_i$ denote a $p \times 1$ vector of covariates that describe the characteristics of the ith advertisement. Typically, $\mathbf{Z}_i$ will consist of components such as media expense (in dollars), venue type (monthly magazine, daily newspaper, etc.), advertisement format (full, page, half page, color, etc.), offer type (free shipment, payment schedule, etc.), and seasonal indicators.

To reflect the fact that effectiveness of advertisement i is a function of time, $N_i(t)$ is described by a NHPP with intensity function

$$\lambda_i(t) = \frac{d}{dt} E[N_i(t)], \tag{1}$$

where $E[\cdot]$ denotes the expectation. The MPPM assumes that the intensity function of the ith advertisement is related to the covariate vector $\mathbf{Z}_i$ via

$$\lambda_i(t, \mathbf{Z}_i) = \lambda_0(t)\, e^{\boldsymbol{\beta}'\mathbf{Z}_i}, \tag{2}$$

where $\lambda_0(t)$ is the baseline intensity function and $\boldsymbol{\beta}$ is a $p \times 1$ vector of parameters. Note that (2) incorporates both the time and covariate effects on call intensity.

The MPPMs have been considered in survival analysis by Sinha (1993). The MPPM can be thought as a counting process alternative to the proportional hazards model (PHM) of Cox (1972b), and is sometimes referred to as the proportional intensities model (PIM) in reliability modeling literature (Merrick et al. 2005). Similar to the PHM, the PIM implies that ratio of intensity functions

$$\frac{\lambda_i(t, \mathbf{Z}_i)}{\lambda_j(t, \mathbf{Z}_j)} = e^{\boldsymbol{\beta}'(\mathbf{Z}_i - \mathbf{Z}_j)} \tag{3}$$

does not depend on time. The cumulative intensity (or the mean-value) function of the MPPM is given by

$$\Lambda_i(t, \mathbf{Z}_i) = \Lambda_0(t) e^{\boldsymbol{\beta}'\mathbf{Z}_i}, \tag{4}$$

where $\Lambda_0(t) = \int_0^t \lambda_0(s)\, ds$. Given $\Lambda_0(t)$, $\boldsymbol{\beta}$, and $\mathbf{Z}_i$, the distribution of $N_i(t)$, the call volume in any interval of length t is given by

$$P(N_i(t) = n \mid \Lambda_0(t), \boldsymbol{\beta}, \mathbf{Z}_i) = \frac{(\Lambda_0(t) e^{\boldsymbol{\beta}'\mathbf{Z}_i})^n}{n!} \exp(-\Lambda_0(t)\, e^{\boldsymbol{\beta}'\mathbf{Z}_i}). \tag{5}$$

Similarly, conditional on $\Lambda_0(\cdot)$, $\boldsymbol{\beta}$, and $\mathbf{Z}_i$, the probability distribution of number of calls in a time interval $(s, t]$ is given by the Poisson model

$$P(N_i(t) - N_i(s) = n \mid \Lambda_0(\cdot), \boldsymbol{\beta}, \mathbf{Z}_i) = \frac{(\Lambda_i(t, \mathbf{Z}_i) - \Lambda_i(s, \mathbf{Z}_i))^n}{n!} \times \exp\{-[\Lambda_i(t, \mathbf{Z}_i) - \Lambda_i(s, \mathbf{Z}_i)]\}. \tag{6}$$

In modeling the baseline intensity $\Lambda_0(t)$ of the modulated Poisson process, a parametric form can be specified. For example, an appropriate form for $\Lambda_0(t)$ in the analysis of call arrival data can be the *power law model* which is also used in reliability and survival analysis. This is given by

$$\Lambda_0(t) = \gamma\, t^{\alpha}, \tag{7}$$

where $\alpha > 0$ and $\gamma > 0$. In the power law model, values of $\alpha < 1$ imply that the effectiveness of an advertisement deteriorates with time. This is typically what is expected in call volume generated by a given advertisement. The specification of the above form also implies that the distribution of the time to the first call arrival is a Weibull density with shape parameter α. For this reason, the Poisson process with the power law intensity function is referred to as the *Weibull process*.

We can investigate the appropriateness of the power law model by obtaining a scatter plot of cumulative number of calls against time in the log scale. Taking logs on both sides of (7) yields

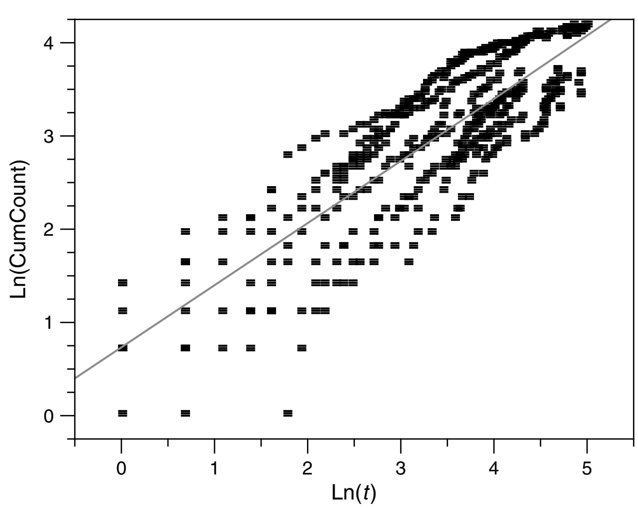
$$\log[\Lambda_0(t)] = \log \gamma + \alpha \log t. \tag{8}$$

The above implies that if the plot of the log of expected number of calls against the log of time is approximately linear, then the power law model is appropriate. In Figure 1, we show such a scatter plot for a group of monthly advertisements that are in a similar cost range. The linear fit that is shown in the figure suggests that the parametric form (7) is reasonable for the cumulative intensity function. In our illustration in §4, we will use the power law as the intensity function of the MPPM. However, it is possible to consider other parametric forms for $\Lambda_0(\cdot)$ or to use a nonparametric setup as discussed by Jongbloed and Verbaken (2004) for modeling the arrival process of problem reports on software products. Alternatively, a semiparametric Bayesian modeling strategy, as in Gelfand (1999), can be considered.

Recent studies in modeling aggregate daily arrivals to call centers have noted the inappropriateness of the standard Poisson process models with deterministic arrival intensities. Avramidis et al. (2004) point out that arrival counts in nonoverlapping time partitions are correlated and propose alternate stochastic models for arrival counts.

In what follows, we will present a Bayesian analysis of the MPPM whose baseline intensity function is given by the power law model (7). The Bayesian

Figure 1 **Scatter Plot of Cumulative Number of Calls vs. Time in Log Scale**



approach requires that uncertainty about all unknown parameters of the MPPM, that is, α, γ, and covariate parameter vector $\boldsymbol{\beta}$, is described probabilistically by specifying a prior distribution $p(\alpha, \gamma, \boldsymbol{\beta})$. Note that the results (5) and (6), as well as the independent increments property of the NHPP, hold only conditional on the parameters α, γ, and $\boldsymbol{\beta}$. In other words, unconditionally, the $N_i(t)$ process will not have independent increments. Thus, the Bayesian treatment of the MPPM provides an alternate modeling framework to alleviate the difficulties that arise in the analysis of call arrival data.

