Incorporating Time-Varying Covariates in the BG/NBD Model

Peter S. Fader
www.petefader.com

Bruce G. S. Hardie
www.brucehardie.com

Chun-Yao Huang
National Taiwan University†

July 2020

1 Introduction

We present an extension to the BG/NBD model (Fader et al. 2005) in which the timing of transactions is influenced by time-varying covariates. This note builds on work presented in Huang (2002), and the presentation of the derivations assumes that the reader is familiar with the BG/NBD model derivations (Fader et al. 2019).

Preliminaries

- We observe a (new) customer making $x + 1$ transactions with the firm at times $t_0, t_1, \ldots, t_x$. By convention, $j = 0$ corresponds to the customer's first-ever transaction with the firm; $j > 0$ corresponds to repeat transactions by the customer. The t_j are measured in calendar time with some arbitrary origin, where $t_0 \geq 0$.

- We are interested in modeling the customer's repeat buying in the time interval $(t_0, T]$, where T is the censoring point that corresponds to the end of the model calibration period. This is typically a fixed point in calendar time, common across all customers.

- We assume that the unit of time (e.g., week, month) is chosen such that it can be assumed that the value of each covariate remains constant within that unit of time (but will vary over time). Transactions can occur in fractions of this time period (e.g., day when the unit of time is week).

- We assume that the calendar time origin is aligned with the start of the unit of time. For example, if week is the unit of time and we are working on a calendar where the first day of the week is Sunday, 0 corresponds to midnight Saturday/Sunday.

When individual-level inter-transaction times are iid exponential with rate parameter λ, it is easy to incorporate the effects of time-varying covariates using the standard proportional hazards framework. We summarise the basic results presented in Fader et al. (2004), making minor notational changes.

Let $\mathbf{z}(t)$ denote the vector of covariates at time t and $\boldsymbol{\beta}$ the effects of these covariates. According to the proportional hazards framework, the with-covariates hazard function is

$$\begin{aligned} h(t \mid \lambda, \boldsymbol{\beta}; \mathbf{z}(t)) &= \lambda e^{\mathbf{z}(t)\boldsymbol{\beta}'} \\ &\equiv \lambda A(t)\,. \end{aligned}$$

It is **very** important to note that this vector of covariates does not include an intercept.

The probability that a transaction has not occurred by t, given the last transaction occurred at t_j, is given by the survivor function

$$\begin{aligned} S(t \mid t_j; \lambda, \boldsymbol{\beta}; \mathbf{Z}(t)) &= \exp\Big[-\int_{t_j}^{t} h(u \mid \lambda, \boldsymbol{\beta}; \mathbf{z}(u))\,du\Big] \\ &= \exp\Big[-\lambda\Big(\int_0^t A(u)\,du - \int_0^{t_j} A(u)\,du\Big)\Big]\,, \end{aligned}$$

where $\mathbf{Z}(t)$ represents the covariate path up to time t.

Given our previously stated assumption that the time-varying covariates remain constant within each unit of time,[1]

$$\begin{aligned} \int_0^t A(u)\,du &= \sum_{i=1}^{\lfloor t \rfloor} A(i) + (t - \lfloor t \rfloor) A(\lceil t \rceil) \\ &\equiv C(t)\,, \end{aligned}$$

[1] That is, for t falling in the ith time period (i.e., $i-1 < t \le i$, $i = 1, 2, 3, \ldots$), $\mathbf{z}(t) = \mathbf{z}(\lceil t \rceil) = \mathbf{z}(i)$.

where $\lfloor\,\rfloor$ is the floor function and $\lceil\,\rceil$ is the ceiling function. Letting $C(t_j, t) = C(t) - C(t_j)$, we can write the survivor function of the with-covariates inter-transaction time distribution as

$$S(t|t_j; \lambda, \boldsymbol{\beta}; \mathbf{Z}(t)) = e^{-\lambda C(t_j, t)} . \tag{1}$$

Noting that $\partial C(t_j, t)/\partial t = A(\lceil t \rceil)$, it follows that the pdf of the with-covariates inter-transaction time distribution is

$$f(t|t_j; \lambda, \boldsymbol{\beta}; \mathbf{Z}(t)) = \lambda A(\lceil t \rceil) e^{-\lambda C(t_j, t)} . \tag{2}$$

