The Institute for Integrating Statistics in Decision Sciences

Technical Report TR-2017-7

Bayesian Modeling of Abandonments in Ticket Queues

Kaan Kuzu
Operations Management Business Statistics
University of Wisconsin-Milwaukee, USA

Refik Soyer
Department of Decision Sciences
The George Washington University, USA

Bayesian Modeling of Abandonments in Ticket Queues

Kaan Kuzu
Sheldon B. Lubar School of Business, Supply Chain Operations Management and Business Statistics,
University of Wisconsin-Milwaukee, Milwaukee, WI 53201, kuzu@uwm.edu

Refik Soyer
School of Business, Department of Decision Sciences,
The George Washington University, Washington DC, 20052, soyer@gwu.edu

Ticket queues arise often in public and private sector operations. The service providers in ticket queues have limited information on customer abandonments because abandonment time data is interval censored. The censored nature of data poses challenges in modeling and analysis of abandonments and in developing staffing policies. To alleviate such challenges, we present a Bayesian framework for analysis of abandonments in ticket queues. In doing so, we propose parametric and semi-parametric modulated Poisson process models to describe the abandonment behavior and develop their Bayesian analysis using Markov chain Monte Carlo methods. We implement our models on actual abandonment data from a bank' s ticket queue and illustrate how the proposed Bayesian framework can be used to provide insights for operations managers in predicting abandonments and in developing optimal server allocation policies.

Key words: ticket queue; Bayesian queues; Bayesian nonparametric; interval censored data; Markov chain Monte Carlo.

1. Introduction and Background:

Service providers desire to identify customer abandonments because abandonments result in opportunity costs (e.g. lost revenues and business). It is easy to achieve this in call centers or internet queues, where abandonments can be observed and recorded in real-time. However, other queuing arrangements, such as ticket queues, do not provide such capability. Ticket queues (TQ) are systems that issue tickets to the arriving customers upon arrival and are used commonly in public sector (e.g., Department of Motor Vehicles, government agencies) and private sector (e.g. banks, retail stores, hospitals). Customers are called for service in the increasing order of the ticket numbers, but do not form physical lines. Although TQs record customer arrival times, they cannot capture customer abandonment times. An abandonment is realized only when the ticket for the abandoned customer is called for service and that customer does not show up. As such, abandonment data collected by TQs is interval censored, which makes prediction of customer abandonments a challenging problem.

Modeling of TQs and their operational characteristics have gained limited attention in the literature. The original work by Xu et al. (2007) analyzes a single server ticket queue (in which customers abandon only by balking) where authors find that service completion rates in ticket queues can be much lower than that in physical (stand-in-line) queues. Kuzu (2010) extends this work to multi server ticket queues with both forms of abandonment (balking and reneging) by using Markov chain models. Xiao et al. (2015) use renewal reward theorem to study the problem of optimal staffing in ticket queues. Kuzu (2015) finds customers are willing to wait longer in ticket queues than in physical queues. Pender and Jennings (2016) develop fluid models for heavy traffic systems with patient customers and both forms of abandonment. They show that ticket and physical queues are asymptotically undistinguishable under such settings. Unlike the literature on patience related studies in call centers; see for example, Mandelbaum and Zeltyn (2004, 2009), Aksin et al. (2013, 2016) and Ibrahim et al. (2017), and in emergency rooms; see for example, Batt and Terwiesch (2015) and Bolandifar et al. (2016), there is a dearth of literature in modeling and analysis of abandonment data in TQs. Some of the more recent work include Kuzu et al. (2017) who study the effect of factors that customers observe during waits (e.g., number of servers operating or number of tickets ahead) on abandonment probability.

In this manuscript, we introduce a counting process model for describing abandonment behavior in TQs and for analyzing interval censored abandonment data. In doing so, we consider abandonment counts in non-overlapping time intervals as shown in Table 1. According to Table 1, there were no abandonments between 9:00:00AM and 9:13:11AM. This was followed by one abandonment between 9:13:12AM and 9:24:24AM. No abandonments took place between 9:24:25AM and 11:37:51AM, while two people abandoned between 11:37:52AM and 11:56:13AM.

Table 1 An example for Nonoverlapping Abandonment Intervals

Interval start	Interval end	Abandonment count
9:00:00	9:13:11	0
9:13:12	9:24:24	1
9:24:25	11:37:51	0
11:37:52	11:56:13	2

From a managerial point of view, the primary objective of modeling abandonment counts is to assess the effectiveness of staff allocation policies on the abandonment behavior of customers. The secondary objective is to develop a staff allocation policy that results in the lowest cost on a given day, considering the trade-off between staffing and abandonment costs. Such managerial decisions would significantly benefit from the assessment of customer abandonments. In this manuscript, our focus is not on developing an optimization model; rather, we focus on building a framework for predicting abandonment counts to assist managers in workforce planning. For this purpose, we propose a modulated Poisson process model that accounts for relevant time and covariate (e.g., day of the week or number of servers) effects on the abandonment process. We also introduce an extension of the proposed model by allowing a non-parametric intensity function and develop its Bayesian analysis.

Our contributions in this paper are threefold: First, to the best of our knowledge, we provide the first statistical study of abandonments in TQs using actual data. Second, we develop a semi-parametric Bayesian model to predict abandonment counts for any interval of a given day in TQs. The Bayesian model involves use of gamma processes; see van Noortwick (2009); which have not been previously considered in the queueing literature. Although there is a rich literature in Bayesian queueing analysis starting in 1990s; see for example, Armero (1994) and the more recent work by Ramirez-Cobo et al. (2010) and Aktekin and Soyer (2012), nonparametric Bayesian models for queues have not gained much attention until very recently. Further, the recent nonparametric approaches proposed by Aktekin (2014) and Ventz and Muliere (2014) have not considered queueing systems with abandonments. Thus, the semi-parametric Bayesian modeling of abandonments proposed here also presents a contribution to the Bayesian queueing literature. The ability to predict customer abandonments over any desired time interval would provide valuable input for determining optimal server allocation policies. Thus, our third contribution is providing a formal framework to assist managers in staffing decisions for TQs.

For the remainder of this manuscript, we first present a modulated Poisson process model for describing the abandonment process in TQs. Next, we introduce parametric and semi-parametric models, develop Bayesian analysis for the modulated Poisson process model and discuss how the model can be used for predicting future abandonments. Then, using a real ticket queue data set, we illustrate the implementation of the Bayesian models and provide managerial insights for server allocation. Finally, we summarize our findings and outline potential areas of future work.

2. A Model for Abandonments in Ticket Queues

2.1. Ticket Queue Data

The ticket queue data used in the study is obtained from two branches ("A_1 and B_1") of a Turkish bank. Each branch offers teller (e.g. cash/check deposit/withdrawal, currency exchange, etc.) services to walk-in customers. A customer arriving to a branch picks up a numbered ticket from a kiosk. Receipt of a ticket initiates a transaction record with fields shown in Table 2.

A new customer arrival creates the time stamp "*Taken time*" in the system. If any of the servers is idle, the newly arriving customer will start service immediately. If arriving customers find all servers to be busy, they may elect to join the queue or abandon immediately. Customers joining the queue may abandon after waiting some time or continue to wait until service start. Abandonment of a customer will be recognized only when that customer is called for service and does not show up.

Table 2 Data Fields for Transactions

Data Field	*Description*
Branch no	Branch the transaction is completed at.
Date	Date of the transaction in (DD/MM/YYYY) format.
Taken time	The time the customer picks up a ticket from the kiosk.
Call time	The time the server calls the ticket number for service.
Finish time	The time the transaction is completed.
Waiting time	Waiting time of the customer before being called for service (= Call time - Taken time).
Transaction time	The time it takes to complete the transaction (= Finish time - Call time).

When a server becomes idle, if there are tickets that have not been called for service yet, the server will call the next ticket number. This event creates the time stamp "*Call time*". On the other hand, completion of a customer transaction is reported by the time stamp"*Finish time*", but this entry does not provide us with complete information on whether the customer has completed service or not. Note that, "*Transaction time*" is reported for each customer. During our interviews, bank officials indicated that it usually takes at least 30 seconds for a customer to see her ticket number is called, approach the server to explain the transaction request and receive service. Thus, transactions taking less than 30 seconds are considered no show transactions in their internal reports. Therefore, we use the same assumption to distinguish the customers who receive service from those who abandon without receiving service. This assumption allows us to identify the number of abandonments in various time intervals within a day as shown in Table 1.

2.2. Modulated Poisson Process Model (MPPM)

Interval count data typically arise in applications in reliability (see for example; Merrick et al. (2005)) and survival analysis (see Sinha (1993)) literatures. Analysis of such data has also been considered by Soyer and Tarimcilar (2008) in modeling call center arrivals. It is common to model interval count data by a nonhomogeneous Poisson process (NHPP).

In modeling abandonments, it is important to capture the effect of waiting time on the abandonment behavior. There are two competing views in literature. The first view states that customers become more patient during their waits, i.e., the longer they wait, the more they become willing to stay until receiving service. This is described as *sunk cost effect* in literature (e.g., Kuzu 2015, Yu et al. 2017). The second view argues that the waiting time will frustrate the customers and make them impatient, leading to higher abandonment probability for longer waits. The effect of waiting time can be described by using a NHPP whose intensity function captures the two views. In addition, one is also interested in assessing the impact of factors like month of the year, day of the week and branch where the transaction takes place on the abandonment behavior. This can be incorporated into the model by modulating the intensity function of the NHPP via such covariates as suggested originally by Cox (1972). In what follows, we introduce a modulated Poisson Process model (MPPM) to describe the abandonments in TQs.

Let $N_i(t)$ denote the number of abandonments observed during a time interval of length t on a given day i and let $\mathbf{Z}_i$ denote a vector of covariates associated with the ith day which may also depend on the time interval for some covariates. We assume that $N_i(t)$ is described by a NHHP with cumulative intensity function $\Lambda_i(t)$. The MPMM assumes that the cumulative intensity function of day i is related to the covariate vector $\mathbf{Z}_i$ via

$$\Lambda_i(t, \mathbf{Z}_i) = \Lambda_0(t)e^{\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{Z_i}}, \tag{1}$$

where $\Lambda_0(t)$ is the baseline cumulative intensity function and $\boldsymbol{\beta}$ is a vector of coefficients. Note that (1) considers both the time and covariate effects on abandonment intensity and represents the expected number of abandonments up to time t during day i, that is, $E[N_i(t)] = \Lambda_i(t, \mathbf{Z}_i)$.

Given $\Lambda_0(t), \boldsymbol{\beta}$ and $\mathbf{Z}_i$, the distribution of the number of abandonments in any time interval of length t on day i, $N_i(t)$, is given by the Poisson model

$$P(N_i(t) = n|\Lambda_0(t), \boldsymbol{\beta}, \mathbf{Z}_i) = \frac{\left(\Lambda_0(t)e^{\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{Z_i}}\right)^n}{n!} \times \exp\left(-\Lambda_0(t)e^{\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{Z_i}}\right). \tag{2}$$

Moreover, conditional on $\Lambda_0(t), \boldsymbol{\beta}$ and $\mathbf{Z}_i$, the probability distribution of the number of abandonments in time interval $(s, t]$ on day i is obtained as

$$P(N_i(t) - N_i(s) = n|\Lambda_0(t), \boldsymbol{\beta}, \mathbf{Z}_i) = \frac{(\Lambda_i(t, \mathbf{Z}_i) - \Lambda_i(s, \mathbf{Z}_i))^n}{n!} \times \exp\left\{-\left[\Lambda_i(t, \mathbf{Z}_i) - \Lambda_i(s, \mathbf{Z}_i\right]\right\}. \tag{3}$$

We note that, conditional on $(\Lambda_0(t), \boldsymbol{\beta}, \mathbf{Z}_i)$, the MPPM (2) preserves the independent increments property of the NHPPs. However, by taking a Bayesian perspective and treating the cumulative intensity function (1) as an unknown quantity, unconditionally, we allow for correlated counts in the nonoverlapping time intervals in a given day. More specifically, as pointed out by Soyer and Tarimcilar (2008), unconditionally the $N_i(t)$ will have dependent increments.

2.3. Bayesian Modeling the Baseline Intensity Function

One strategy in Bayesian modeling of the cumulative intensity function (1) is to assume a parametric form for the baseline cumulative intensity $\Lambda_0(t)$ as in Soyer and Tarimcilar (2008). Alternatively, a nonparametric form can be used for $\Lambda_0(t)$ as considered in Merrick and Soyer (2017). In both cases, a parametric prior model is specified for the coefficient vector $\boldsymbol{\beta}$ of the MPPM.