3. Bayesian Analysis of the Modulated Poisson Process Model

We assume that the call volume data is available in terms of interval arrivals as shown in Table 1, but the presented approach can easily be modified to incorporate time of arrivals data. For advertisement i, the process $N_i(t)$ is observed at r_i time intervals with endpoints $t = t_{i1}, \ldots, t_{ir_i}$ and $t_{i1} < \cdots < t_{ir_i}$, $i = 1, \ldots, m$, and the call volume data generated by advertisement i is given by

$$D_i = \{N_i(t_{ij}) = n_i(t_{ij}),\ j = 1, \ldots, r_i, \mathbf{Z}_i\},$$

where $\mathbf{Z}_i$ denotes the covariate vector associated with advertisement i. Using the independent increments property of the NHPP conditional on the parameters, given data on advertisement i, the likelihood function of $\Lambda_0(\cdot)$ and $\boldsymbol{\beta}$ is given by

$$L_i(\Lambda_0(\cdot), \boldsymbol{\beta}; D_i) = \prod_{j=1}^{r_i} \frac{([\Lambda_0(t_{ij}) - \Lambda_0(t_{i,j-1})]e^{\boldsymbol{\beta}'\mathbf{Z}_i})^{n_i(t_{ij}) - n(t_{i,j-1})}}{[n_i(t_{ij}) - n(t_{i,j-1})]!} \times \exp\{-([\Lambda_0(t_{ij}) - \Lambda_0(t_{i,j-1})]e^{\boldsymbol{\beta}'\mathbf{Z}_i})\}. \tag{9}$$

Conditional on $\Lambda_0(\cdot)$ and $\boldsymbol{\beta}$, the $N_i(t)$s are assumed to be independent. Thus, given the call counts for each $N_i(t)$ at $t_{i1}, \ldots, t_{ir_i}$, where $t_{i1} < \cdots < t_{ir_i}$, for $i = 1, \ldots, m$, the likelihood function of $\Lambda_0(\cdot)$ and $\boldsymbol{\beta}$ is

$$L(\Lambda_0(\cdot), \boldsymbol{\beta}; D) = \prod_{i=1}^{m} L_i(\Lambda_0(\cdot), \boldsymbol{\beta}; D_i), \tag{10}$$

where $D = (D_1, \ldots, D_m)$.

We will assume the power law model for the intensity function $\Lambda_0(\cdot)$. The Bayesian analysis requires specification of a joint prior distribution $p(\alpha, \gamma, \boldsymbol{\beta})$ for the power (α), and scale (γ) parameters of the intensity function; and for $\boldsymbol{\beta}$, the parameters of the covariate vector $\boldsymbol{\beta}$. Note that the values and signs of the coefficients in the $\boldsymbol{\beta}$ vector provide information about the effectiveness of a covariate in generating calls. For example, if the covariate is the media expense, then a positive value of the coefficient implies that higher levels of advertising expenses result in a higher level of call intensity. In our development, we will assume independent priors on these parameters. More specifically, uncertainty about the covariate vector $\boldsymbol{\beta}$ will be described by a multivariate normal prior, denoted as $\boldsymbol{\beta} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with a specified mean vector $\boldsymbol{\mu}$ and a covariance matrix $\boldsymbol{\Sigma}$. In specifying the prior parameters of this distribution, we can incorporate any expert judgment on the effect of media expenses, advertising strategy, etc. into the analysis; see Lindley (1983) for a formal framework for the use of expert opinion in Bayesian analysis.

In specifying the prior distribution of the parameters in the power law model, we note the fact that both α and γ are nonnegative random variables, and furthermore values of $\alpha < 1$ (>1) reflect the fact that the effectiveness of the advertisements deteriorates (improves) by time. In our development, we will specify a gamma prior on α and a lognormal prior on γ, but one can use other prior distributions for these parameters. For example, a lognormal prior can be used on α. Again, the prior parameters of these distributions can be chosen in such a way to reflect our knowledge or ignorance about them. Our experience with the application presented in §5 suggests that the results are robust to the selection of the prior distributions.

Given the above choice of priors or any other form, a fully Bayesian analysis of the MPPM cannot be obtained analytically. In other words, the joint posterior distribution

$$p(\alpha, \gamma, \boldsymbol{\beta} \mid D) \propto L(\Lambda_0(\cdot), \boldsymbol{\beta}; D) p(\alpha, \gamma, \boldsymbol{\beta}) \tag{11}$$

is not in an analytically tractable form. However, a Gibbs sampler similar to that used in Dellaportes and Smith (1993) for analysis of PHM can be developed to simulate from the posterior distribution (11). Once we have samples from the posterior distribution, we can obtain predictions of type (5) and (6). The Gibbs sampler enables us to simulate from (11) by recursively sampling from full conditional distributions $p(\alpha \mid \gamma, \boldsymbol{\beta}, D)$, $p(\gamma \mid \alpha, \boldsymbol{\beta}, D)$, and $p(\boldsymbol{\beta} \mid \alpha, \gamma, D)$; see Gelfand and Smith (1990) for a review of the Gibbs sampler and implementation details. Given the above selection of priors, the full conditional distributions for α, γ, and $\boldsymbol{\beta}$ cannot be obtained in any known form. Thus, sampling from these requires use of special simulation algorithms. However, it can be easily shown that all of these distributions are logconcave. Thus, the adaptive rejection algorithm of Gilks and Wild (1992) can be used for sampling from these distributions.

Given the posterior sample

$$\{\alpha^{(l)},\ \gamma^{(l)}, \boldsymbol{\beta}^{(l)}\}_{l=1}^{S}$$

from the joint posterior, we can make call arrival predictions for any advertisement for any time interval. Note that the posterior predictive distribution of number of arrivals in an interval of t time units for advertisement i is given by

$$P(N_i(t)=n \mid D) = \int P(N_i(t)=n \mid \alpha, \gamma, \boldsymbol{\beta}, \mathbf{Z}_i)\, dp(\alpha, \gamma, \boldsymbol{\beta} \mid D), \quad (12)$$

where $P(N_i(t) = n \mid \alpha, \gamma, \boldsymbol{\beta}, \mathbf{Z}_i)$ is given by (5). The above integral cannot be evaluated analytically, but we can compute it using a Monte Carlo integral approximation as

$$P(N_i(t)=n \mid D) \simeq \frac{1}{S}\sum_{l=1}^{S} P(N_i(t)=n \mid \alpha^{(l)}, \gamma^{(l)}, \boldsymbol{\beta}^{(l)}, \mathbf{Z}_i) \quad (13)$$

for $n=0,1,2,\ldots$. Also, we can obtain expected number of calls in a time interval as

$$\mathrm{E}(N_i(t) \mid D) \simeq \frac{1}{S}\sum_{l=1}^{S} E(N_i(t) \mid \alpha^{(l)}, \gamma^{(l)}, \boldsymbol{\beta}^{(l)}, \mathbf{Z}_i),$$

where

$$\mathrm{E}(N_i(t) \mid \alpha,\ \gamma,\ \boldsymbol{\beta}, \mathbf{Z}_i) = \Lambda_i(t, \mathbf{Z}_i) = \gamma t^{\alpha} e^{\boldsymbol{\beta}'\mathbf{Z}_i}.$$

Similarly, we can approximate the probability of n calls in the interval $(s, t]$, $s<t$, for advertisement i as

$$P(N_i(t)-N_i(s)=n \mid D) \simeq \frac{1}{S}\sum_{l=1}^{S} P(N_i(t)-N_i(s)=n \mid \alpha^{(l)},\ \gamma^{(l)},\ \boldsymbol{\beta}^{(l)}, \mathbf{Z}_i), \quad (14)$$

where $n=0,1,2,\ \ldots,$ and $P(N_i(t) - N_i(s) = n \mid \alpha, \gamma, \boldsymbol{\beta}, \mathbf{Z}_i)$ is given by (6). We can also evaluate the joint distribution of number of calls for any given type of advertisements using the conditional independence of $N_i(t)$s given $\alpha, \gamma, \boldsymbol{\beta}, \mathbf{Z}_i$, for $i=1,2,\ldots,m$.