This is equivalent to assuming that transactions can be characterized by a non-homogeneous Poisson process (NHPP). Let the random variable $X(t_a, t_b)$ denote the number of transactions occurring in the time interval $(t_a, t_b]$. It follows that

$$P(X(t_a, t_b) = x \mid \lambda, \boldsymbol{\beta}; \mathbf{Z}(t_b)) = \frac{[\lambda C(t_a, t_b)]^x e^{-\lambda C(t_a, t_b)}}{x!} \tag{3}$$

and

$$E(X(t_a, t_b) \mid \lambda, \boldsymbol{\beta}; \mathbf{Z}(t_b)) = \lambda C(t_a, t_b) . \tag{4}$$

2 Model Likelihood Function

Recall that the BG/NBD model is based on the following six assumptions:

1. Customers go through two stages in their "lifetime" with a specific firm: they are "alive" for some period of time, then become permanently inactive (i.e., "die").

2. While alive, the number of transactions made by a customer follows a Poisson process with transaction rate λ. This is equivalent to assuming that the time between transactions is distributed exponential with transaction rate λ,

$$f(t_j \mid t_{j-1}; \lambda) = \lambda e^{-\lambda(t_j - t_{j-1})} , \quad t_j > t_{j-1} \geq 0 .$$

3. Heterogeneity in transaction rates across customers follows a gamma distribution with shape parameter r and scale parameter α:

$$g(\lambda \mid r, \alpha) = \frac{\alpha^r \lambda^{r-1} e^{-\lambda\alpha}}{\Gamma(r)} .$$

4. After any (repeat) transaction, a customer dies with probability p.[2] Therefore the point at which the customer dies is distributed across transactions according to a geometric distribution with pmf

$$P(\text{die immediately after } j\text{th transaction}) \\ = p(1-p)^{j-1}\,, \quad j = 1, 2, 3, \ldots\,.$$

5. Heterogeneity in death probabilities follows a beta distribution with parameters a and b:

$$g(p \mid a, b) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}\,, \quad 0 < p < 1\,.$$

6. The transaction rate λ and the death probability p vary independently across customers.

We modify the second assumption, now assuming that inter-transaction times are distributed according to the *with-covariates* exponential distribution, (1) and (2). Following the logic of the derivation presented in Fader et al. (2005, 2019), it follows that

$$L(\lambda, p, \boldsymbol{\beta} \mid \mathbf{Z}(T), [t_0, \ldots, t_x]) = \left\{ \prod_{j=1}^{x} A(\lceil t_j \rceil) \right\} \left\{ (1-p)^x \lambda^x e^{-\lambda C(t_0, T)} \right. \\ \left. + \mathbf{1}_{x>0}\, p(1-p)^{x-1} \lambda^x e^{-\lambda C(t_0, t_x)} \right\}. \quad (5)$$

To remove the conditioning on the unobserved transaction rate and death probability, we take the expectation of (5) over the distributions of λ and p:

$$L(r, \alpha, a, b, \boldsymbol{\beta} \mid \mathbf{Z}(T), [t_0, \ldots, t_x]) \\ = \int_0^1 \int_0^\infty L(\lambda, p, \boldsymbol{\beta} \mid \mathbf{Z}(T)[t_0, \ldots, t_x]) g(\lambda \mid r, \alpha) g(p \mid a, b) d\lambda dp \\ = \left\{ \prod_{j=1}^{x} A(\lceil t_j \rceil) \right\} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)} \left\{ \frac{B(a, b+x)}{B(a,b)} \left(\frac{1}{\alpha + C(t_0, T)} \right)^{r+x} \right. \\ \left. + \mathbf{1}_{x>0} \frac{B(a+1, b+x-1)}{B(a,b)} \left(\frac{1}{\alpha + C(t_0, t_x)} \right)^{r+x} \right\}. \quad (6)$$

[2]The standard BG/NBD model does not allow for "death" immediately after the first (trial) transaction. See Batislam et al. (2007, 2008), Hoppe and Wagner (2007), and Wagner and Hoppe (2008) for a variant of the standard model that allows for this. Harman (2016) incorporates the effects of time-varying covariates into the transaction process of this modified BG/NBD model, relying on simulation methods to compute quantities such as expected sales.