The parametric Bayesian modeling strategy requires specification of a parametric form for $\Lambda_0(t)$ like the *power law* model

$$\Lambda_0(t; \boldsymbol{\theta}) = \alpha t^{\gamma}, \tag{4}$$

where $\boldsymbol{\theta} = (\alpha, \gamma)$ and $\alpha > 0, \gamma > 0$. The power law model (4) is widely used in modeling repairable systems; see for example, Pievatolo and Ruggeri (2004). It is also used in modeling call center arrivals by Soyer and Tarimcilar (2008). The baseline intensity function for (4) is given by

$$\lambda_0(t; \boldsymbol{\theta}) = \frac{d}{dt}\Lambda_0(t; \boldsymbol{\theta}) = \alpha\gamma t^{\gamma-1} \tag{5}$$

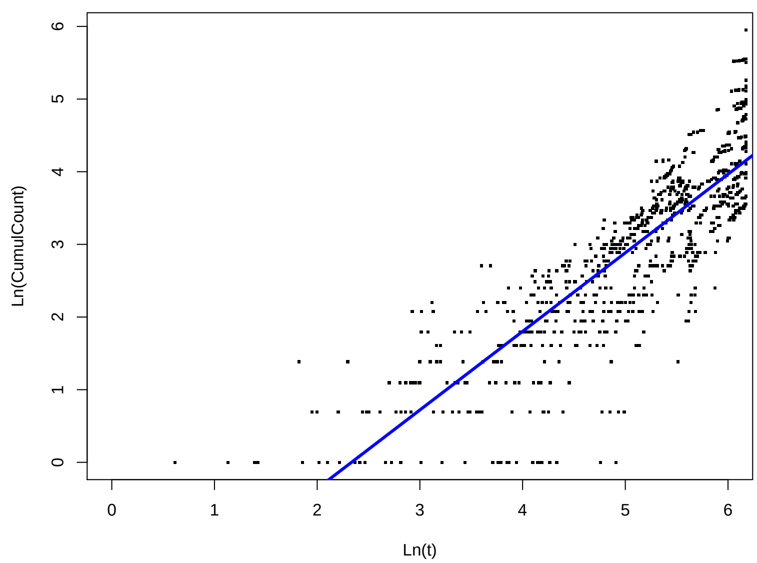
where $\gamma > 1$ ($\gamma < 1$) implies that customers' impatience increases (decreases) by time. Thus, the latter case would capture the sunk cost effect.

One way to infer the appropriateness of the power law model is via a scatter plot of the cumulative number of abandonments against time in the log scale, that is, by using the relationship

$$\log[\Lambda_0(t)] = \log\alpha + \gamma\log t. \tag{6}$$

Approximate linearity of such a plot will imply appropriateness of the power law model. A scatter plot of data from A_1 branch on Fridays, as seen in Figure 1, provides evidence of linearity in the log scale. Plots for the other days and branches provide similar forms, motivating us to use the power law model in our development and to compare it with the nonparametric model. Specification of the parametric Bayesian MPPM is completed by describing uncertainty about the unknown parameters $\boldsymbol{\theta}$ of the power law model and $\boldsymbol{\beta}$ of the covariates via a prior distribution $p(\boldsymbol{\theta}, \boldsymbol{\beta})$.

Figure 1 Sample Scatter Plot for Abandonments at A_1 branch on Fridays



As previously mentioned, an alternative modeling strategy is to consider a nonparametric form for the baseline cumulative intensity $\Lambda_0(t)$ (or equivalently for intensity $\lambda_0(t)$). In the Bayesian framework, this is achieved by treating the baseline cumulative intensity function $\Lambda_0(t)$ as a stochastic process and specifying a model that allows for a wide variety of different forms for $\Lambda_0(t)$. In doing so, we consider $\Lambda_0(t)$ as an unknown sample path of a stochastic process and specify prior distributions for the partitions of the sample path. This approach provides more flexibility than the parametric model by allowing more variability in the baseline cumulative intensity.

As pointed out by Merrick and Soyer (2017), because $\Lambda_0(t)$ is a nondecreasing function taking values in $[0,\infty)$ and there is no restriction on the size of any instantaneous jumps of the $\Lambda_0(t)$, a suitable stochastic process for $\Lambda_0(t)$ is a gamma process. By assuming a gamma process, we assume that all increments (that is, partitions) of the sample path $\Lambda_0(t)$ have gamma distributions and they are independent; see for example, Singpurwalla (1997).

We let $M(t)$ be a best guess for the baseline cumulative intensity function, c be a positive real number and assume that the distribution of any partition, $(t_{j-1}, t_j]$ of $[0,\infty)$ is given by

$$[\Lambda_0(t_j) - \Lambda_0(t_{j-1})] \backsim G(cM(t_j) - cM(t_{j-1}), c), \tag{7}$$

where $X \backsim G(a, b)$ denotes that X has a gamma distribution with shape parameter a and scale parameter b. It follows from this construction that $\Lambda_0(t)$ is a gamma process with $M(t)$ being a best guess and c is a measure of certainty about the best guess given the prior information D_0,

$$(\Lambda_0(t)|D_0) \backsim G(cM(t), c), \tag{8}$$

for all values of t. In our development, we will refer to (8) as the gamma process prior which implies that $E[\Lambda_0(t)|D_0] = M(t)$ and $V[\Lambda_0(t)|D_0] = M(t)/c$. Note that $M(t)$ and c are specified in (8) where higher values of c describes a strong prior belief in the mean function $M(t)$. A suitable choice for $M(t)$ can be the power law model (4). The Bayesian model is completed by specifying a prior distribution $p(\boldsymbol{\beta})$ on the covariate parameters which are assumed to be independent of the $\Lambda_0(t)$ a priori. Since the Bayesian modeling strategy consists of a nonparametric treatment of the baseline cumulative intensity and a parametric specification of the effect of covariates, the resulting Bayesian MPPM for abandonments is referred to as the *semi-parametric* Bayesian model.

3. Bayesian Analysis of the MPPMs

In the TQ system, for day i, the process $N_i(t)$ is observed at r_i time intervals with end points $(t_{i,0}, t_{i,1}, \ldots, t_{i,r_i})$ where $t_{i,0} < t_{i,1} < \ldots < t_{i,r_i}$ and $t_{i,0} = 0$. Let $n_{i,j}$ denote the number of abandonments observed on the jth interval $(t_{i,j-1}, t_{i,j}]$ of day i, that is, $n_{i,j}$ is the realization of $N_i(t_{i,j}) - N_i(t_{i,j-1})$. We denote the observed interval count data for day i by D_i which is given by

$$D_i = \{n_{i,j}, \mathbf{Z}_{i,j}, j = 1, \ldots, r_i\} \tag{9}$$

where $\mathbf{Z}_{i,j}$ denotes the covariate vector associated with the j th interval $(t_{i,j-1}, t_{i,j}]$ of day i. Using the independent increments property of the NHPP given the parametric intensity function $\Lambda_0(t;\boldsymbol{\theta})$ and $\boldsymbol{\beta}$, the likelihood function $L_i(\Lambda_0(t;\boldsymbol{\theta}), \boldsymbol{\beta}; D_i)$ for day i is given by the product of Poisson probabilities as

$$L_i(\Lambda_0(t;\boldsymbol{\theta}), \boldsymbol{\beta}; D_i) = \prod_{j=1}^{r_i} \frac{\left([\Lambda_0(t_{i,j};\boldsymbol{\theta}) - \Lambda_0(t_{i,j-1};\boldsymbol{\theta})]\, e^{\boldsymbol{\beta}^T \mathbf{Z}_{i,j}}\right)^{n_{i,j}}}{n_{i,j}!} \times \exp\left\{-\left([\Lambda_0(t_{i,j};\boldsymbol{\theta}) - \Lambda_0(t_{i,j-1};\boldsymbol{\theta})]\, e^{\boldsymbol{\beta}^T \mathbf{Z}_{i,j}}\right)\right\}. \tag{10}$$

We assume that conditional on $\Lambda_0(t;\boldsymbol{\theta})$ and $\boldsymbol{\beta}$ the $N_i(t)$s are independent NHPPs. Given prior information and interval count data from m days, i.e., given $D = (D_0, D_1, \ldots, D_m)$, the complete likelihood is obtained as a product of $L_i(\Lambda_0(t;\boldsymbol{\theta}), \boldsymbol{\beta}; D_i)$'s, that is,

$$L(\Lambda_0(t;\boldsymbol{\theta}), \boldsymbol{\beta}; D) = \prod_{i=1}^{m} L_i(\Lambda_0(t;\boldsymbol{\theta}), \boldsymbol{\beta}; D_i). \tag{11}$$

Because $\Lambda_0(t;\boldsymbol{\theta})$ has the parametric form (4) in our development, we can replace $L(\Lambda_0(t;\boldsymbol{\theta}), \boldsymbol{\beta}; D)$ with $L(\boldsymbol{\theta}, \boldsymbol{\beta}; D)$ in (11) which is obtained as a product of $L_i(\boldsymbol{\theta}, \boldsymbol{\beta}; D_i)$'s given by (10). The joint posterior distribution of $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ given D is obtained via the Bayes' law as

$$p\left(\boldsymbol{\theta}, \boldsymbol{\beta} | D\right) \propto L\left(\boldsymbol{\theta}, \boldsymbol{\beta}; D\right) p\left(\boldsymbol{\theta}, \boldsymbol{\beta}\right). \tag{12}$$

As pointed out by Soyer and Tarimcilar (2008), the posterior distribution (12) can not be obtained analytically for any given form of the prior $p\left(\boldsymbol{\theta}, \boldsymbol{\beta}\right)$, but Markov chain Monte Carlo (MCMC) methods such as Gibbs sampling can be used to draw samples from the joint posterior. Implementation of the Gibbs sampler involves successive draws from the full conditional posterior distributions $p\left(\boldsymbol{\theta} | \boldsymbol{\beta}, D\right)$ and $p\left(\boldsymbol{\beta} | \boldsymbol{\theta}, D\right)$. If these full conditionals are not available as in any known distributional forms, then samples can be drawn from $p\left(\boldsymbol{\theta} | \boldsymbol{\beta}, D\right)$ and $p\left(\boldsymbol{\beta} | \boldsymbol{\theta}, D\right)$ using either Metropolis-Hastings method; see for example, Chib and Greenberg (1995), or the adaptive rejection sampling algorithm of Gilks and Wild (1992). In our development, we will assume prior independence of $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ and use a Gibbs sampler with a Metropolis-Hastings step.

For the parametric MPPM, once G samples from the posterior distribution of $p(\boldsymbol{\theta}, \boldsymbol{\beta} | D)$ are obtained using the Gibbs sampler, we can obtain predictions for any interval in a given day. Assume that, given observed interval count data D, we want to predict number of abandonments for a time interval $(t_{p,j}, \, t_{p,j+1}]$ of day p. Then, the posterior predictive distribution of the number of abandonments can be approximated by the Monte Carlo average

$$P(N(t_{p,j+1}) - N(t_{p,j}) = n | D) \simeq \frac{1}{G} \sum_{g=1}^{G} P(N(t_{p,j+1}) - N(t_{p,j}) = n | \, \boldsymbol{\theta}^{(g)}, \, \boldsymbol{\beta}^{(g)}). \tag{13}$$

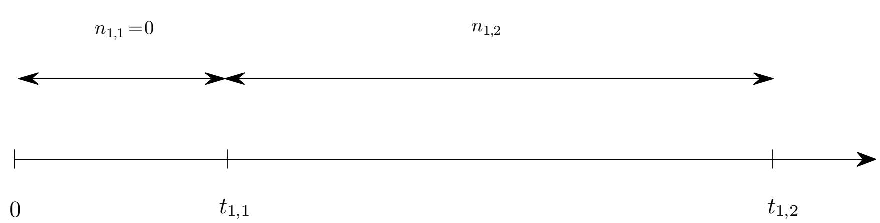
The above approximation can be used to compute posterior predictive distributions for any interval of day p.

3.1. Bayesian Inference for the Semi-Parametric Model

Gamma process priors in nonparametric Bayesian analysis of NHPPs have been considered by Sinha (1993) and Kuo and Ghosh (2001). Kuo and Ghosh (2001) assume a gamma process for the cumulative intensity function of a NHPP and show that the posterior distribution for the cumulative intensity function is obtained also as a gamma process if all event times of the NHPP are observed. Merrick and Soyer (2017) have noted that, for interval count data, the posterior distribution is not a gamma process and thus the Bayesian updating with a gamma process prior is not straightforward. The authors have proposed a Gibbs sampler with a data augmentation step to be able to draw samples from the posterior distribution of the baseline cumulative intensity function $\Lambda_0(t)$. In what follows, we adapt the data augmentation approach of Merrick and Soyer (2017) to our case and discuss modifications for developing predictions from the MPPM.