As pointed out by one of the referees, in reality it is not uncommon to have cases where linkage between the ad and the calls may not exist for all calls. These ad-independent calls can be considered as if they were generated by the same call arrival process. The model presented in the above can be modified to account for such cases by introducing a random component to the baseline cumulative intensity in (7). Specifically, for all ad-independent calls, we assume that the cumulative intensity is given by

$$\Lambda_0^u(t) = \delta_u \Lambda_0(t) = \delta_u \gamma\, t^{\alpha}, \quad (15)$$

where δ_u is a random component which rescales the baseline intensity function (7) to reflect the behavior of all ad-independent calls. Because there is no covariate information for the ad-independent calls, the intensity function (4) reduces to (15) as δ_u replaces the covariate effect term $e^{\boldsymbol{\beta}'\mathbf{Z}_i}$ in (4). In the Bayesian approach, uncertainty about δ_u is described probabilistically implying that the joint prior $p(\alpha, \gamma, \beta, \delta_u)$ needs to be specified. Given that we observe r_u time intervals with endpoints $t_{u1} < \cdots < t_{ur_u}$, and the call arrivals $n_u(t_{uj})$, $j=1,\ldots,r_u$, we can write the likelihood function of γ, δ_0, and α for the ad-independent calls as

$$\prod_{j=1}^{r_u} \frac{\left(\gamma\delta_u\left(t_{uj}^{\alpha} - t_{u,\,j-1}^{\alpha}\right)\right)^{n_u(t_{uj})-n_u(t_{u,\,j-1})}}{[n_u(t_{uj}) - n_u(t_{u,\,j-1})]!} \times \exp\left\{-\gamma\delta_u\left(t_{uj}^{\alpha} - t_{u,\,j-1}^{\alpha}\right)\right\}.$$

By assuming conditional independence as before, the overall likelihood function is obtained as the product of the above term with (10). The Bayesian analysis of the new model can be developed using the Gibbs sampler to evaluate the posterior and predictive distributions. However, because our data consists of calls that are all linked to an ad, this model will not be used in our illustration in §4.

Even though the presented framework has emphasized call volume predictions for specific advertisements, it is possible to use our model for prediction of call volume generated by all ads at different time periods. Such predictions are more valuable for operations managers. If we consider K ads, then, given $\alpha, \gamma, \boldsymbol{\beta}$, and $\mathbf{Z}_i$, we have K conditionally independent NHPPs, $N_i(t)$s, with cumulative intensity functions $\Lambda_i(t, \mathbf{Z}_i)$, $i=1,\ldots,K$. Because operations managers may be interested in the cumulative number of calls generated by the K ads, we can use the fact that superposition of K independent NHHPs, is also an NHPP (cf. Çınlar 1975, pp. 87–88). That is, $N(t)=\sum_{i=1}^{K} N_i(t)$ is an NHPP with the cumulative intensity function given by

$$\Lambda(t,\ \mathbf{Z}) = \sum_{i=1}^{K} \Lambda_i(t, \mathbf{Z}_i) 1(t - t_{i0}), \quad (16)$$

where t_{i0} is the issue date of the ith ad, $1(\cdot)$ is the indicator function, and $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_K)$. The posterior predictive distribution of the cumulative number of calls can then be obtained via

$$P(N(t)=n \mid D) = \int P(N(t)=n \mid \alpha,\ \gamma,\ \boldsymbol{\beta}, \mathbf{Z})\, dp(\alpha, \gamma, \boldsymbol{\beta} \mid D).$$

The above can be computed by using a Monte Carlo integral approximation as in (13) by replacing $N_i(t)$ with $N(t)$. A numerical illustration of the posterior predictive distribution of the cumulative number of calls is presented in §4.

3.1. Modeling Heterogeneity in Advertisements

The MPPM of (4) implies that differences between the intensity functions of advertisements are adequately captured by the included covariates of the model. It is possible that in many cases, certain covariate information may not be available or the covariates included in the model will not reflect potential heterogeneity in different advertisements. Modeling heterogeneity in advertisements is essential not only for improving the predictive performance of the call arrival model, but also to assess the effectiveness of different advertising and promotion strategies in creating the call volume.

Recently, Bayesian models have been used in marketing literature for modeling consumer heterogeneity; see, for example, Allenby and Rossi (1999). However, these models have not included point process models such as the MPPM and they have not considered analysis of call center arrival data.

A common strategy to model heterogeneity is to consider a random-effects-type extension of the MPPM. This involves reparameterization of the intensity function as

$$\Lambda_0^i(t) = \gamma_i t^{\alpha}, \tag{17}$$

where the parameter γ_i, for $i = 1, 2, \ldots, m$, is assumed to be drawn from a known mixing distribution G. This type of extension results in a hierarchical Bayes-type setup in the sense of Lindley and Smith (1972). More specifically, in the intensity function (17), we define

$$\log \gamma_i = \theta + \phi_i, \tag{18}$$

where the ϕ_is are the random-effect terms. We assume that the ϕ_is are conditionally independent normal random variables

$$\phi_i \mid \tau \sim N(0, 1/\tau),$$

where the unknown precision is described by the gamma prior, $\tau \sim G(a_\tau, b_\tau)$, with specified parameters a_τ and b_τ. As before, α will have a gamma prior and the coefficient vector $\boldsymbol{\beta}$ is assumed to follow a multivariate normal prior. Furthermore, in this setup, we can also specify a normal prior for θ in (18), which can be incorporated into the multivariate normal prior of $\boldsymbol{\beta}$. In our setup, we assume that apriori $\boldsymbol{\beta}$, θ, and α are independent of the ϕ_is and τ.

Under the random-effects-type model, the posterior distribution of interest is

$$p(\alpha, \theta, \boldsymbol{\beta}, \boldsymbol{\phi}, \tau \mid D) \propto \prod_{i=1}^{m} L_i(\Lambda_0^i(\cdot), \boldsymbol{\beta}; D_i)\, p(\alpha, \theta, \boldsymbol{\beta}, \boldsymbol{\phi}, \tau), \tag{19}$$

where $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_m)$. As in the original model, the posterior distribution (19) cannot be obtained analytically. Thus, all the posterior and predictive analyses need to be done using a Gibbs sampler. In this case, the required full conditionals also include $p(\tau \mid \boldsymbol{\phi}, \alpha, \boldsymbol{\beta}, D)$ and $p(\phi_i \mid \phi^{(-i)}, \tau, \alpha, \boldsymbol{\beta}, D)$, where $\phi^{(-i)} = \{\phi_j \mid j \neq i, j = 1, \ldots, m\}$, as well as the full conditionals of α and $\boldsymbol{\beta}$.

It follows from the above setup that the full conditional distribution of τ can be obtained as a gamma distribution given by

$$(\tau \mid \boldsymbol{\phi}, \alpha, \boldsymbol{\beta}, D) \sim G\left(a_\tau + m/2, b_\tau + \sum_{i=1}^{m} \phi_i^2 \Big/ 2\right).$$

The remaining full conditional distributions do not follow any standard forms. However, because they are all logconcave, the adaptive rejection sampling method can be used to sample from these at each iteration of the Gibbs sampler. Similar to the original model, the predictive distributions (13) and (14) are obtained once posterior draws are available from the Gibbs sampler.

3.2. Model Comparison

In comparing the original MPPM with its random-effects extension presented above, computation of the Bayes factors (see Kass and Raftery 1995 for a comprehensive review) is difficult. The marginal likelihoods under the two competing models cannot be directly approximated from the Gibbs sampler to compute the Bayes factor.

An alternative is to use a model selection criterion such as the deviance information criterion (DIC) of Spiegelhalter et al. (2002). For a generic parameter vector Θ, DIC is defined as

$$\text{DIC} = \overline{\text{D}} + p_{\text{D}}, \tag{20}$$

where $\text{D} = -2\log \mathscr{L}(\Theta)$, is two times the negative log-likelihood, $\overline{\text{D}} = E_{\Theta\mid \text{data}}[\text{D}]$, and $p_{\text{D}} = \overline{\text{D}} - \text{D}(\hat{\Theta})$, where $\hat{\Theta}$ is the posterior mean. The DIC has the general "fit+complexity" form used by many model selection criteria. In (20), $\overline{\text{D}}$ represents the "goodness of the fit" of the model, where p_{D} represents a complexity penalty as reflected by the effective number of parameters of the model. Note that p_{D} also includes the contribution of the random-effects parameters to the model.