Note that we now need to know the customer's entire purchase history, $[t_0, \ldots, t_x]$; "recency" and "frequency"' are no longer sufficient statistics.

When $\boldsymbol{\beta} = \mathbf{0}$, $A(t) = 1$ and $C(t) = t$. Therefore (6) reduces to the basic BG/NBD model likelihood when the effects of the time-varying covariates are "switched off".[3]

3 Expected Sales

We wish to derive an expression for $E[X(t_0, t)]$, where the random variable $X(t_0, t)$ denotes the number of (repeat) transactions made by t for a customer acquired at t_0.

Given that the number of transactions (while alive) follows a NHPP and recalling (4), $E[X(t_0, t)]$ is simply $\lambda C(t_0, t)$ if the customer is alive at t. For a customer who dies at $\omega \leq t$, the expected number of transactions in the time interval $(t_0, \omega]$ is $\lambda C(t_0, \omega)$.

The probability that a customer acquired at t_0 is still alive at ω is

$$\begin{aligned} P(\Omega > \omega \,|\, \lambda, p, \boldsymbol{\beta}; \mathbf{Z}(\omega), t_0) &= \sum_{j=0}^{\infty} (1-p)^j \frac{[\lambda C(t_0, \omega)]^j e^{-\lambda C(t_0, \omega)}}{j!} \\ &= e^{-\lambda p C(t_0, \omega)} \,. \end{aligned}$$

This implies that the pdf of the time at which a customer dies is given by

$$g(\omega \,|\, \lambda, p, \boldsymbol{\beta}; \mathbf{Z}(\omega), t_0) = \lambda p A(\lceil \omega \rceil) e^{-\lambda p C(t_0, \omega)} \,.$$

It follows that the expected number of (repeat) transactions occurring by t for a customer acquired at t_0 is given by

$$\begin{aligned} E(X(t_0, t) \,|\, \lambda, p, \boldsymbol{\beta}; \mathbf{Z}(t)) &= \lambda C(t_0, t) P(\Omega > t \,|\, \lambda, p, \boldsymbol{\beta}; \mathbf{Z}(t), t_0) \\ &\quad + \int_{t_0}^{t} \lambda C(t_0, \omega) g(\omega \,|\, \lambda, p, \boldsymbol{\beta}; \mathbf{Z}(\omega), t_0) d\omega \\ &= \frac{1}{p} - \frac{1}{p} e^{-\lambda p C(t_0, t)} \,. \end{aligned} \tag{7}$$

(See the Appendix for details of the derivation.)

---

[3]Note that we have not included the effects of time-varying covariates in the death process. It is not practical to include the effects of covariates, be they time-invariant or time-varying, into p and then allow for unobserved heterogeneity using the beta distribution. Anyone wishing to let p be a function of time-varying covariates will have to replace the BG components of the basic BG/NBD model. (See, for example, Braun et al. (2015), who use what can be called the G2G+covariates model (Fader and Hardie 2020) to allow the probability of dying immediately after the jth transaction to be a function of the characteristics of that transaction.) See Fader and Hardie (2007) for a discussion of how to incorporate the effects of time-invariant covariates into the death process.

Taking the expectation of (7) over the distributions of λ and p results in the following expression for the expected number of transactions in the time interval $(t_0, t]$:

$$\begin{aligned} &E(X(t_0,t)\,|\,r,\alpha,a,b,\boldsymbol{\beta};\mathbf{Z}(t)) \\ &\quad = \frac{a+b-1}{a-1}\left[1-\left(\frac{\alpha}{\alpha+C(t_0,t)}\right)^r {}_2F_1\big(r,b;a+b-1;\tfrac{C(t_0,t)}{\alpha+C(t_0,t)}\big)\right]. \quad (8)\end{aligned}$$

When $\boldsymbol{\beta} = \mathbf{0}$ (i.e., no covariate effects), $C(t_0,t) = t - t_0$. Therefore (8) reduces to the basic BG/NBD model expression for $E[X(t_0,t)]$. (Note that $t_0 = 0$ in all the expressions presented in the original paper.)