Since we use a nonparametric model for the baseline cumulative intensity, $\Lambda_0(t_{i,j};\boldsymbol{\theta})$ terms will be replaced by $\Lambda_0(t_{i,j})$ in (10), which yields a likelihood function $L(\Lambda_0(t),\boldsymbol{\beta};D)$, given data D. Suppose that for day i the process $N_i(t)$ is observed for two intervals $(0,t_{1,1}]$ and $(t_{1,1},t_{1,2}]$ so that the data is given by $D_i=\{n_{1,1},n_{1,2}\}$ as shown in Figure 2, where $n_{1,1}=0$ for the interval $(0,t_{1,1}]$. Assuming no covariate information is available for day i, the cumulative intensity for the ith day

Figure 2 The data observed for a single day.



is given by $\Lambda_i(t)=\Lambda_0(t)$, where $\Lambda_0(t)$ is defined by the gamma process prior in (8). Using the independent increments property of the NHPP, the likelihood function of $\Lambda_0(t)$ given D_i is written as

$$L_i(\Lambda_0(t);D_i)=\prod_{j=1}^{2}\frac{(\Lambda_0(t_{1,j})-\Lambda_0(t_{1,j-1}))^{n_{i,j}}}{n_{i,j}!}\times\exp\{-(\Lambda_0(t_{1,j})-\Lambda_0(t_{1,j-1}))\}, \quad (14)$$

where $t_{1,0}=0$.

By using the independent increments property of the gamma process and the Bayes' law, the posterior distribution of $\Lambda_0(t)$ is obtained as

$$(\Lambda_0(t)|D_i)\backsim G(cM(t),c+1),\ \text{for } t\le t_{1,1} \quad (15)$$

$$(\Lambda_0(t_{1,2})-\Lambda_0(t_{1,1})|D_i)\backsim G(c(M(t_{1,2})-M(t_{1,1}))+n_{1,2},c+1),\ \text{for } t_{1,1}<t\le t_{1,2} \quad (16)$$

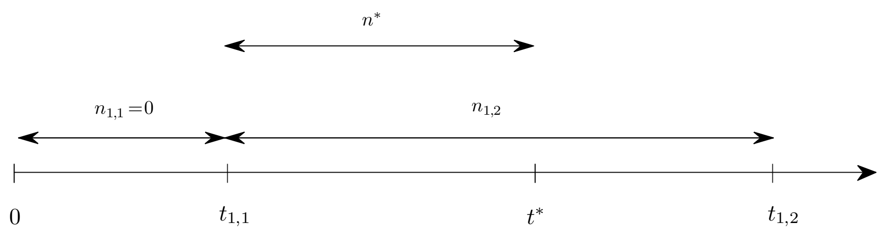
$$(\Lambda_0(t)-\Lambda_0(t_{1,2})|D_i)\backsim G(c(M(t)-M(t_{1,2})),c),\ \text{for } t>t_{1,2}. \quad (17)$$

Thus, the distribution of $\Lambda_0(t_{1,2})$ can be obtained as the sum of two independent gamma random variables, that is,

$$\Lambda_0(t_{1,2})=[\Lambda_0(t_{1,2})-\Lambda_0(t_{1,1})]+\Lambda_0(t_{1,1}). \quad (18)$$

Suppose that, for prediction purposes, the posterior distribution of $\Lambda_0(t^*)$ is required, where $t_{1,1} < t^* \le t_{1,2}$, and, n^* denoting the number of abandonments between $t_{1,1}$ and t^* is unknown as shown in Figure 3. As pointed out by Merrick and Soyer (2017), one can obtain the posterior distribution

Figure 3 The prediction problem for a single day.



of $\Lambda_0(t^*)$ via data augmentation; see for example, Tanner and Wong (1987). Specifically, if $N_i(t^*) - N_i(t_{1,1}) = n^*$ is known, then the distribution of $\Lambda_0(t^*)$ can be updated as the sum of independent gamma random variables $[\Lambda_0(t^*) - \Lambda_0(t_{1,1})] + \Lambda_0(t_{1,1})$, where $(\Lambda_0(t_{1,1})|D_i)$ is given by (15) and

$$(\Lambda_0(t^*) - \Lambda_0(t_{1,1})|D_i, n^*) \backsim G(c(M(t^*) - M(t_{1,1})) + n^*, c+1). \tag{19}$$

Similarly, we can obtain

$$(\Lambda_0(t_{1,2}) - \Lambda_0(t^*)|D_i, n^*) \backsim G(c(M(t_{1,2}) - M(t^*)) + (n_{1,1} - n^*), c+1). \tag{20}$$

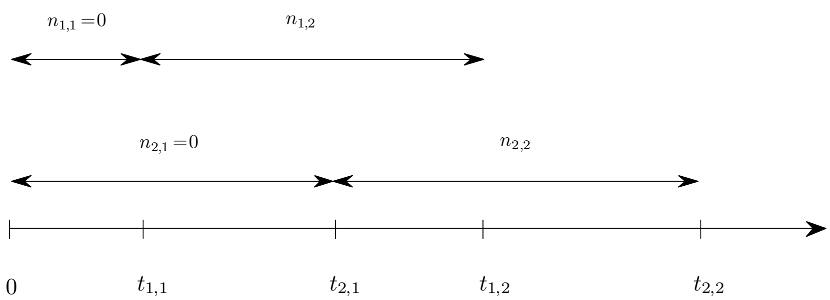
The conditional distribution of $(N_i(t^*) - N_i(t_{1,1}))$ given $\Lambda_0(t)$ and D_i can be obtained as a Binomial distribution

$$(N_i(t^*) - N_i(t_{1,1}) = n^*|\Lambda_0(t), D_i) \sim Bin\left(n_{1,2}, \frac{\Lambda_0(t^*) - \Lambda_0(t_{1,1})}{\Lambda_0(t_{1,2}) - \Lambda_0(t_{1,1})}\right), \tag{21}$$

using the properties of NHPP; see Ross (1989, p. 242). Following Merrick and Soyer (2017), a Gibbs sampler can be used to draw samples from the posterior distributions of $[\Lambda_0(t_{1,2}) - \Lambda_0(t^*)]$, $[\Lambda_0(t_{1,2}) - \Lambda_0(t^*)]$ and $[N_i(t^*) - N_i(t_{1,1})]$.

In the ticket queue abandonment data, there are cases where days are observed for different but overlapping intervals, as shown in Figure 4. Updating $\Lambda(t)$ given data then requires the use of a data augmentation step in the Gibbs sampler as discussed earlier. In Figure 4, we have $n_{1,1} = n_{2,1} = 0$ for days 1 and 2, but since the the observed counts $n_{1,2}$ and $n_{2,2}$ are coming from overlapping intervals, $(\Lambda_0(t_{1,2}) - \Lambda_0(t_{1,1}))$ and $(\Lambda_0(t_{2,2}) - \Lambda_0(t_{2,1}))$ cannot be updated separately. However, similar to the case of prediction in a single day, we can data augment using unobserved counts n_1^*

Figure 4 **The case of overlapping intervals for two days.**



and n_2^* as shown in Figure 5. We note that, conditional on the counts n_1^* and n_2^*, using the independent increments property of the gamma process, we can update the baseline cumulative intensity over $(0, t_{2,2}]$, as

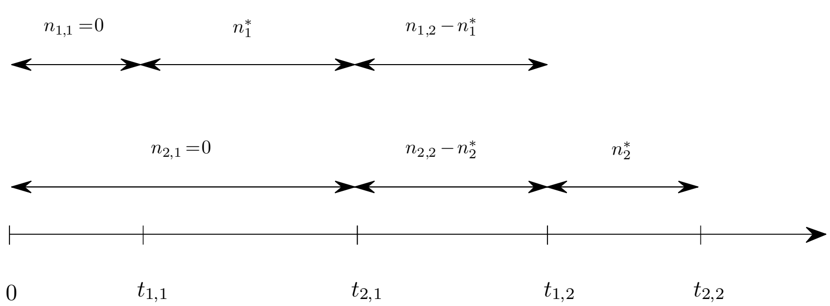
$$(\Lambda_0(t)|D) \backsim G(cM(t), c+2), \text{ for } t \le t_{1,1},$$

$$(\Lambda_0(t_{2,1}) - \Lambda_0(t_{1,1})|n_1^*, D) \sim G(c[M(t_{2,1}) - M(t_{1,1})] + n_1^*, c+2), \text{for } t_{1,1} < t \le t_{2,1},$$

$$(\Lambda_0(t_{1,2}) - \Lambda_0(t_{2,1})|n_1^*, n_2^*, D) \sim G(c[M(t_{1,2}) - M(t_{2,1})] + (n_{1,2} - n_1^*) + (n_{2,2} - n_2^*), c+2), \text{for } t_{2,1} < t \le t_{1,2},$$

$$(\Lambda_0(t_{2,2}) - \Lambda_0(t_{1,2})|n_2^*, D) \sim G(c[M(t_{2,2}) - M(t_{1,2})] + n_2^*, c+1), \text{for } t_{1,2} < t \le t_{2,2}.$$

Figure 5 **The abandonment counts required for data augmentation.**



Similar to the case of prediction in one day, conditional on the baseline cumulative intensity $\Lambda_0(t)$ over $(0, t_{2,2}]$, the latent counts can be obtained as independent binomial random variables given by

$$\left(N_1(t_{2,1}) - N_1(t_{1,1})|n_{1,2}, \frac{\Lambda_0(t_{2,1}) - \Lambda_0(t_{1,1})}{\Lambda_0(t_{1,2}) - \Lambda_0(t_{1,1})}\right) \sim Bin\left(n_{1,2}, \frac{\Lambda_0(t_{2,1}) - \Lambda_0(t_{1,1})}{\Lambda_0(t_{1,2}) - \Lambda_0(t_{1,1})}\right) \tag{22}$$

and

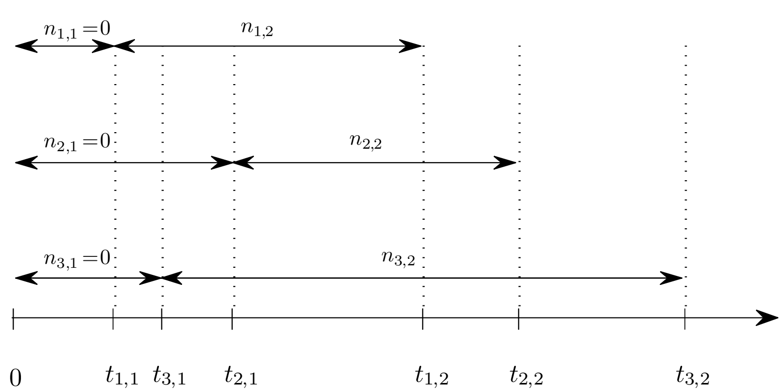
$$\left(N_2(t_{2,2}) - N_2(t_{1,2})|n_{2,2}, \frac{\Lambda_0(t_{2,2}) - \Lambda_0(t_{1,2})}{\Lambda_0(t_{2,2}) - \Lambda_0(t_{2,1})}\right) \sim Bin\left(n_{2,2}, \frac{\Lambda_0(t_{2,2}) - \Lambda_0(t_{1,2})}{\Lambda_0(t_{2,2}) - \Lambda_0(t_{2,1})}\right). \tag{23}$$

Draws from (22) and (23) facilitate the data augmentation step in the Gibbs sampler. We note that incorporation of the covariates into the model also requires draws from the full conditional distribution of $\boldsymbol{\beta}$, vector of covariate parameters. This can be achieved using a Metropolis-Hastings step within the Gibbs sampler as discussed in the parametric case. Incorporation of covariates will effect the full conditional distributions of $\Lambda_0(t)$ in the scale parameter. A more general data augmentation algorithm involving covariates will be discussed next.

3.2. A General Data Augmentation Algorithm Considering Covariates

The data augmentation for the semi-parametric model may become quite complicated as the number of overlapping intervals increases with the number of days and one needs a systematic approach. Following Merrick and Soyer (2017), an efficient way of doing this is to break the possible intervals into a partition defined by the endpoints of all intervals. This is illustrated in Figure 6 for the case of three different days each following a NHPP. For any observed interval that has now been broken up into sub-intervals, data augmentation is used on all but the one of these sub-intervals, because the number of abandonments in the remaining interval is known given the total number of abandonments in the whole interval and the number of abandonments in the other sub-intervals. For day i, the augmented abandonment counts will now follow a multinomial distribution. In the sequel, we now describe a general data augmentation algorithm that follows the

Figure 6 The case of three overlapping intervals.



approach of Merrick and Soyer (2017) considering covariates. In doing so, we also discuss how their approach can be modified for making abandonment predictions for regular time intervals such as hourly intervals.

We first determine the intervals that will be used for the data augmentation. Let $t_1^* < t_2^* < \cdots < t_q^*$ denote the q ordered values amongst the interval endpoints $t_{i,j}$ for $j = 1, \ldots, r_i$ and $i = 1, \ldots, m$. Also, we denote the unknown number of abandonments in the interval $[t_k^*, t_{k+1}^*)$ for day i by $N_{i,k}^*$

and define $B_k^* = \{i : t_k^* \leq t_{i,j} < t_{k+1}^*\}$ for $k = 1, \ldots, q$, as the set of days that have an abandonment count that spans the interval $[t_k^*, t_{k+1}^*)$. Let $S_{i,j}^* = \{t_k^* : t_{i,j-1} \leq t_k^* < t_{i,j}\}$ denote the set of all interval endpoints that fall in interval $[t_{i,j-1}, t_{i,j})$ and $m_{i,j}^* = |S_{i,j}^*|$ be the number of interval endpoints in this set. Furthermore, define the ordered list of members of $S_{i,j}^*$ by $\{\ell_{i,j}^1, \ldots, \ell_{i,j}^{m_{i,j}^*}\}$ with $\ell_{i,j}^{m_{i,j}^*+1} = t_{i,j}$.