4. A Real-Life Data Illustration

In what follows, we present an illustration of the model using the real call center data. The data used in the illustration is only a subset of the real data set we have received from the company. Because of the company's specific and structured advertising policy, each call is traced to the advertisement that led the call. All the characteristics of the advertisement are known, such as its cost, specials offered in the advertisement, and its format. Our previous analysis indicated that the cost, which is named as media dollars, has a significant effect on the calls, both in its volume and its distribution. This variable is used as a covariate in the model. The results that are presented in this section are based on 142 different print media advertisements. The number of time intervals, where call volumes are available, vary between 3 and 99. In our analysis, we will use media dollars (in $000), that is, the cost for the advertisement, print media type (weekly or monthly), and the offer type as covariates. There will be a dummy variable which will capture the effect of media type and the monthly ads will be used as the reference group. There are three offer types and these are captured by two dummy variables in the model. The standard offer type will be used as the reference in the model, and the first and second offer types represent interest-bearing installment option and free originating and return shipment, respectively. In our model, we also include an interaction term of cost and media type to allow for differences in cost effect of different media type.

We first consider the original MPPM without the random-effect terms, and refer to this model as the *fixed-effects model* in our discussion. We note that for the case where we include the offer type and cost of the advertising as covariates, the log of the cumulative intensity function (4) can be written as

$$\log[\Lambda_i(t, \mathbf{Z}_i)] = \theta + \alpha \log t + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{1i} Z_{2i} + \beta_4 Z_{3i} + \beta_5 Z_{4i}, \quad (21)$$

where $\theta = \log\gamma$ in (7), Z_{1i} is cost of the advertisement, Z_{2i} is the dummy variable for media type, and Z_{3i} and Z_{4i} are the dummy variables representing the second and third offer types. The model also includes an interaction term of cost and media type. In (21), $\boldsymbol{\beta} = (\beta_1\, \beta_2\, \beta_3\, \beta_4\, \beta_5)$ is the covariate vector, and for the monthly advertisements with standard offer type, the cumulative intensity becomes

$$\Lambda_i(t, \mathbf{Z}_i) = \gamma\, t^{\alpha} \exp(\beta_1 Z_{1i}). \quad (22)$$

In our analysis, flat but proper priors are used for all the unknown parameters. Specifically, we use independent normal priors with mean zero and variance 100 for all β_ks and θ in (21). The prior for α is specified as a gamma distribution with mean one and variance 100. All the computations are done using the BUGS programming environment of Spiegelhalter et al. (1996). In implementation of the Gibbs sampler, the results are obtained after an initial burn-in sample of 10,000 iterations. We have experienced no convergence problems, and runs with different starting points gave us very similar results.

Unlike the classical approach, all inferences in Bayesian paradigm are described by posterior probability distributions of the unknown quantities. Summary statistics such as mean, mode, and variance of the posterior distributions can also be reported. In

Figure 2 **Posterior Distributions of α and β_1**

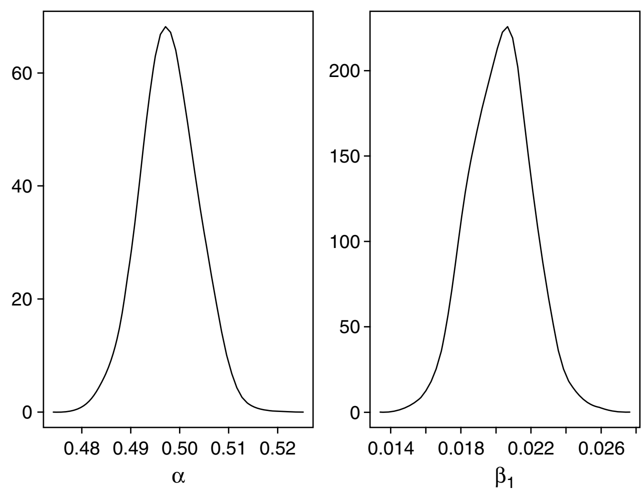


Figure 2, we present the posterior distributions of α and β_1 of the fixed-effects model. The posterior distribution of α which is concentrated in the interval (0.48, 0.52) indicates that the effectiveness of the advertisements deteriorates by time. In other words, for a fixed level of media expense and using a particular offer, as the advertisement ages, the expected number of calls during any time interval diminishes exponentially. This is an expected result and is consistent with the advertisement literature. In Figure 2, the value of β_1 shows the impact of each $1,000 spent on the total number of calls received due to that specific advertisement. The distribution of β_1 is concentrated in the interval (0.014, 0.026) indicating, as expected, that for a given offer type, increasing media expense implies an increase in the expected number of calls during any time interval.

The posterior distributions of β_2 and β_3 are shown in Figure 3. Both of these distributions are concentrated in the positive region, implying that weekly ads are more effective than the monthly ads in generating the call volume. Note that the positive values of β_3 represents a higher impact of ad cost for weekly ads on the expected number of calls received.

Posterior distributions of β_4 and β_5 that are shown in Figure 3 represent uncertainty about expected change in the call volume (in log scale) as a result of using a special offer. Based on the concentration of the posterior distribution of β_4 somewhere around (0.1, 0.3), the expected change in call volume due to the interest-bearing installment option seems to be positive relative to the standard offer. Similarly, as implied by the posterior distribution of β_5, the "free originating and return shipment" option also yields an expected increase in call volume beyond the standard offer.

Figure 4 displays the posterior predictive distributions of the number of calls for four different

Figure 3 **Posterior Distributions of β_is, $i = 2, \ldots, 5$**

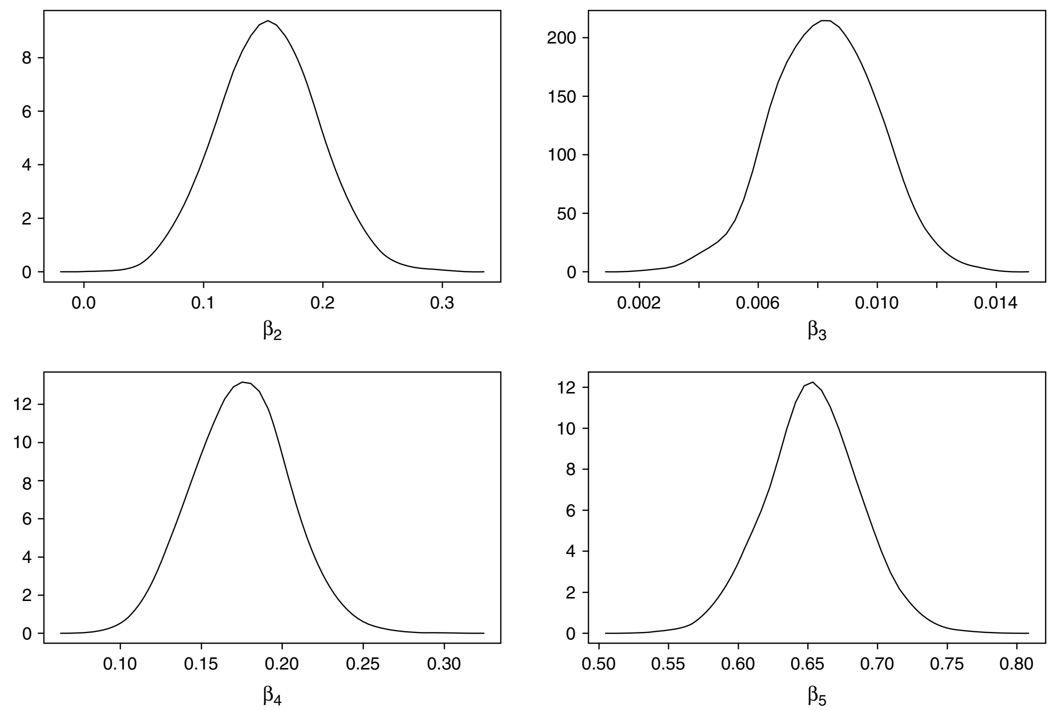


Figure 4 **Posterior Predictive Distributions of the Number of Calls for a Monthly Ad with $Z_1 = \$4{,}500$ and Standard Offer Type Under the Fixed-Effects Model**