4 Distribution of Repeat Transactions

In order to derive an expression for $P(X(t_0,t) = x)$, we recall the fundamental relationship between inter-event times and the number of events: $X(t_0,t) \geq x \Leftrightarrow T_x\,|\,t_0 \leq t$, where $T_x\,|\,t_0$ is the random variable denoting the time of the xth repeat transaction for a customer acquired at t_0. Given our assumption regarding the nature of the death process,

$$\begin{aligned} P(X(t)=x) = P(&\text{alive after } x\text{th transaction})\,P(T_x \leq t \text{ and } T_{x+1} > t) \\ &+ \mathbf{1}_{x>0}\,P(\text{die after } x\text{th transaction})P(T_x \leq t)\,. \end{aligned}$$

Now, $P(T_x \leq t \text{ and } T_{x+1} > t)$ is simply the NHPP probability that $X(t_0,t) = x$, (3). Noting that $T_x\,|, t_0$ cannot be less than t if only $0, 1, \ldots, x-2$, or $x-1$ (repeat) transactions are made in the time interval $(t_o, t]$,

$$P(T_x \leq t) = 1 - \sum_{j=0}^{x-1} P(X(t_0,t) = j)\,.$$

Therefore

$$\begin{aligned} &P(X(t_0,t) = x\,|\,\lambda,p,\boldsymbol{\beta};\mathbf{Z}(t)) \\ &\quad = (1-p)^x \frac{[\lambda C(t_0,t)]^x e^{-\lambda C(t_0,t)}}{x!} \\ &\qquad + \mathbf{1}_{x>0}\,p(1-p)^{x-1}\left[1 - e^{-\lambda C(t_0,t)}\sum_{j=0}^{x-1}\frac{[\lambda C(t_0,t)]^j}{j!}\right]. \quad (9)\end{aligned}$$

Taking the expectation of (9) over the distributions of λ and p results in the following expression for the distribution of the number of repeat transactions in the time interval $(t_0, t]$:

$$
\begin{aligned}
&P(X(t_0,t) = x \,|\, r, \alpha, a, b, \boldsymbol{\beta}; \mathbf{Z}(t)) \\
&\quad = \frac{B(a, b+x)}{B(a,b)} \frac{\Gamma(r+x)}{\Gamma(r)x!} \left(\frac{\alpha}{\alpha + C(t_0,t)} \right)^r \left(\frac{C(t_0,t)}{\alpha + C(t_0,t)} \right)^x \\
&\qquad + \mathbf{1}_{x>0} \frac{B(a+1, b+x-1)}{B(a,b)} \Bigg[1 - \left(\frac{\alpha}{\alpha + C(t_0,t)} \right)^r \\
&\qquad\quad \times \left\{ \sum_{j=0}^{x-1} \frac{(\Gamma(r+j)}{\Gamma(r)j!} \left(\frac{C(t_0,t)}{\alpha + C(t_0,t)} \right)^j \right\} \Bigg] . \qquad (10)
\end{aligned}
$$

When $\boldsymbol{\beta} = \mathbf{0}$, this reduces to the basic BG/NBD model expression for $P(X(t_0,t) = x)$.

5 Conditional Expectations

We wish to derive an expression for the expected number of transactions in the time interval $(T, T + \Delta t]$ for a customer with transaction history $[t_0, \ldots, t_x]$.

If the customer is alive at T (i.e., $\omega > T$), it follows from (7) that

$$
\begin{aligned}
&E(X(T, T+\Delta t) \,|\, \lambda, p, \boldsymbol{\beta}; \mathbf{Z}(T+\Delta t); \omega > T) \\
&\quad = \frac{1}{p} - \frac{1}{p} e^{-\lambda p C(T, T+\Delta t)} . \qquad (11)
\end{aligned}
$$

It follows from (5) that

$$
\begin{aligned}
&P(\Omega > T \,|\, \lambda, p, \boldsymbol{\beta}; \mathbf{Z}(T), [t_0, \ldots, t_x]) \\
&\quad = \frac{\left\{ \prod_{j=1}^{x} A(\lceil t_j \rceil) \right\} (1-p)^x \lambda^x e^{-\lambda C(t_0,T)}}{L(\lambda, p, \boldsymbol{\beta} \,|\, \mathbf{Z}(T), [t_0, \ldots, t_x], T)} . \qquad (12)
\end{aligned}
$$