For example, in Figure 5, we have two days of data each with two intervals, that is, $r_1 = r_2 = 2$, with $q = 5$ interval endpoints: $t_1^* = 0, t_2^* = t_{1,1}, \ldots, t_5^* = t_{2,2}$. These provide us with index sets $B_1^* = \emptyset$, $B_2^* = \{1, 2\}$, $B_3^* = \{1, 2\}$, $B_4^* = \{1, 2\}$ and $B_5^* = \{2\}$. For day 1, $N_{1,1}^* = 0$ and the number of unknown abandonments in $[t_2^*, t_3^*)$ is $N_{1,2}^*$ which is shown by n_1^* in Figure 5. $N_{1,3}^*$, the number of abandonments in $[t_3^*, t_4^*)$ is shown by $(n_{1,2} - n_1^*)$ and $N_{1,4}^* = 0$. Similarly, for day 2, we have $N_{2,1}^* = N_{2,2}^* = 0$, $N_{2,3}^* = (n_{2,2} - n_2^*)$ and $N_{2,4}^* = n_2^*$ in Figure 5. Furthermore, the set of interval endpoints are obtained as $S_{1,1}^* = \{t_1^*\}$, $S_{1,2}^* = \{t_2^*, t_3^*\}$, $S_{2,1}^* = \{t_1^*, t_2^*\}$ and $S_{2,2}^* = \{t_3^*, t_4^*\}$ giving us $m_{1,1}^* = 1$, $m_{1,2}^* = 2$, $m_{2,1}^* = 2$ and $m_{2,2}^* = 2$. It follows from the above that $\{\ell_{1,1}^1 = t_1^*\}$, $\{\ell_{1,2}^1 = t_2^*, \ell_{1,2}^2 = t_3^*\}$, $\{\ell_{2,1}^1 = t_1^*, \ell_{2,1}^2 = t_2^*\}$ and $\{\ell_{2,2}^1 = t_3^*, \ell_{2,2}^2 = t_4^*\}$.

Once we define the above components, and obtain $N^* = (N_{i,k}^*; i = 1, \ldots, m, k = 1, \ldots, q-1)$, then at each iteration of the Gibbs sampler we can draw from the full conditional distributions $p(\Lambda_0(t)|N^*, \boldsymbol{\beta}, D)$, $p(N^*|\Lambda_0(t), \boldsymbol{\beta}, D)$ and $p(\boldsymbol{\beta}|N^*, \Lambda_0(t), D)$. For drawing samples from $p(\Lambda_0(t)|N^*, \boldsymbol{\beta}, D)$, we can update $\Lambda_0(t_{k+1}^*) - \Lambda_0(t_k^*)$ by using the independent increments property of the gamma process. More specifically, given N^*,D and $\boldsymbol{\beta}$, vector of covariate parameters, we can easily show that

$$(\Lambda_0(t_{k+1}^*) - \Lambda_0(t_k^*) \, |N^*, \boldsymbol{\beta}, D) \backsim G(c[M(t_{k+1}^*) - M(t_k^*)] + \sum_{i \in B_{k+1}^*} N_{i,k}^*, c + \sum_{i \in B_{k+1}^*} e^{\boldsymbol{\beta}^T \boldsymbol{Z}_{i,k}^*}) \tag{24}$$

for $k = 1, \ldots, q-1$, where $\boldsymbol{Z}_{i,k}^*$ denotes the covariate vector associated with the interval $[t_k^*, t_{k+1}^*)$.

To obtain the full conditional distribution of $N_{i,k}^*$'s we consider the vector $\underline{N_{i,j}^*} = (N_{i,k}^* : t_k^* \in S_{i,j}^*)$ containing $N_{i,k}^*$'s that lie in the interval $[t_{i,j-1}, t_{i,j})$ for day i. Given $\Delta = \{\Lambda(t_{k+1}^*) - \Lambda(t_k^*); k = 1, \ldots, q-1\}$, using the properties of NHPPs we can obtain the full conditional of $\underline{N_{i,j}^*}$' as a multinomial given by

$$(\underline{N_{i,j}^*}|\Delta, D) \backsim \; Mult(n_{i,j}, p_{i,j,1}^*, \ldots, p_{i,j,m_{i,j}^*}^*), \tag{25}$$

where

$$p_{i,j,h}^* = \frac{\Lambda_0(\ell_{i,j}^{h+1}) - \Lambda_0(\ell_{i,j}^h)}{\Lambda_0(\ell_{i,j}^{m_{i,j}^*+1}) - \Lambda_0(\ell_{i,j}^1)}, \tag{26}$$

for $h = 1, \ldots, m_{i,j}^* - 1$. Obviously, if $m_{i,j}^* = 1$, then $N_{i,k}^* = n_{i,j}$. It is important to note that $\underline{N_{i,j}^*}$'s are drawn as independent multinomials across days i and intervals $j = 1, \ldots, r_i$, for a given day.

The full conditional posterior distribution of $\boldsymbol{\beta}$ is not available in a familiar form and therefore, a Metropolis-Hastings step can be used at each iteration of the Gibbs sampler to draw samples from $p(\boldsymbol{\beta}|N^*, \Lambda_0(t), D)$; see for example, Chib and Greenberg (1995).

Computation of posterior predictive distributions of the number of abandonments for a time interval $(t_{p,j}, t_{p,j+1}]$ of any day p in semi-parametric model requires extending the general data augmentation approach discussed above. More specifically, if we want to predict the number of abandonments in r_p nonoverlapping intervals $(t_{p,j}, t_{p,j+1}]$, $j = 1, 2, \ldots, r_p$, of day p then we define a new set of interval points $t_1^* < t_2^* < \cdots < t_q^*$ based on the observed data D as well as the r_p prediction intervals. Once this new set is defined, the posterior predictive distributions of $N(t_{p,j+1}) - N(t_{p,j})$ for $j = 1, 2, \ldots, r_p$ can be obtained using the Gibbs sampler with Metropolis-Hastings step by drawing from full conditional distributions (24), (25) and $p(\boldsymbol{\beta}|N^*, \Lambda_0(t), D)$.

4. Analysis of Actual Ticket Queue Data

In this section, we analyze the ticket queue data set using the parametric and semi-parametric models and illustrate the accuracy of the predictive models. The data set introduced in §2.1 contains 111 days of interval count data coming from 2 different branches ("A_1 and B_1") covering three months (October to December). The bank operates 9:00AM till 5:00PM Monday to Friday.

Table 3 presents the daily average abandonment counts for both branches across months. Within a given month, the number of days we have data for varies due to national holidays (e.g. religious festivals or public holidays). In Table 3, average number of abandonments on each day of the week is reported with the number of days we have data for given in parentheses. Mondays have higher abandonment counts as a result of the higher traffic intensity observed on those days.

Table 3 Daily Average Abandonment Counts

	A_1 Branch			B_1 Branch		
	October	November	December	October	November	December
Monday	316.5 (4)	214.8 (4)	275.3 (3)	154.5 (4)	143.0 (4)	164.0 (3)
Tuesday	96.5 (3)	81.3 (4)	104.0 (4)	66.0 (3)	29.0 (4)	75.5 (4)
Wednesday	105.0 (3)	95.0 (4)	101.5 (4)	64.3 (3)	48.0 (4)	103.5 (4)
Thursday	124.3 (3)	46.3 (4)	93.3 (3)	81.7 (3)	27.8 (4)	44.7 (3)
Friday	177.4 (5)	122.7 (3)	162.5 (4)	95.2 (5)	57.8 (4)	109.3 (4)

Next, we explain how to construct the abandonment count data on nonoverlapping intervals using a portion of data from Tuesday, December 30, at A_1 branch as shown in Table 4. Table 4a lists the time stamps for the abandoned customers. Using this data, we construct abandonment counts in each non-overlapping interval by finding the set of abandoned customers whose possible abandonment times within that interval do not overlap with possible abandonment times of any other customers in the consecutive non-overlapping interval. In other words, abandonment counts in a non-overlapping interval consists of the number of consecutively abandoned customers who take their tickets and are called for service within that interval. If no customers abandon within an interval, the abandonment count is 0. We illustrate such non-overlapping abandonment intervals by separating them with highlights in Table 4a. Table 4b shows the constructed list of intervals with their respective abandonment counts.

Table 4 Constructing Non-overlapping Intervals for Tuesday, December 30, at A_1 branch - (Portion of Day)

(a) Abandonment data

Taken time	Call time	Finish time
9:08:32 AM	9:12:43 AM	9:13:07 AM
9:17:16 AM	9:19:08 AM	9:19:31 AM
9:20:22 AM	9:21:06 AM	9:21:35 AM
9:38:37 AM	9:44:45 AM	9:44:57 AM
9:48:41 AM	9:50:51 AM	9:51:08 AM
9:56:19 AM	9:57:49 AM	9:58:14 AM
10:08:00 AM	10:10:06 AM	10:10:09 AM
10:13:59 AM	10:15:34 AM	10:15:52 AM
10:28:30 AM	10:30:03 AM	10:30:23 AM
10:32:29 AM	10:37:10 AM	10:37:21 AM
10:48:30 AM	10:55:20 AM	10:55:34 AM
10:55:12 AM	11:08:40 AM	11:08:49 AM
10:55:31 AM	11:08:49 AM	11:09:04 AM
10:56:46 AM	11:09:04 AM	11:09:13 AM
11:01:51 AM	11:12:25 AM	11:12:34 AM
11:02:22 AM	11:12:35 AM	11:12:44 AM
11:03:38 AM	11:14:03 AM	11:14:13 AM
11:05:48 AM	11:17:15 AM	11:17:37 AM
11:08:30 AM	11:17:49 AM	11:18:06 AM
11:08:41 AM	11:19:53 AM	11:20:09 AM
11:17:17 AM	11:21:06 AM	11:21:21 AM
11:20:54 AM	11:24:48 AM	11:25:09 AM
11:26:31 AM	11:36:38 AM	11:36:51 AM
11:27:47 AM	11:37:42 AM	11:37:49 AM
11:33:55 AM	11:41:53 AM	11:42:11 AM
11:48:03 AM	12:03:15 PM	12:03:35 PM
11:48:27 AM	12:03:39 PM	12:04:04 PM
11:53:31 AM	12:09:47 PM	12:09:57 PM
11:53:53 AM	12:09:59 PM	12:10:10 PM
11:53:59 AM	12:10:11 PM	12:10:19 PM
11:54:20 AM	12:10:14 PM	12:10:37 PM
11:55:16 AM	12:10:19 PM	12:10:24 PM
12:02:34 PM	12:13:54 PM	12:14:03 PM
12:07:29 PM	12:22:55 PM	12:23:20 PM
12:21:53 PM	12:33:45 PM	12:34:00 PM

(b) Non-overlapping intervals

Interval start	Interval end	Abandonment count
9:00:00 AM	9:08:31 AM	0
9:08:32 AM	9:12:43 AM	1
9:12:44 AM	9:17:15 AM	0
9:17:16 AM	9:19:08 AM	1
9:19:09 AM	9:20:21 AM	0
9:20:22 AM	9:21:06 AM	1
9:21:07 AM	9:38:36 AM	0
9:38:37 AM	9:44:45 AM	1
9:44:46 AM	9:48:40 AM	0
9:48:41 AM	9:50:51 AM	1
9:50:52 AM	9:56:18 AM	0
9:56:19 AM	9:57:49 AM	1
9:57:50 AM	10:07:59 AM	0
10:08:00 AM	10:10:06 AM	1
10:10:07 AM	10:13:58 AM	0
10:13:59 AM	10:15:34 AM	1
10:15:35 AM	10:28:29 AM	0
10:28:30 AM	10:30:03 AM	1
10:30:04 AM	10:32:28 AM	0
10:32:29 AM	10:37:10 AM	1
10:37:11 AM	10:48:29 AM	0
10:48:30 AM	11:24:48 AM	12
11:24:49 AM	11:26:30 AM	0
11:26:31 AM	11:41:53 AM	3
11:41:54 AM	11:48:02 AM	0
11:48:03 AM	12:33:45 PM	10

Following this approach, we obtain the abandonment counts on nonoverlapping intervals for each day. For example, Figure 7 displays the actual cumulative abandonment counts for Tuesday, December 30 at A_1 branch. In addition, Table 5 presents summary statistics for abandonment counts and nonoverlapping intervals for the whole data set.