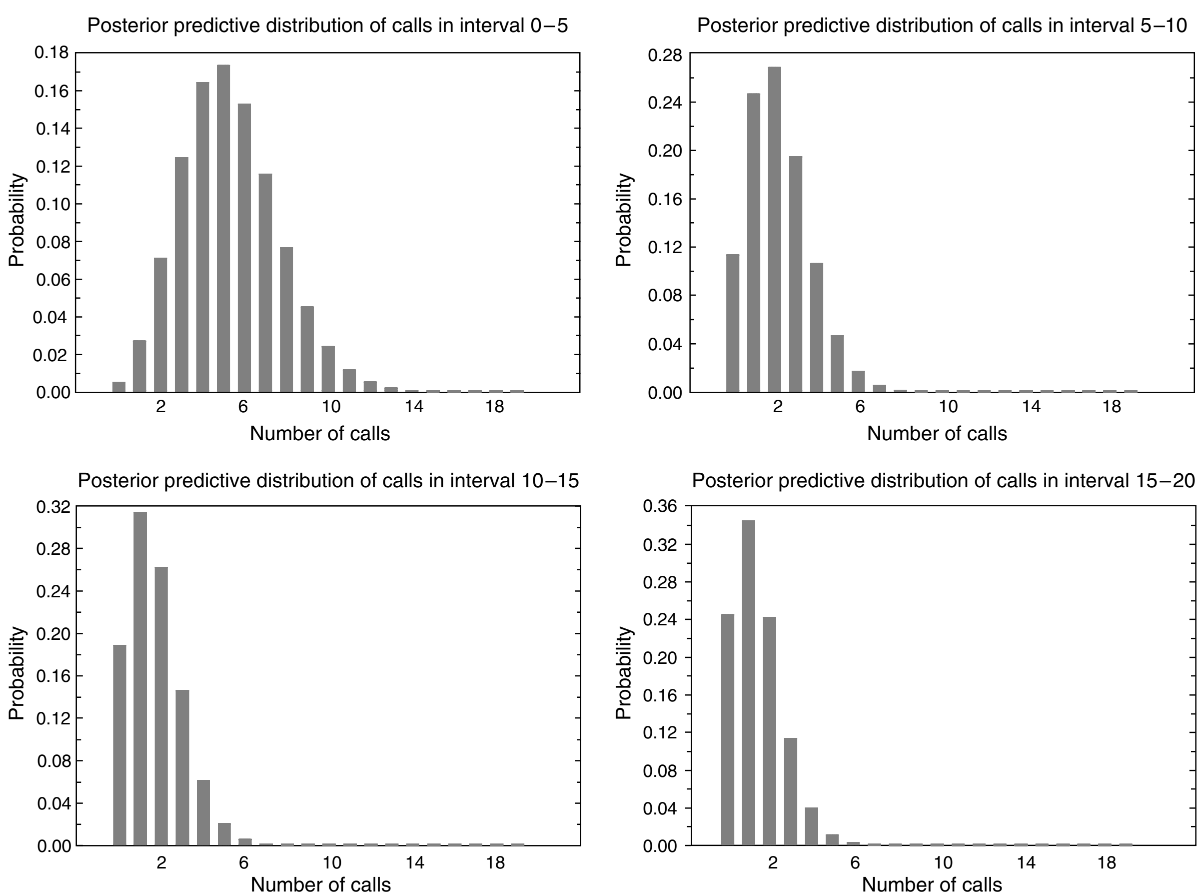
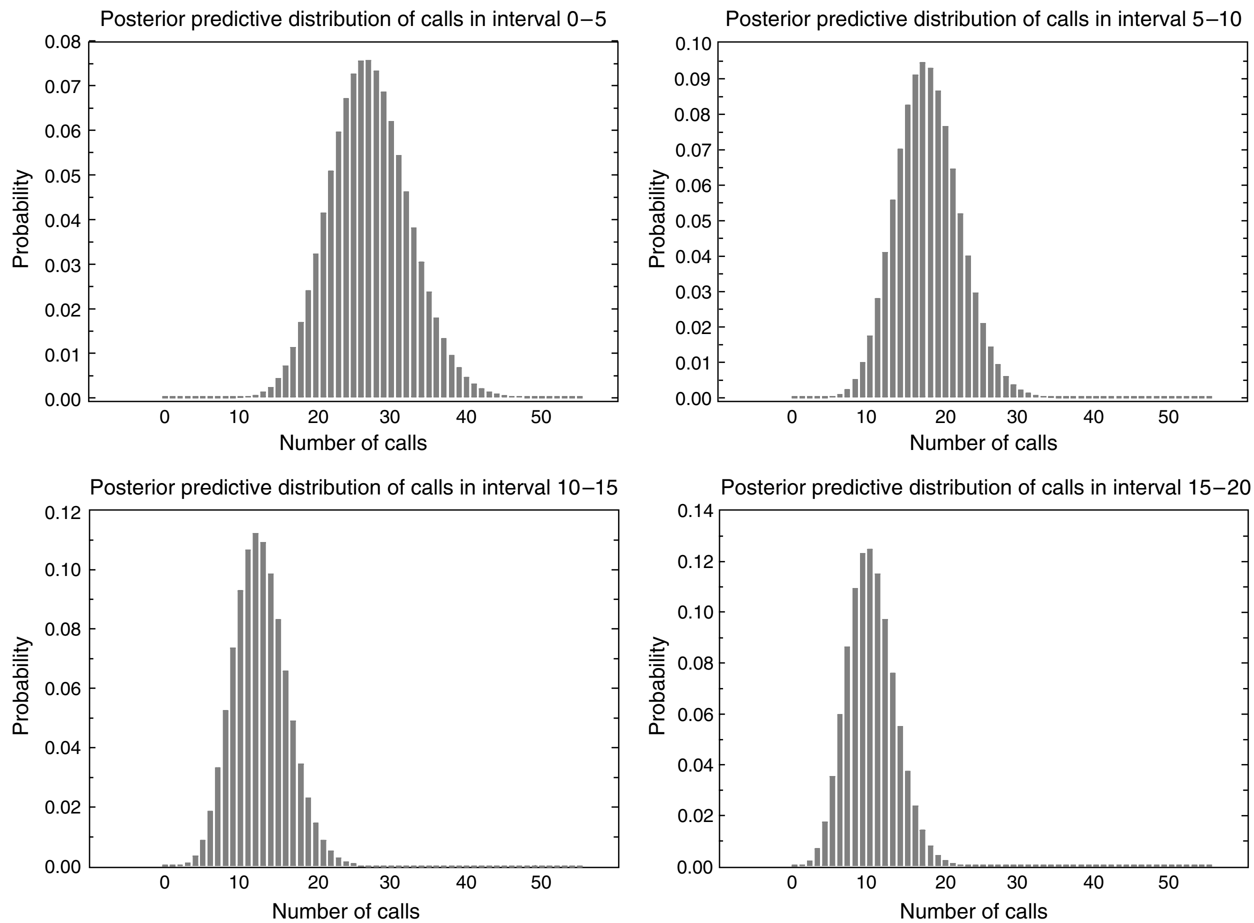


Figure 5 Posterior Predictive Distributions of the Total Number of Calls for a Set of Ads Under the Fixed-Effects Model



time intervals for a monthly advertisement that costs $4,500 and uses the third offer type, that is, free originating and return shipment option. Specifically, we present the distribution of the number of calls during the periods of 0–5, 5–10, 10–15, and 15–20 days. As expected, the distribution of the number of calls becomes more concentrated around smaller values as we move from 0–5 to 15–20 days, where the precision increases.