Now, the joint posterior distribution of λ and p for a customer with transaction history $[t_0, \ldots, t_x]$ is given by

$$
\begin{aligned}
&g(\lambda, p \,|\, r, \alpha, a, b, \boldsymbol{\beta}; \mathbf{Z}(T), [t_0, \ldots, t_x]) \\
&\quad = \frac{L(\lambda, p, \boldsymbol{\beta} \,|\, \mathbf{Z}(T), [t_0, \ldots, t_x]) g(\lambda \,|\, r, \alpha)\, g(p \,|\, a, b)}{L(r, \alpha, a, b, \boldsymbol{\beta} \,|\, \mathbf{Z}(T), [t_0, \ldots, t_x], T)} . \qquad (13)
\end{aligned}
$$

Integrating the product of (11)–(13) over λ and p (and simplifying) gives us the following expression for the conditional expectation:

$$E(X(T,T+\Delta t)\,|\,r,\alpha,a,b,\boldsymbol{\beta};\mathbf{Z}(T+\Delta t),[t_0,\ldots,t_x])$$
$$= \frac{\dfrac{a+b+x-1}{a-1}(1-\mathsf{AB})}{1+\mathbf{1}_{x>0}\,\dfrac{a}{b+x-1}\left(\dfrac{\alpha+C(t_0,T)}{\alpha+C(t_0,t_x)}\right)^{r+x}}\,, \qquad (14)$$

where

$$\mathsf{A} = \left(\frac{\alpha+C(t_0,T)}{\alpha+C(t_0,T+\Delta t)}\right)^{r+x}$$

and

$$\mathsf{B} = {}_2F_1\big(r+x, b+x; a+b+x-1; \tfrac{C(T,T+\Delta t)}{\alpha+C(t_0,T+\Delta t)}\big)\,.$$

When $\boldsymbol{\beta} = \mathbf{0}$, this reduces to the basic BG/NBD model expression for $E(X(T,T+\Delta t)\,|\,x,t_x,T)$.

6 The Special Case of Seasonal Effects[4]

In many cases, the desire to incorporate the effects of time-varying covariates in the BG/NBD model is driven by the need to capture seasonality in sales. If these seasonal effects are represented by dummy variables (and we have a small number of "seasons"), it is possible to simplify the model likelihood function.

To illustrate, suppose we have a "regular season" and a "high season". Assuming week is the chosen unit of time, we have one covariate defined as $\mathbf{z}(t) = 1$ if week $\lceil t \rceil$ is in the high season, 0 otherwise. Let W_{high} be the set of weeks (in the calendar year) associated with the high season, and

$$x_{\text{high}} = \sum_{j=1}^{x} \mathbf{1}_{\lceil t_j \rceil \in W_{\text{high}}}\,,$$

i.e., the number of purchases made in high-season weeks.

We now have

$$A(\lceil t \rceil) = \begin{cases} 1 & \text{if } \lceil t \rceil \text{ is a regular-season week} \\ \eta_{\text{high}} = \exp(\beta) & \text{if } \lceil t \rceil \text{ is a high-season week} \end{cases}$$
$$= \eta_{\text{high}}^{\mathbf{1}_{\lceil t \rceil \in W_{\text{high}}}}\,.$$

[4]We thank Dan McCarthy for suggesting this special case.

It follows that

$$\prod_{j=1}^{x} A(\lceil t_j \rceil) = \eta_{\text{high}}^{x_{\text{high}}} .$$

We can replace $C(t)$ with

$$D(t) = \sum_{i=1}^{\lfloor t \rfloor} \eta_{\text{high}}^{\mathbf{1}_{\lceil i \rceil \in W_{\text{high}}}} + (t - \lfloor t \rfloor)\eta_{\text{high}}^{\mathbf{1}_{\lceil t \rceil \in W_{\text{high}}}}$$

and $C(t_a, t_b)$ with $D(t_a, t_b) = D(t_b) - D(t_a)$.