Figure 7 Actual Abandonments for Tuesday, December 30 at A_1 branch

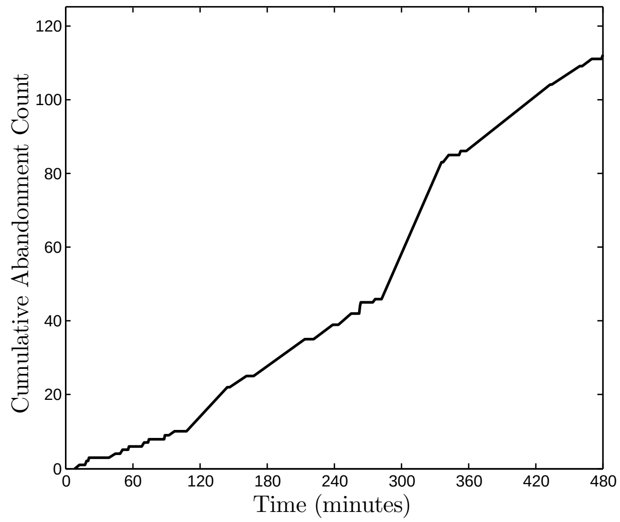


Table 5 Summary Statistics for Non-overlapping Abandonment Intervals Data

	A_1 Branch			B_1 Branch		
	October	November	December	October	November	December
Number of days	18	19	18	18	20	18
Total number of abandonments	3,158	2,117	2,578	1,730	1,222	1,779
Minimum no. of intervals on a day	4	5	2	2	2	2
Maximum no. of intervals on a day	69	71	58	35	41	35

4.1. Implementation of Bayesian Models

We analyze each branch' s data for each month separately using the parametric and semi-parametric models. We use the means of the posterior distributions of α and γ values obtained from the parametric runs to specify the power law model, which is used as the best guess $M(t)$ for the baseline cumulative intensity function in the semi-parametric model.

We consider three covariates in the MPPM for abandonment counts. First covariate is "day of the week"; we set Monday as the reference day due to the high abandonment counts (see Table 3) resulting from higher traffic intensity observed on Mondays and generate 4 dummy variables for capturing the effects of the remaining weekdays. The second covariate, $s_{i,j}$, is the "number of servers" on jth interval of day i and is time-dependent. In practice, the number of servers working on a given day is not fixed, but rather floating, and no information is available for these floating number of servers in the raw data. To estimate the server numbers, we first construct the chronological events list for all customers by sorting each *taken time*, *call time* and *finish time* record, and then identify, for every second of the operating time, the number of customers who were simultaneously receiving service. This approximation is used as a proxy for the number of servers working in each second. To infer the number of servers in each non-overlapping interval, we compute the time average of the number of servers using above approximations. Our discussions with the bank managers suggest that there is higher traffic intensity on a day if that day is the first day or the last day of the month (coinciding with salary and utility payment days), or if that day is right before or right after a major holiday (religious or public). We model the effect of higher intensity on such days on abandonment counts using a "holiday/salary day" covariate. Thus, our covariate vector $\mathbf{Z}_{i,j}$ in (9) consists of four dummy variables for the day of the week effect, one variable for the number of servers and another dummy variable for the holiday/salary day effect.

We assume independent normal priors with mean 0 and variance 100 for all elements of $\boldsymbol{\beta}$ in (12). For γ, we use a gamma prior distribution with mean 1 and variance 100. As such, we use flat but proper priors for all unknown parameters. In implementation of the Gibbs samples, the results are obtained by using 10,000 iterations after an initial burn-in sample of 10,000 iterations. For the parametric case, we sampled every 20th iteration for posterior analysis. We did not experience any convergence problems in any runs. Finally, we also tested runs with different initial starting points and obtained very similar results.

A similar approach was used for the semi-parametric model. The parameters of the baseline intensity function α and γ are set to the posterior mean values obtained from the parametric model as $e^{-1.987}$ and 1.294, respectively. In specifying the precision factor c, we use the value minimizing the mean absolute deviation between the actual and predicted interval counts as discussed in §4.2. Figures 20 and 21 in Appendix A present trace and auto correlation plots for components of $\boldsymbol{\beta}$ obtained by the semi-parametric model using December data at A_1 branch with $c = 1$. Both figures indicate that the Gibbs sampler with Metropolis-Hastings step result in convergence for $\boldsymbol{\beta}$ values and behave properly. We did not face any convergence issues in any combination of runs and the distributions for other combinations are available upon request.

In the Bayesian approach, all inferences for unknown parameter distributions are described by posterior probability distributions where statistics such as mean, variance and mode of the distribution can be obtained. Table 6 below displays the posterior distribution statistics of the covariate parameters using the parametric and semi-parametric (with $c = 1$) models for December at A_1 branch, where $\beta_{\mathrm{TUE}}, \ldots, \beta_{\mathrm{FRI}}$ are the coefficients for the day of the week effects and β_{SER} and β_{HOL} are the coefficients for server and holiday/salary day effects. Table 6 reveals that the covariate parameters have tighter bounds in the semi-parametric model. We perform a similar analysis for all months in both branches, and the details are omitted for brevity.

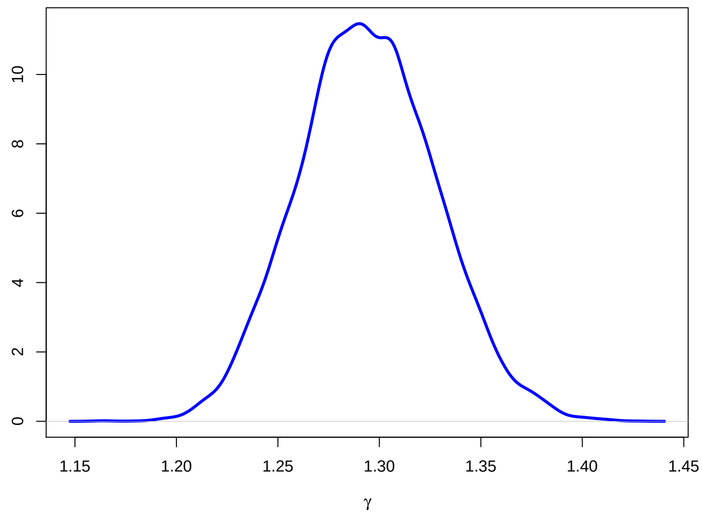
Figure 8 displays that the posterior distribution of γ is concentrated in the interval (1.18, 1.42) and indicates that the expected number of abandonments during any time interval at A_1 branch in December increases with time. Similar results are obtained for all months in the two branches.

Table 6 Statistics for Posterior Distributions of Covariate Parameters at A_1 Branch for December

(a) Parametric Model | (b) Semi-Parametric Model

Node	Mean	Std.Dev.	2.5%	97.5%	Median	Node	Mean	Std.Dev.	2.5%	97.5%	Median
β_{TUE}	-0.490	0.104	-0.694	-0.281	-0.490	β_{TUE}	-0.478	0.068	-0.611	-0.345	-0.478
β_{WED}	-0.709	0.093	-0.893	-0.528	-0.708	β_{WED}	-0.668	0.061	-0.787	-0.549	-0.668
β_{THU}	-0.587	0.108	-0.803	-0.376	-0.586	β_{THU}	-0.558	0.077	-0.711	-0.408	-0.557
β_{FRI}	-0.422	0.084	-0.585	-0.256	-0.422	β_{FRI}	-0.343	0.052	-0.446	-0.241	-0.344
β_{SER}	-0.368	0.032	-0.430	-0.306	-0.369	β_{SER}	-0.372	0.020	-0.411	-0.332	-0.372
β_{HOL}	0.350	0.070	0.212	0.489	0.349	β_{HOL}	0.398	0.045	0.307	0.486	0.399

Figure 8 Posterior Distribution of γ at A_1 branch for December



Figures 9 and 10 display the posterior distributions of components of $\boldsymbol{\beta}$ in the semi-parametric model for December at A_1 branch. Figure 9 indicates that the number of abandonments on Tuesday through Friday are less than that on Monday with probability ≈ 1. Figure 10 suggests that each additional server observed by customers decreases the abandonment counts, as expected. Finally, on a day right before or right after a holiday or on a salary/utility payment day, the number of abandonments increases compared to a regular day (see Figure 10).

Figure 9 Posterior Distributions of the "Day of the Week" effect at A_1 branch for December

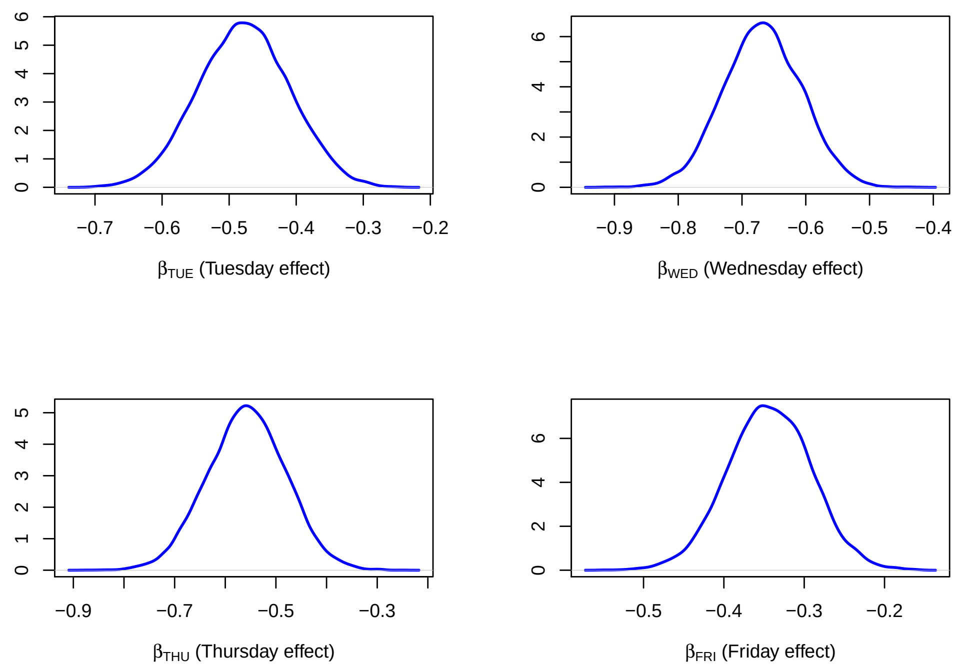
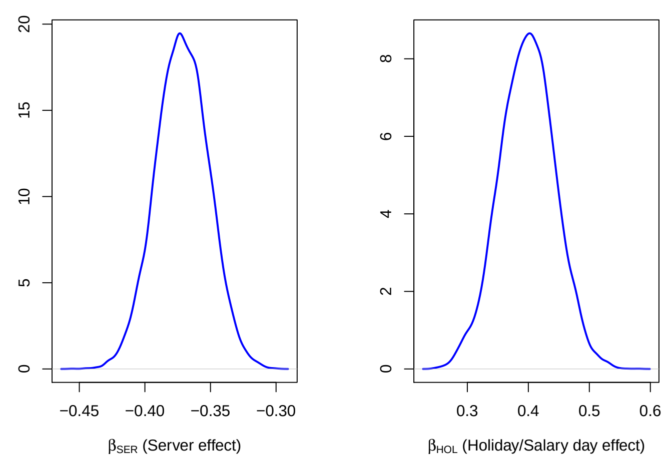
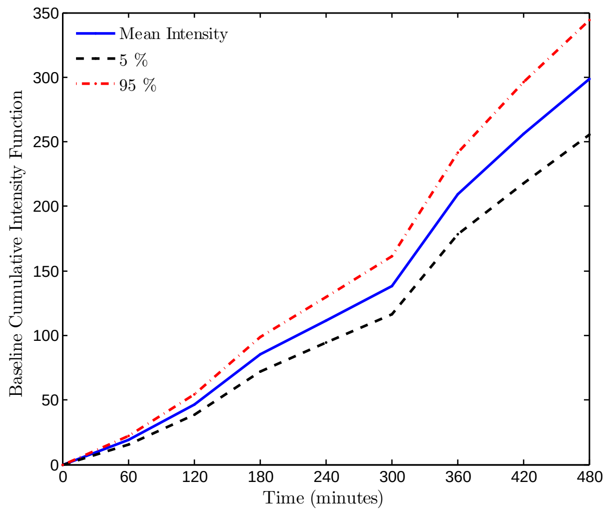


Figure 10 Posterior Distributions of the "Server" and "Holiday/Salary day" effect at A_1 branch for December



It is also possible to obtain the posterior baseline cumulative intensity function for the semi-parametric model as illustrated in Figure 11 by using $c = 1$ for 60 minute intervals. This corresponds to the cumulative intensity function of Mondays which is used as the baseline case in our model. Because 50% values for the baseline cumulative intensity function coincide with the mean intensity values, we omitted them in Figure 11.

Figure 11 **Baseline Cumulative Intensity Function of the Semi-Parametric Model ($c = 1$) for December at A$_1$**



4.2. Comparison of Accuracy for the Parametric and Semi-Parametric Models

We compare the actual and predicted number of abandonments using the parametric and semi-parametric models. In each model, we consider predictions based on the posterior mean and posterior mode for the number of counts in each interval.