As discussed in §3, it is possible to obtain cumulative posterior predictions for a set of K ads using our model. In Figure 5, we present the posterior predictive distributions of the total number of calls for a set of four ads that are issued at different dates. The covariate data for the selected ads is presented in Table 2. Similar to Figure 4, the posterior predictive distributions are shown for periods of 0–5, 5–10, 10–15, and 15–20 days. As expected, the mean of these predictive distributions are much higher than the ones in the single-ad case.

We have done a similar analysis using the random-effects-type model of §3.1. In so doing, we again used proper but noninformative priors in all cases including the random-effect term ϕ_i and the precision parameter τ. We have used normal priors with mean zero and variance 100 for all fixed-effect parameters and θ. A gamma distribution with mean one and variance 100 was used for α and τ, the precision parameter of the random-effect terms. As before, an

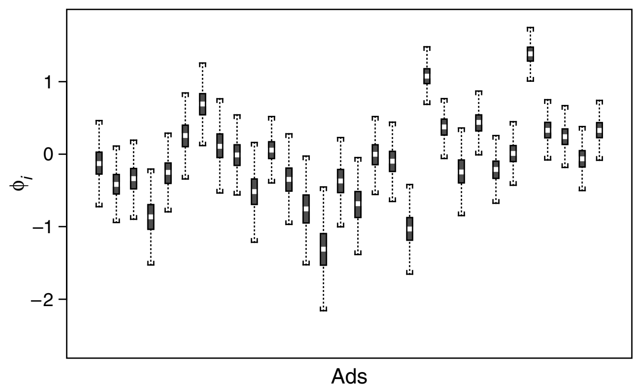
Table 2 Data on Ads Used for Total Number of Calls Prediction

Ad	Issue date	Media type	Cost ($)	Offer type
1	0	Weekly	7,500	2
2	1	Monthly	4,500	3
3	1	Monthly	3,700	2
4	7	Monthly	9,000	1

Table 3 Comparison of Posterior Means and Standard Deviations

	Fixed-effects model		Random-effects model	
	Mean	Std. dev.	Mean	Std. dev.
α	0.4977	0.0058	0.5012	0.0060
θ	0.7730	0.0484	0.5138	0.1593
β_1	0.0203	0.0017	0.0274	0.0069
β_2	0.1542	0.0408	0.0951	0.1601
β_3	0.0083	0.0017	0.0066	0.0082
β_4	0.1746	0.0289	0.1388	0.1231
β_5	0.6536	0.0331	0.8338	0.1459

Figure 6 **Posterior Distributions of ϕ_i for Selected Advertisements**



initial burn-in run of 10,000 iterations was used in the Gibbs sampler, and 2,000 samples were obtained using a lag of 25 between successive samples. No convergence problems were experienced in the Gibbs sampler. A comparison of the posterior means and standard deviations of the common parameters of the fixed- and random-effects models are given in Table 3. We note that in all cases, except the distribution of α, as expected, we have much higher variability under the random-effects model. The distribution of α is very similar in both cases. When we compare the distributions of β_5 under the two models, we see that the positive effect of the offer type two on call volume is more pronounced under the random-effects model. On the other hand, the effect of the first offer type on expected call volume, as implied by the distribution of β_4, is less pronounced. The posterior probability $\Pr(\beta_4 > 0 \mid D)$ can be obtained as 0.875 under the random-effects model.

Comparison of the posterior distributions of β_2 and β_3 show that the relative effectiveness of weekly ads in generating the call volume is less pronounced under the random-effects model. The posterior probability $\Pr(\beta_2 > 0 \mid D)$ can be obtained as 0.726 under the random-effects model, whereas under the fixed-effects model, the posterior distribution was concentrated in the positive region. Similarly, under the random-effects model, the posterior probability $\Pr(\beta_3 > 0 \mid D) = 0.795$ implies weaker evidence for the higher impact of ad cost for weekly ads on the expected number of calls received.

In Figure 6, we present the posterior distributions of the random-effects parameter ϕ_i for 30 arbitrar-

Figure 7 **Posterior Predictive Distributions of the Number of Calls for a Monthly Ad with $Z_1 = \$4{,}500$ and Standard Offer Type Under the Random-Effects Model**