To compare the accuracy of models, we first use mean absolute deviation (MAD). Let $\widehat{n}_{i,j}$ denote the predicted number of abandonments during the ith day's jth interval, and $e_{i,j} = n_{i,j} - \widehat{n}_{i,j}$ denote the prediction error made for the ith day's jth interval. We compute MAD for each month and each branch by

$$MAD = \frac{\sum_i \sum_{j=1}^{r_i} |e_{i,j}|}{\sum_i r_i}, \tag{27}$$

where r_i is the number of intervals for day i as previously defined. MADs for each month and branch are shown in Table 7 to compare the prediction accuracy of the parametric and semi-parametric models based on the posterior mean and mode.

We also compare the accuracy of the predictions by comparing the actual and predicted number of abandonments on each day. We use a weighted approach to compute the absolute percent deviation by using the abandonment counts on each day as weights. As such, absolute percent deviation (APD) for a data set is computed by

$$APD = \frac{\sum_i |A_i - F_i| \times A_i}{\sum_i A_i} \times 100\%, \tag{28}$$

where $A_i = \sum_{j=1}^{r_i} n_{i,j}$ and $F_i = \sum_{j=1}^{r_i} \widehat{n}_{i,j}$ are the actual and predicted abandonment counts for day i. The APD comparison is presented in Table 8 for both models.

Tables 7 and 8 indicate that the semi-parametric model outperforms the parametric model in all cases. In the semi-parametric model, the precision factor c that minimizes MAD is not constant

Table 7 Prediction Accuracy based on MAD

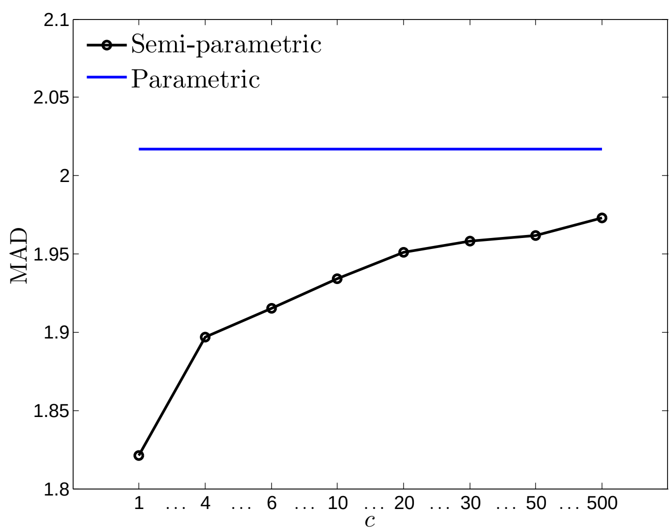
		Parametric Model		Semi-parametric Model	
Branch	Month	Mean	Mode	Mean	Mode
A_1	October	2.28	1.91	1.97	1.72
	November	1.99	1.92	1.89	1.81
	December	2.02	1.92	1.82	1.75
B_1	October	2.82	2.75	2.46	2.39
	November	1.88	1.81	1.65	1.58
	December	2.53	2.42	2.04	1.97

Table 8 Prediction Accuracy based on APD

		Parametric Model		Semi-parametric Model	
Branch	Month	Mean	Mode	Mean	Mode
A_1	October	19.70%	17.07%	17.46%	14.57%
	November	30.52%	28.58%	24.53%	25.98%
	December	13.18%	15.63%	12.94%	14.66%
B_1	October	27.83%	32.14%	24.13%	24.16%
	November	32.77%	43.86%	27.51%	33.47%
	December	17.73%	20.85%	16.00%	16.47%

across months or branches. As expected, when c increases, the prediction accuracy of the semi-parametric model gets closer to that of the parametric model, as illustrated in Figure 12.

Figure 12 MAD vs c for A_1 branch in December (Posterior Mean Case)



In Table 9, for both models, we report the percentage of $n_{i,j}$ (i.e., actual interval counts) that are covered by 95% posterior probability intervals (PI) for each month at each branch. We also present coverage percentages based on one and two posterior standard deviations (σ) around the posterior predictive means. Table 10 illustrates the results for the whole data set containing 3,300 intervals. The semi-parametric model predicts the actual counts with higher accuracy than the parametric model does. For 88.3% of all intervals, the semi-parametric model' s PI cover the actual count data (see Table 10). For illustrative purposes, Figure 13 presents one and two σ limits around the cumulative posterior mean of the abandonment counts for Tuesday, December 30, at A_1 branch. For this day, all the actual cumulative abandonment counts are within the 2σ limits.

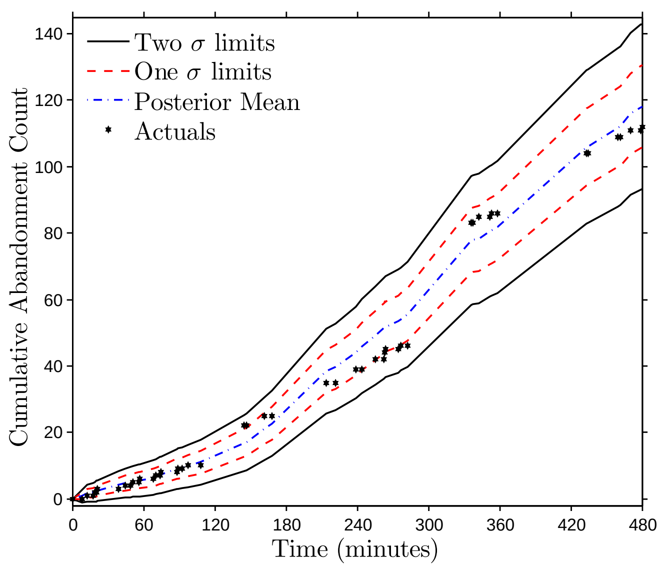
Table 9 Percentage of Actual Abandonments Within Various Prediction Limits

		Parametric Model			Semiparametric Model		
Branch	Month	$\pm\sigma$ (%)	$\pm2\sigma$ (%)	95% PI (%)	$\pm\sigma$ (%)	$\pm2\sigma$ (%)	95% PI (%)
A_1	October	45.6	78.1	88.1	50.4	82.6	91.0
	November	44.1	73.6	86.0	45.3	75.9	87.5
	December	47.7	78.9	88.9	47.8	78.9	88.6
B_1	October	43.1	68.6	78.5	42.2	71.4	83.9
	November	45.6	69.7	85.2	45.2	71.1	87.7
	December	49.3	81.4	86.9	51.7	82.7	89.2

Table 10 Percentage of Actual Abandonments Within Various Prediction Limits - Whole Data Set

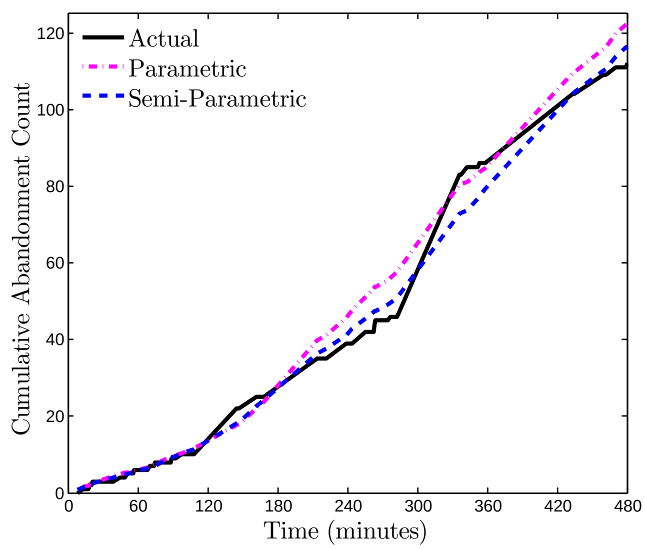
Model	$\pm\sigma$ (%)	$\pm2\sigma$ (%)	95% PI (%)
Parametric	45.8	75.2	86.1
SemiParametric	47.2	77.4	88.3

Figure 13 **σ Limits Around the Cumulative Posterior Mean vs. Actual Abandonment Counts for Dec 30 at A_1**



Thus far, we have completed the analysis in this section retrospectively by using all the available data for all days in a given month at a given branch. We can also compare the two modeling approaches by using "Out of sample" predictions, where we exclude the data for a specific day's abandonment counts. We note that, as discussed in §3.2, in the semi-parametric model, this requires a new MCMC run with data augmentation where we incorporate the specific intervals for which predictions are required. Thus, for illustrative purposes, we present out of sample predictions using posterior mean counts for a specific day – Tuesday, December 30 at A_1 branch – in Figure 14. For this day, absolute deviation per interval is 1.45 and 1.54 for the semi-parametric and parametric models, respectively. Similarly, absolute percent deviation for total abandonment counts is 4.03% and 9.16% for the semi-parametric and parametric models, respectively. As such, semi-parametric model also provides better out of sample predictions than the parametric model.

Figure 14 Actual vs. Out of Sample Predicted Abandonments for Tuesday, December 30, at A_1



We also provide results for out of sample predictions on a salary payment day at A_1 branch in Appendix B.1 by considering posterior mean counts for predictions. Similar analysis can be performed for B_1 branch and we present two such examples in Appendices B.2 and B.3.

Finally, Figures 15 and 16 illustrate the posterior predictive distributions of the abandonment counts during the morning and afternoon hours, respectively, for Tuesday, December 30, with the actual number of servers used on that day. We observe that the predicted abandonment counts are higher in the afternoon hours compared to the morning hours.

Figure 15 **Abandonment Count Distribution for Tuesday, Dec 30, Morning with Actual Server Allocation**

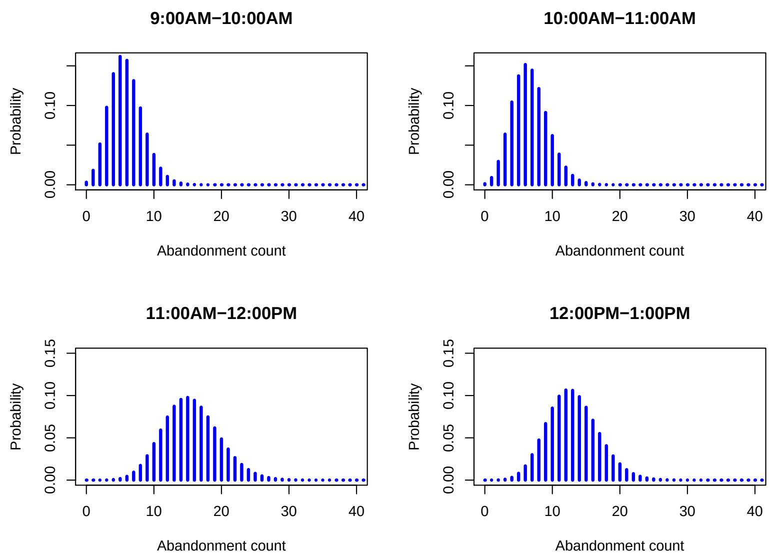
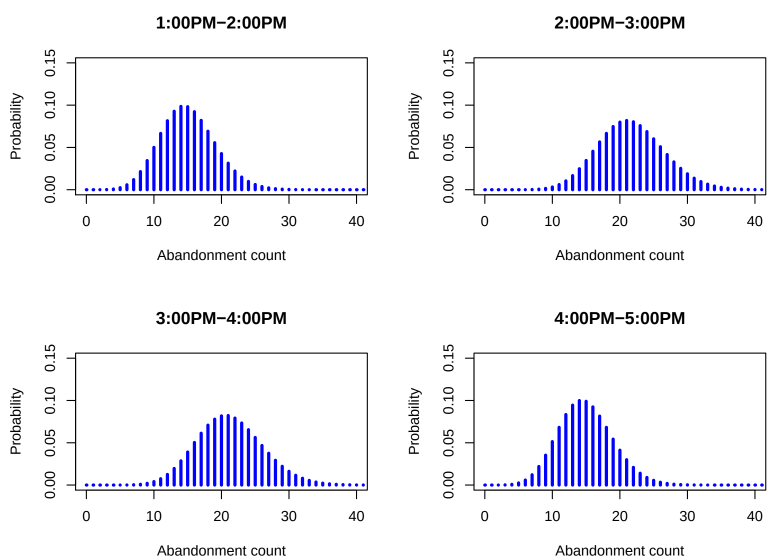


Figure 16 **Abandonment Count Distribution for Tuesday, Dec 30, Afternoon with Actual Server Allocation**



5. Insights on Server Allocation Policies from Bayesian Analysis

The proposed models can provide valuable managerial insights on the staffing decisions. Because each day has its own characteristics, we could determine the optimal staff allocation policy for each day, independent from the decisions on the other days. Next, we discuss how we can use the MPPMs to develop the server allocation policy that minimizes the total costs of labor and abandonment.

Let h be the staffing cost per hour and a be the (opportunity) cost of an abandonment to the firm. We assume that the firm can modify the number of servers employed throughout the day by using "*1-hour*, *2-hour*, and *4-hour*" policies. Let $s_{i,j}$ denote the number of servers used during the jth hour of day i. We set $s_{i,j} \in S = \{2,3,4,5\}$. Also, let S^p denote the set of all possible server allocation combinations under the policy that requires staffing decisions every p hours. For example, under the "1-hour" policy, we identify the optimal allocation policy by evaluating $4^8 = 65,536$ possible server allocations.