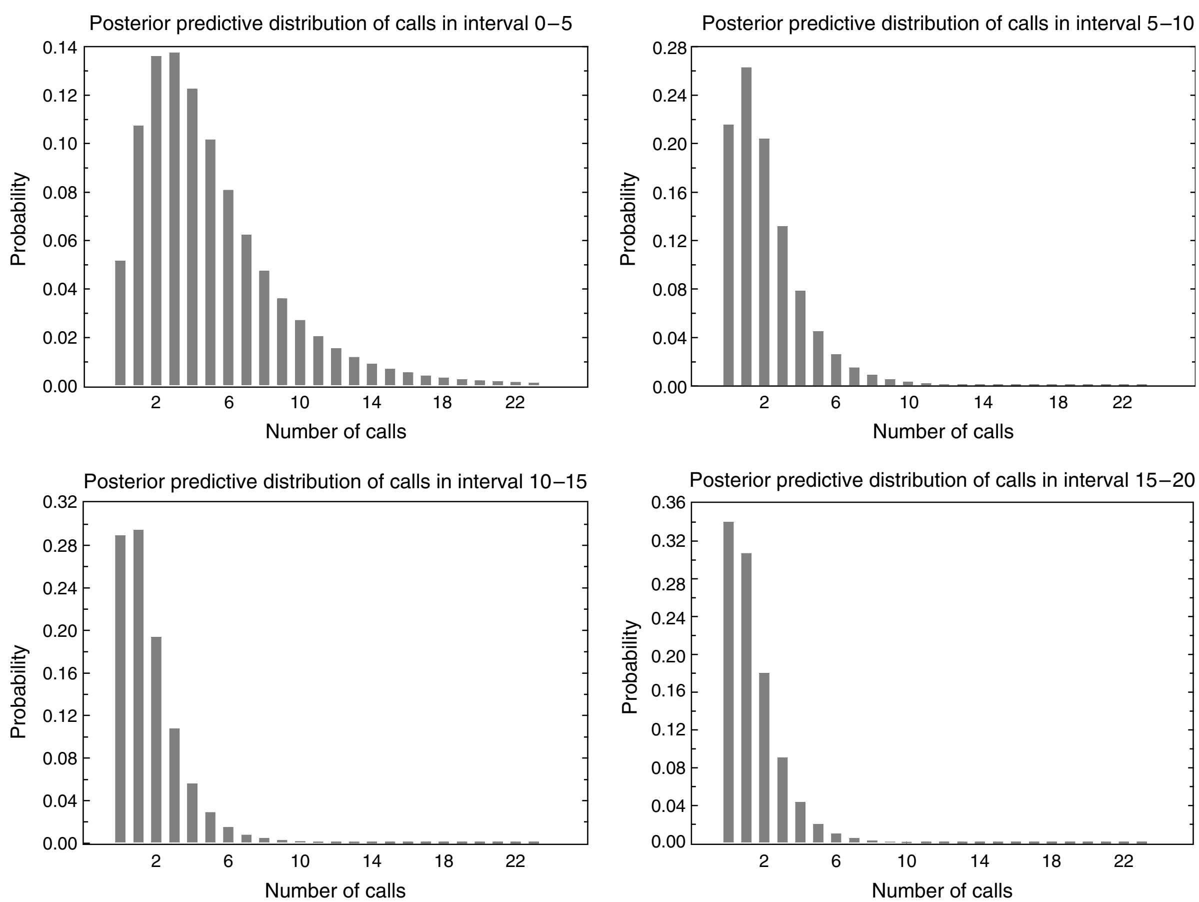
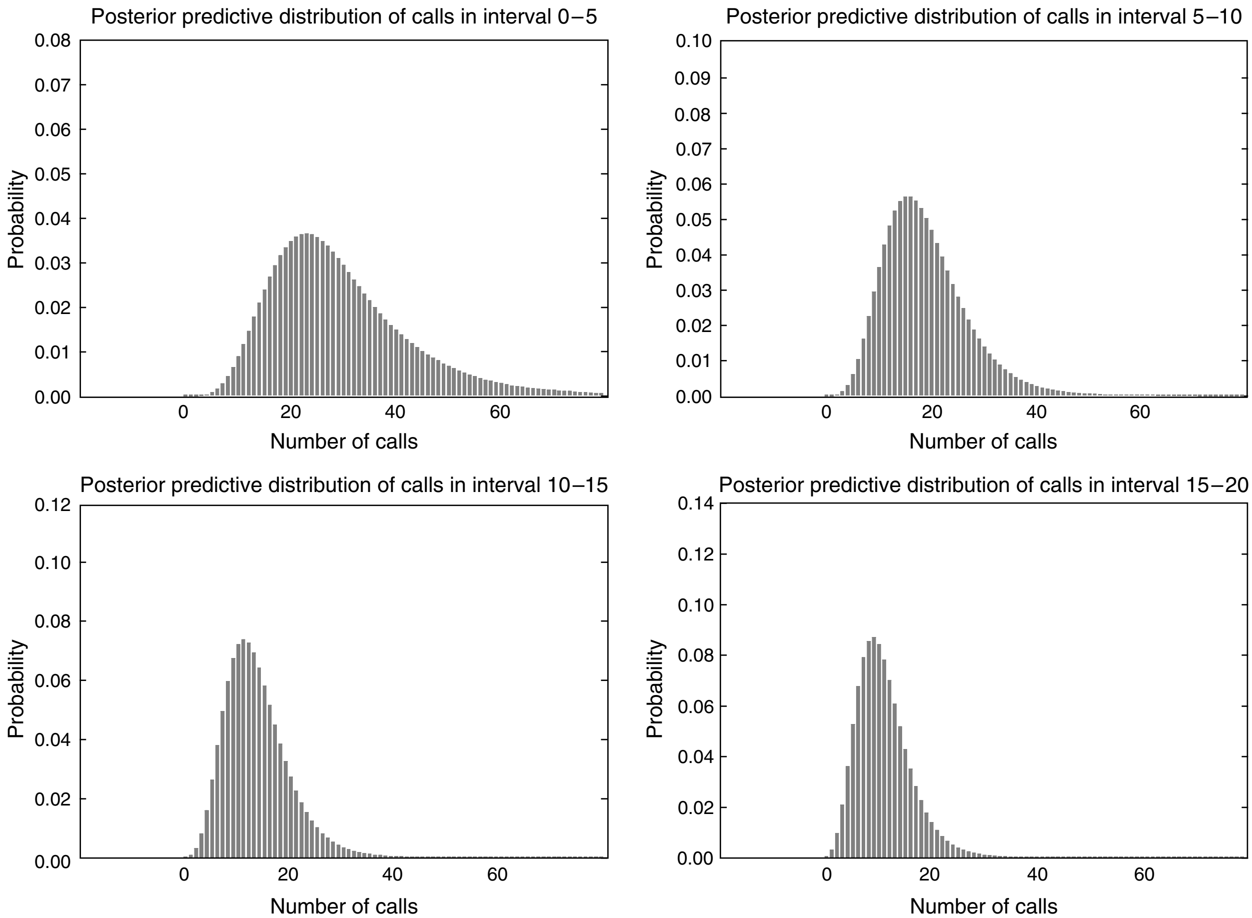


Figure 8 Posterior Predictive Distributions of the Total Number of Calls for a Set of Ads Under the Random-Effects Model



ily selected advertisements. We note that there are clear differences between these distributions. This suggests the presence of heterogeneity in intensity of call arrivals generated by specific advertisements.

These differences also yield a more diffused posterior predictive distribution for the number of calls during the intervals of 0–5, 5–10, 10–15, and 15–20 days as shown in Figure 7. If we compare the posterior predictive distributions given in Figure 7 with those of Figure 4, we see that the posterior predictive distributions under the random-effects model are more right skewed.

Similarly, we obtain more diffused posterior predictive distributions for the total number of calls for the set of four ads given in Table 2. These distributions are shown in Figure 8. As expected, they are more right skewed than the ones presented in Figure 5 based on the fixed-effects model.

4.1. Assessment of Model Fit and Model Comparison

A formal Bayesian comparison of model fit can be made by using the DIC introduced in §3.2. It may also be insightful to compare the quality of fit of Bayesian models presented here with those of classical approaches. However, such a comparison cannot be made using DIC or other Bayesian criteria. Thus, we will use some traditional fit measures based on prediction errors such as mean absolute error (MAE) and mean square error (MSE). We estimate the regression function $\log[\Lambda_i(t, \mathbf{Z}_i)]$ in (21) using a traditional least squares (LS) approach and predict the number of calls for each interval associated with each

Table 4 Comparison of the Models Using Fit Measures

Model	MAE	MSE
LS regression	1.651	7.622
Fixed-effects	1.444	5.138
Random-effects	1.346	4.218

Table 5 Percentage of Actual Calls Within Various Prediction Limits

Model	$\pm\sigma_c$ (%)	$\pm 2\sigma_c$ (%)	95% PI (%)
Fixed-effects	52.6	66.3	87.4
Random-effects	56.2	69.9	90.5

Table 6 Percentage of Actual Calls Within Prediction Limits Using Five-Day Intervals

Model	$\pm\sigma_c$ (%)	$\pm 2\sigma_c$ (%)	95% PI (%)
Fixed-effects	49.9	85.9	88.2
Random-effects	61.4	88.6	90.6

ad. For all 142 ads in our data set, the model provides predictions for 5,905 time intervals. Using these, we compute the MAE and MSE and compare them with their counterparts from the Bayesian models. The Bayesian predictions are based on the posterior predictive means of the number of calls for each interval. This comparison is reported in Table 4, which illustrates that the Bayesian models provide better fit to the actual data. Furthermore, this comparison suggests that the random-effects model performs better than the fixed-effects model.

In what follows, we will provide a more detailed comparison of these two models using Bayesian fit measures. In Table 5, we present the percentage of the actual number of calls that are within the one and two posterior standard deviations (σ_c) of predictive means for all 5,905 time intervals. Because the posterior predictive distribution of the number of calls are skewed, we also present the percentage of calls that are covered by 95% probability intervals (PI) for the two models.

Another comparison can be made on period basis for any prespecified time intervals for any ad. This is presented in Table 6 using five-day intervals for all ads. For illustrative purposes, in Figure 9, we present one σ_c limits around the posterior mean for a specific ad and the actual number of calls for each five-day interval.