With these assumptions, we compute the expected total cost of staffing and abandonment for day i given server allocation $s_{i,j}$ by

$$\mathbb{E}[C_i] = \sum_j h * s_{i,j} + a * \mathbb{E}[N_i] \tag{29}$$

where $\mathbb{E}[N_i]$ is the expected value of the predicted number of abandonments on day i, obtained by the MCMC methods previously discussed. Our goal is to find the optimal hourly server allocation that minimizes Eq. (29), and we achieve this goal by exhaustively evaluating S^p, the set of all possible server combinations.

We set $h = \$5$ based on the reported average salary of bank personnel in the annual report of "The Banks Association of Turkey" for the year our data is obtained. We also set $a \in \{0.25, 0.5, 0.75, 1, 1.25, 1.5, 2, 2.5\}$. For a given (h,a) pair, we identify the optimal server allocation that would minimize the expected sum of staffing and predicted abandonment costs on a given day. Once we identify the optimal server allocation, we compare its expected total cost with the actual costs of that day under the current policy used by the bank in question. For actual cost computations, we consider the time-averaged number of servers in each non-overlapping interval and the actual abandonment counts for that day.

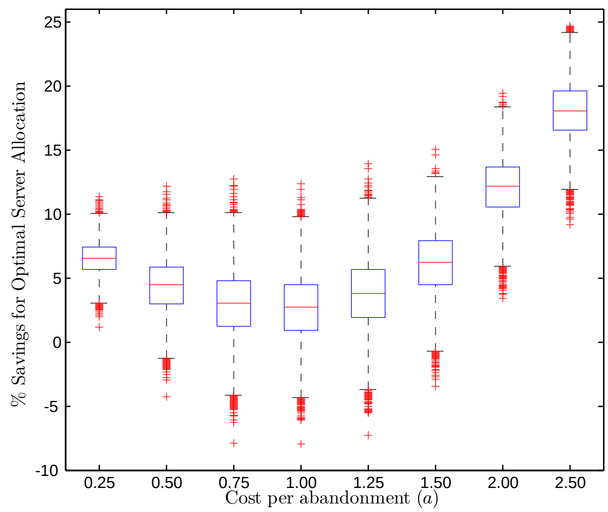
In Table 11, using the semi-parametric model, we show the optimal server allocations for Tuesday, December 30, at A_1 branch with the 1-hour policy. We evaluate the server state space exhaustively to identify the optimal staff allocation. Note that the number of servers required increases when the abandonment cost per unit (a) becomes larger than 0.75. In addition, for $a \geq 1$, the number of servers required for each hour increases, in general, during a day.

Using the posterior samples obtained for a given a and the actual operations costs, we obtain the percentage savings by using the optimal allocation of servers under 1-hour policy as shown in Figure 17. This can provide the managers with a better understanding of how their costs can be improved by varying the server levels (for a given level of abandonment cost (a)), and accordingly allow them to better assess the risks of under or overstaffing. Note that the form of the percent savings plot (e.g., Figure 17) is different for each day and for each branch. Such analysis could provide tailored server allocation strategies for the bank.

Table 11 Optimal Server Allocations and Corresponding Daily Costs for Tuesday, December 30, at A_1 ($h = 5$)

Abandonment Cost (a)	Optimal Server Allocation (hourly)	Predicted # Abandonments	Optimal Daily Cost ($) Expected Value	Optimal Daily Cost ($) Standard Deviation	Actual Daily Cost ($)
0.25	(2,2,2,2,2,2,2,2)	117.01	109.25	1.46	116.87
0.50	(2,2,2,2,2,2,2,2)	117.01	138.50	2.92	144.87
0.75	(2,2,2,2,2,2,2,2)	117.01	167.76	4.38	172.87
1.00	(2,2,2,2,2,3,3,3)	100.54	195.54	5.04	200.87
1.25	(2,2,2,2,2,4,3,3)	96.26	220.32	5.68	228.87
1.50	(2,2,3,3,3,4,4,4)	77.38	241.07	6.52	256.87
2.00	(2,3,4,4,4,5,5,5)	57.54	275.09	8.24	312.87
2.50	(3,4,5,4,5,5,5,5)	49.00	302.49	10.22	368.87

Figure 17 Percent Savings by Optimal Allocation under 1-hr policy for Tuesday, December 30, at A_1 branch



We also compare the actual costs with the optimal costs by the proposed staff allocation to identify potential savings under the three server allocation policies for different levels of abandonment cost a for Tuesday, December 30. For $a \in \{0.25, 0.50, 0.75\}$, percent savings are identical under all three policies. For $a > 0.75$, 1-hr policy provides the highest savings for the branch. However, the differences between the savings from the three policies are not that significant. As such, if service providers find it impractical to modify the number of servers hourly, the 2-hr and 4-hr policies could be implemented easily for a typical Tuesday at A_1 branch. Note that, policy savings comparisons for each day at each branch should be performed individually and will show variations.

Figures 18 and 19 present the plots of the posterior predictive distributions of the abandonment counts during the morning and afternoon hours, respectively, for Tuesday, December 30, at A_1 branch when we use the optimal server allocation under 1-hour policy with $a = 1.25$. For this specific choice of abandonment cost a, we see that the proposed allocation would reduce the abandonment counts mainly in the afternoon hours (see 1:00PM-4:00PM). Similar analysis can be developed for other values of a. Finally, Appendix C provides a similar costs and savings analysis for Monday, December 1, which is a salary payment day.

Figure 18 **Abandonment Count Distribution for Tuesday, December 30, Morning with optimal server allocation and** $a = 1.25$

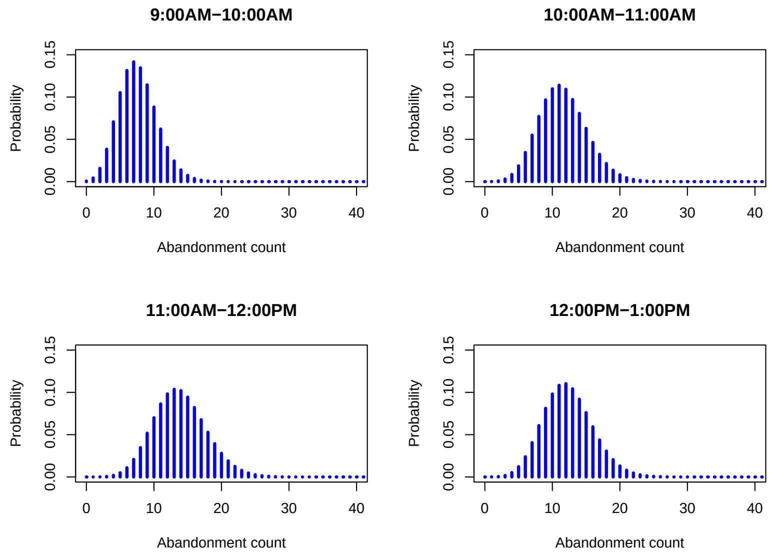
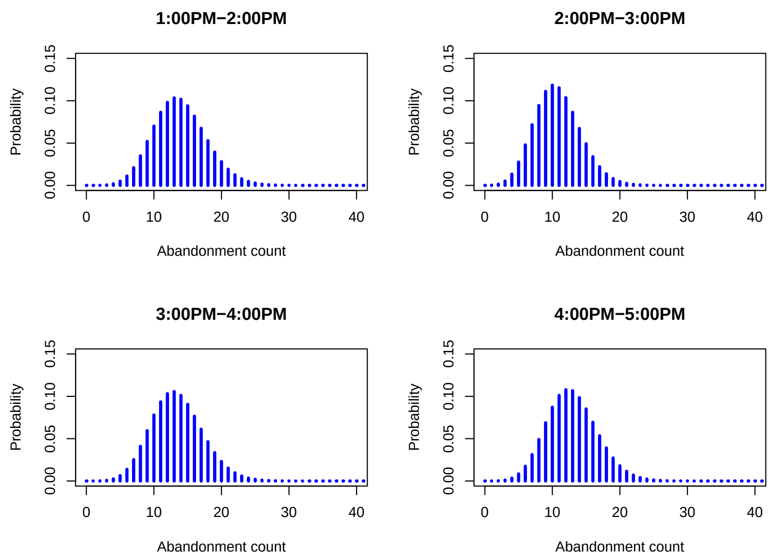


Figure 19 **Abandonment Count Distribution for Tuesday, December 30, Afternoon with optimal server allocation and** $a = 1.25$



6. Concluding Remarks

This article introduces a modulated Poisson process model to describe the abandonment process in ticket queues and to analyze abandonment data which is interval censored by its nature. In addition to providing the first statistical modeling and analysis of real abandonment data in TQs, its emphasis on Bayesian methods distinguishes this work from other modeling attempts in TQs. We introduce parametric and semi-parametric models, develop their Bayesian analysis and show how the models can be used to predict future abandonments. The semi-parametric Bayesian modeling of abandonments proposed here presents a modest contribution to Bayesian queueing literature as well.

The proposed models are implemented on a Turkish bank' s ticket queue data to show that the semi-parametric model provides more accurate abandonment count predictions than the parametric model does. The Bayesian framework is also used to develop managerial insights for server allocation. We illustrate how optimal server allocation policies can be developed, by considering staffing and abandonment costs, for each day of the week at each branch. We also discuss potential cost savings as a result of using the proposed framework.

An area of future research is the analysis and comparison of abandonments across multiple customer classes in TQs. This would provide insights about the effects of customer heterogeneity on abandonment counts. We could also introduce additional covariates to our MPPM model such as the average waiting times for those customers who abandon in each interval. However, interval censored nature of TQ data would require us to use a proxy, such as "offered waiting time" defined in Kuzu et al. (2017), for actual waiting times before abandonment. Finally, the proposed semi-parametric Bayesian modeling framework could be implemented on interval censored data from other domains, such as call center operations as well as reliability and survival analysis.

References

Aksin, Z., B. Ata, S. M. Emadi, C. L. Su. 2013. Structural estimation of callers' delay sensitivity in call centers. *Management Science* **59**(12) 2727–2746.

Aksin, Z., B. Ata, S. M. Emadi, C.L. Su. 2016. Impact of delay announcements in call centers: An empirical approach. *Operations Research* **65**(1) 242–265.

Aktekin, A., R. Soyer. 2012. Bayesian analysis of queues with impatient customers: Applications to call centers. *Naval Research Logistics* **59**(6) 441–456.

Aktekin, T. 2014. Call center service process analysis: Bayesian parametric and semi-parametric mixture modeling. *European Journal of Operational Research* **234**(3) 709–719.

Armero, C. 1994. Bayesian inference in markovian queues. *Queueing Systems* **15**(1) 419–426.

Batt, R.J., C. Terwiesch. 2015. Waiting Patiently: An Empirical Study of Queue Abandonment in an Emergency Department. *Management Sci.* **61** 39–59.

Bolandifar, E., N. DeHoratius, T.L. Olsen, J.L. Wiler. 2016. Modeling the behavior of patients who leave the ed without being seen. Working Paper.

Chib, S., E. Greenberg. 1995. Understanding the metropolis-hasting algorithm. *Amer. Statis.* **49** 327–335.

Cox, D. R. 1972. *Stochastic Point Processes*, chap. The Statistical Analysis of Dependencies in Point Processes. Wiley, New York, 55–66.

Gilks, W. R., P. Wild. 1992. Adaptive rejection sampling for gibbs sampling. *Appl. Statist.* **41** 337–348.

Ibrahim, R., M. Armony, A. Bassamboo. 2017. Does the Past Predict the Future? The Case of Delay Announcements in Service Systems. *Management Science* **63**(6) 1762–1780.

Kuo, L., S. K. Ghosh. 2001. Bayesian nonparametric inference for nonhomogeneous poisson processes. Tech. rep., North Carolina State University. Institute of Statistics Mimeo Series No. 2530.

Kuzu, K. 2010. Analytical and Empirical Investigations of Ticket Queues: Implications for System Performance, Customer Perceptions and Behavior. Ph.D. Thesis, The Pennsylvania State University.

Kuzu, K. 2015. Comparisons of Perceptions and Behavior in Ticket Queues and Physical Queues. *Service Science* **7**(4) 294–314.

Kuzu, K., L. Gao, S. H. Xu. 2017. To wait or not to wait: The theory and practice of ticket queues. Working Paper.

Mandelbaum, A., S. Zeltyn. 2004. The impact of customer' patience on delay and abandonment: Some empirically-driven experiments with the $M/M/n+G$ queue. *OR Spectrum* **26**(3) 377–411.

Mandelbaum, A., S. Zeltyn. 2009. Staffing Many-Server Queues with Impatient Customers: Constraint Satisfaction in Call Centers. *Oper. Res.* **57**(5) 1189–1205.

Merrick, J. R., R. Soyer. 2017. Semiparametric bayesian optimal replacement policies: Application to railroad tracks. *Appl. Stoc. Mod. in Bus. and Ind.* Forthcoming.

Merrick, J. R., R. Soyer, T.A. Mazzuchi. 2005. Are maintenance practices for rail road tracks effective? *Journal of the American Statistical Association* **100** 17–25.

Pender, J., O. Jennings. 2016. Comparisons of standard and ticket queues. *Queueing Systems* **84** 145–202.

Pievatolo, A., F. Ruggeri. 2004. Bayesian reliability analysis of complex reparable systems. *Appl. Stoc. Mod. in Bus. and Ind.* **20** 253–264.

Ramirez-Cobo, P., R. E. Lillo, S. Wilson, M. P. Wiper. 2010. Bayesian Inference for Double Pareto Lognormal Queues. *The Annals of Applied Statistics* **4**(3) 1533–1557.

Ross, S. M. 1989. *Introduction to Probability Models*. 4th ed. Academic Press: New York, NY, USA.

Singpurwalla, N. D. 1997. *In Engineering Probabilistic Design and Maintenance for Flood Protection.*, chap. Gamma Processes and Their Generalizations: An Overview. Kluwer, New York, 67–75.

Sinha, D. 1993. Semiparametric bayesian analysis of multiple time data. *Journal of the American Statistical Association* **88** 979–983.

Soyer, R., M. M. Tarimcilar. 2008. Modeling and Analysis of Call Center Arrival Data: A Bayesian Approach. *Management Science* **54** 266–278.

Tanner, T. A., W. H. Wong. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* **82** 528–549.

van Noortwick, J. M. 2009. A survey of the application of gamma processes in maintenance. *Reliability Engineering and System Safety* **94** 2–21.

Ventz, S., P. Muliere. 2014. Fast and Scalable Nonparametric Bayesian Prediction for the $M/G/1$ Queue. Tech. Report.

Xiao, L., S. Xu, D.D. Yao, H. Zhang. 2015. Optimal staffing for ticket queues. Working paper.

Xu, S.X., L. Gao, J. Ou. 2007. Service Performance Analysis and Improvement for a Ticket Queue with Balking Customers. *Management Sci.* **53** 971–990.

Yu, Qiuping, Gad Allon, Achal Bassamboo. 2017. How do delay announcements shape customer behavior? an empirical study. *Management Science* **63**(1) 1–20.

Appendix A: Sample Trace and Autocorrelation plots for the Semi-parametric Model

Figure 20 **β samples in Gibbs iterations - A$_1$ branch for December**

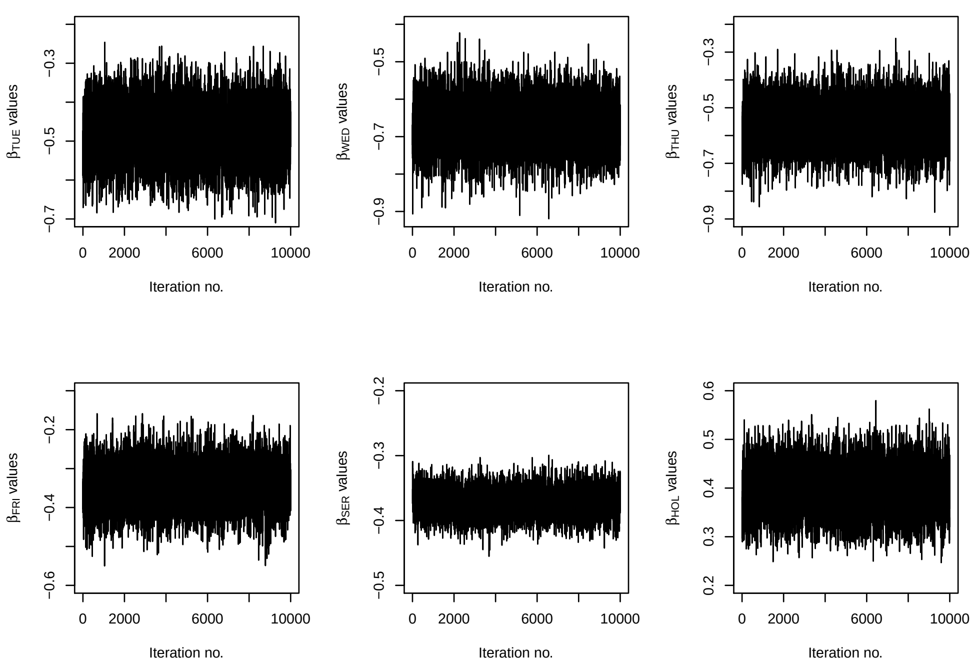
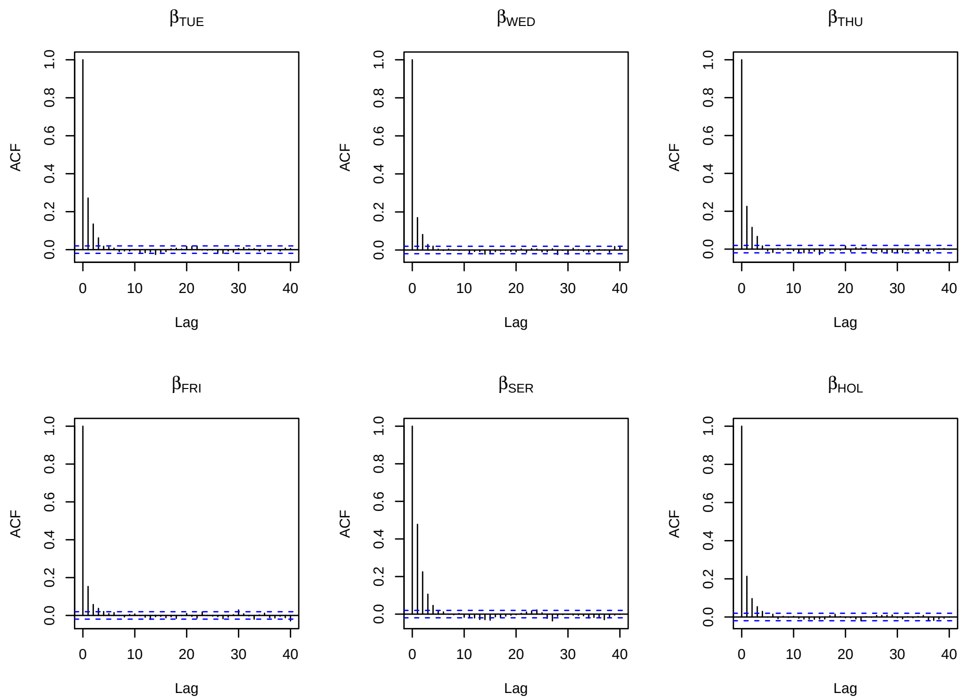


Figure 21 **Autocorrelation values for β samples in Gibbs iterations - A_1 branch for December**

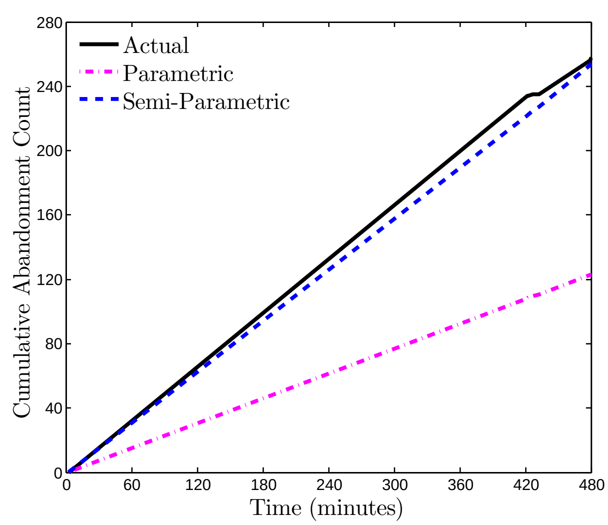


Appendix B: Out of Sample Predictions for additional days

B.1. Predictions for Monday, December 1, a Salary Payment Day at A_1

Figure 22 shows the actual and out of sample predicted (posterior mean) abandonment counts on a salary payment day, Monday, December 1, at A_1. The semi-parametric model provides significantly more accurate estimations even if we exclude a salary payment day' s count data for out of sample predictions. The parametric model' s performance suffers significantly when the salary payment day' s count data is omitted.

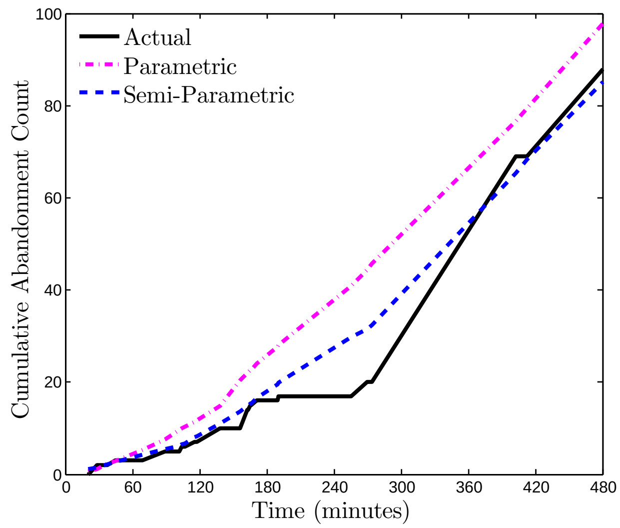
Figure 22 **Actual vs. Out of Sample Predicted Abandonments for Monday, December 1, at A_1**



B.2. Predictions for Wednesday, December 24, at B_1

Figure 23 shows the actual and out of sample predicted (posterior mean) abandonment counts on Wednesday, December 24, at Branch B_1 (which is not a holiday or salary payment day). The semi-parametric model (with $c = 4$) outperforms the parametric model.

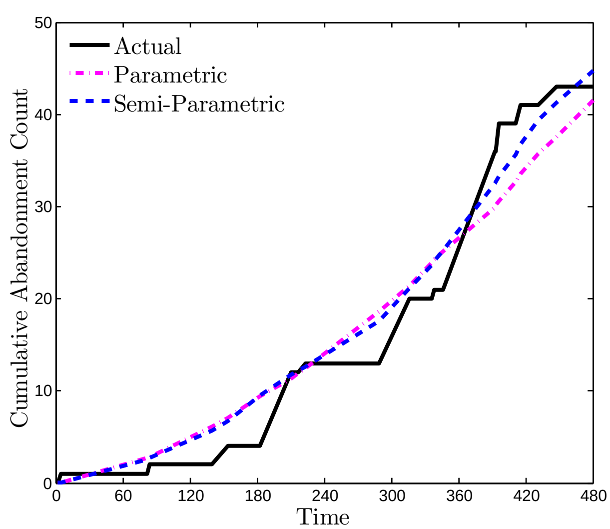
Figure 23 Actual vs. Out of Sample Predicted Abandonments for Wednesday, Dec 24, at B_1



B.3. Predictions for Thursday, December 18, at B_1

Figure 24 shows the actual and out of sample predicted (posterior mean) abandonment counts on Thursday, December 18, at B_1 branch (which is not a holiday or salary payment day). The cumulative abandonment counts are predicted with a higher accuracy by the semi-parametric model (with $c = 4$) than by the parametric model.

Figure 24 Actual vs. Out of Sample Predicted Abandonments for Thursday, Dec 18, at B_1



Appendix C: Cost Savings Analysis for Monday, December 1, a Salary payment day, at A_1

Table 12 outlines the optimal server allocation at each abandonment cost (a) on Monday, December 1, which is a holiday, at A_1 branch using the semi-parametric model predictions and the 1-hour policy. Figure 25 displays the percentage savings by the optimal server allocation under the 1-hour policy.

Table 12 Summary of Server Allocation and Cost Statistics - Monday, December 1, at A_1

Abandonment Cost	Optimal Server Allocation (hourly)	Predicted # Abandonments	Optimal Daily Cost Expected Value	Optimal Daily Cost Standard Deviation	Actual Daily Cost
$a = 0.25$	(2,2,2,2,2,2,2,2)	266.96	146.74	3.66	154.11
$a = 0.50$	(2,2,2,2,2,2,2,2)	266.96	213.48	7.33	218.61
$a = 0.75$	(2,2,2,2,2,4,3,3)	235.02	276.27	11.15	283.11
$a = 1.00$	(2,2,3,2,3,5,4,4)	205.16	330.16	15.16	347.61
$a = 1.25$	(2,3,4,3,4,5,5,5)	176.77	375.97	19.52	412.11
$a = 1.50$	(2,4,5,4,5,5,5,5)	161.50	417.25	23.66	476.61
$a = 2.00$	(3,5,5,5,5,5,5,5)	152.31	494.81	31.94	605.61
$a = 2.50$	(4,5,5,5,5,5,5,5)	150.01	570.03	40.14	734.61

Figure 25 % Savings by Optimal Allocation under 1-hr policy on Monday, December 1, at A_1