Table 7 Comparison of the Models Using DIC

Model	DIC	p_D
Fixed-effects	25,430.4	6.9
Random-effects	23,810.0	129.6

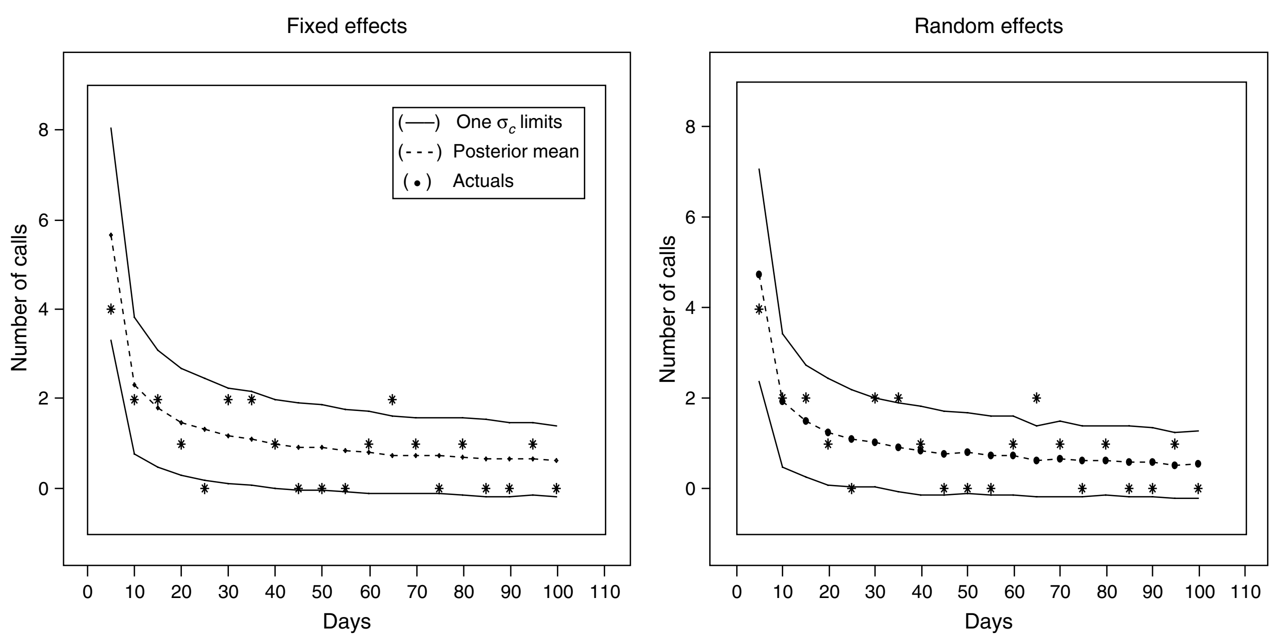
Although both models provide strong fit to the data, in all comparisons the random-effects model performs better.

A formal Bayesian comparison of the two models is made using the DIC in Table 7. We note that the DIC under the fixed-effects model is much higher than the one obtained under the random-effects model. In other words, the data provides strong support in favor of the random-effects model. When we look at the estimated values of the effective number of parameters, that is, the p_D values, we see that in the fixed-effects model this is 6.9, where the actual number is 7. On the other hand, under the random-effects model, the effective number of parameters is 130. We note that eight of these are the actual parameters and the remaining ones represent the random effects. In other words, among the advertisements we analyzed, there are differences that cannot be captured by the fixed covariates.

5. Concluding Remarks

In this paper, we present an MPPM for describing call center arrival data and develop its Bayesian analysis using Gibbs sampling. To incorporate potential heterogeneity in advertisements, we considered a

Figure 9 One σ_c Limits Around the Posterior Mean and the Actual Number of Calls for a Single Ad



random-effects-type extension of the model and discuss Bayesian model comparison.

The proposed approach and the models were implemented using real call arrival data on 142 advertisements, and the type of insights that can be obtained from the Bayesian analysis have been illustrated. The extension of the model to predict the total number of calls was considered, and its Bayesian implementation was illustrated for a set of selected ads.

The analysis of the data on 142 advertisements has shown that the effectiveness of advertisements deteriorates by time and certain offer types increase the call volume. Analysis of the random-effects model suggests the presence of heterogeneity in call volume generating ability of the advertisements. Our comparison of the fixed- and random-effects type models, using fit measures MAE and MSE as well as the DIC, has shown strong support for the random-effects model. The fixed-effects model also provides reasonably accurate results which are encouraging from a predictive point of view.

The presented analysis has assumed a power law model for the baseline cumulative intensity. The appropriateness of this choice for given call volume data can be investigated by considering alternate models for the baseline cumulative intensity and using DIC to compare these models. Alternatively, a semiparametric modeling strategy can be used by treating the baseline cumulative intensity $\Lambda_0(t)$ as a nonparametric form and by specifying a parametric prior for $\boldsymbol{\beta}$. This is an area for future research.

Another area of further study is the formal incorporation of prior opinion into the analysis of the MPPM. An extension of the methodologies presented in Campodonico and Singpurwalla (1995) for using expert opinion in analysis of point process models can be considered for this purpose.

Acknowledgments
The authors thank the editors and referees for their thorough reviews, which substantially improved this paper.

References

Allenby, G. M., P. E. Rossi. 1999. Marketing models of consumer heterogeneity. *J. Econometrics* **89** 57–78.

Andrews, B. H., S. M. Cunningham. 1995. L. L. Bean improves call-center forecasting. *Interfaces* **25** 1–13.

Andrews, B. H., H. L. Parsons. 1993. Establishing telephone-agent staffing levels through economic optimization. *Interfaces* **23** 15–20.

Avramidis, A. N., A. Deslauriers, P. L'Ecuyer. 2004. Modeling daily arrivals to a telephone call center. *Management Sci.* **50** 896–908.

Campodonico, S., N. D. Singpurwalla. 1995. Inference and predictions for Poisson point processes incorporating expert knowledge. *J. Amer. Statist. Assoc.* **90** 220–226.

Çınlar, E. 1975. *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs, NJ.

Cox, D. R. 1972a. The statistical analysis of dependencies in point processes. P. A. W. Lewis, ed. *Stochastic Point Processes*. Wiley, New York, 55–66.

Cox, D. R. 1972b. Regression models and life tables. *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

Dellaportes, P., A. F. M. Smith. 1993. Bayesian inference for generalized linear and proportional hazards models via Gibbs sampling. *Appl. Statist.* **42** 443–459.

Gans, N., G. Koole, A. Mandelbaum. 2003. Telephone call centers: Tutorial, review and research prospects. *Manufacturing Service Oper. Management* **5** 79–141.

Gelfand, A. E. 1999. Approaches for semiparametric Bayesian regression. S. Ghosh, ed. *Asymptotics, Nonparametrics, and Time Series*. Marcel Dekker, New York, 615–638.

Gelfand, A. E., A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* **85** 398–409.

Gilks, W., P. Wild. 1992. Adaptive rejection sampling for Gibbs sampling. *J. Roy. Statist. Soc. Ser. B* **41** 337–348.

Jongbloed, G., G. Koole. 2001. Managing uncertainty in call centers using Poisson mixtures. *Appl. Stochastic Models Bus. Indust.* **17** 307–318.

Jongbloed, G., A. M. H. Verbaken. 2004. Modelling the process of incoming problem reports on released software products. *Appl. Stochastic Models Bus. Indust.* **20** 131–142.

Kass, R. E., A. E. Raftery. 1995. Bayes factors. *J. Amer. Statist. Assoc.* **90** 773–777.

Lindley, D. V. 1983. Reconciliation of probability distributions. *Oper. Res.* **31** 866–880.

Lindley, D. V., A. F. M. Smith. 1972. Bayes estimates for the linear model (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 1–41.

Merrick, J. R., R. Soyer, T. A. Mazzuchi. 2005. Are maintenance practices for railroad tracks effective? *J. Amer. Statist. Assoc.* **100** 17–25.

Rossi, P. E., G. M. Allenby. 2003. Bayesian statistics and marketing. *Marketing Sci.* **22** 304–328.

Sinha, D. 1993. Semiparametric Bayesian analysis of multiple time data. *J. Amer. Statist. Assoc.* **88** 979–983.

Spiegelhalter, D., A. Thomas, N. Best, W. Gilks. 1996. *Bayesian Inference Using Gibbs Sampling Manual* (Version ii). MRC Biostatistics Unit, Cambridge University, Cambridge, UK.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, A. van der Linde. 2002. Bayesian measures of model complexity and fit. *J. Roy. Statist. Soc. Ser. B* **64** 1–34.