The individual-level likelihood function, (5), becomes

$$\begin{aligned} &L(\lambda, p, \eta_{\text{high}} \,|\, W_{\text{high}}, t_0, t_x, x, x_{\text{high}}, T) \\ &\quad = \eta_{\text{high}}^{x_{\text{high}}} \Big\{ (1-p)^x \lambda^x e^{-\lambda D(t_0,T)} + \mathbf{1}_{x>0}\, p(1-p)^{x-1} \lambda^x e^{-\lambda D(t_0,t_x)} \Big\} . \end{aligned}$$

If $W_{\text{high}} = \varnothing$, $x_{\text{high}} = 0$ and $D(t) = t$, and this reduces to the basic BG/NBD individual-level likelihood function.[5]

It follows that

$$\begin{aligned} L(r, \alpha, &a, b, \eta_{\text{high}} \,|\, W_{\text{high}}, t_0, t_x, x, x_{\text{high}}, T) \\ &= \eta_{\text{high}}^{x_{\text{high}}} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)} \left\{ \frac{B(a, b+x)}{B(a,b)} \left(\frac{1}{\alpha + D(t_0, T)} \right)^{r+x} \right. \\ &\qquad \left. + \mathbf{1}_{x>0} \frac{B(a+1, b+x-1)}{B(a,b)} \left(\frac{1}{\alpha + D(t_0, t_x)} \right)^{r+x} \right\} . \qquad (15) \end{aligned}$$

Note that we no longer need to know the customer's entire purchase history, "recency" and "frequency" and "number of high-season transactions" are sufficient statistics.

It follows that

$$\begin{aligned} &E(X(t_0, t) \,|\, r, \alpha, a, b, \eta_{\text{high}}; W_{\text{high}}) \\ &\quad = \frac{a+b-1}{a-1} \left[1 - \left(\frac{\alpha}{\alpha + D(t_0, t)} \right)^r {}_2F_1\big(r, b; a+b-1; \tfrac{D(t_0,t)}{\alpha + D(t_0,t)}\big) \right] , \qquad (16) \end{aligned}$$

[5] One way to think about $D(t)$ is as follows. The Poisson mean of a regular-season week is λ and the Poisson mean of a high-season week is $\lambda\eta_{\text{high}}$. The Poisson mean over the period $(0, t]$ is $\lambda D(t)$. We can think of $D(t)$ as the equivalent number of "regular" weeks in $(0, t]$, ranging from t when it contains no high-season weeks to $\eta_{\text{high}} t$ when all the weeks in $(0, t]$ are high-season weeks.

$$
\begin{aligned}
&P(X(t_0,t) = x \mid r, \alpha, a, b, \eta_{\text{high}}; W_{\text{high}}) \\
&\quad = \frac{B(a, b+x)}{B(a,b)} \frac{\Gamma(r+x)}{\Gamma(r)x!} \left(\frac{\alpha}{\alpha + D(t_0,t)}\right)^r \left(\frac{D(t_0,t)}{\alpha + D(t_0,t)}\right)^x \\
&\qquad + \mathbf{1}_{x>0} \frac{B(a+1, b+x-1)}{B(a,b)} \left[1 - \left(\frac{\alpha}{\alpha + D(t_0,t)}\right)^r \right. \\
&\qquad\qquad \left. \times \left\{ \sum_{j=0}^{x-1} \frac{\Gamma(r+j)}{\Gamma(r)j!} \left(\frac{D(t_0,t)}{\alpha + D(t_0,t)}\right)^j \right\} \right] , \qquad (17)
\end{aligned}
$$

and

$$
\begin{aligned}
&E(X(T, T+\Delta t) \mid (r, \alpha, a, b, \eta_{\text{high}}; W_{\text{high}}; t_0, t_x, x, x_{\text{high}}, T) \\
&\quad = \frac{\dfrac{a+b+x-1}{a-1}(1 - \mathsf{AB})}{1 + \mathbf{1}_{x>0} \dfrac{a}{b+x-1} \left(\dfrac{\alpha + D(t_0,T)}{\alpha + D(t_0,t_x)}\right)^{r+x}} , \qquad (18)
\end{aligned}
$$

where

$$
\mathsf{A} = \left(\frac{\alpha + D(t_0,T)}{\alpha + D(t_0, T+\Delta t)}\right)^{r+x}
$$

and

$$
\mathsf{B} = {}_2F_1\big(r+x, b+x; a+b+x-1; \tfrac{D(T,T+\Delta t)}{\alpha + D(t_0, T+\Delta t)}\big) .
$$

This can easily be extended to more seasonal periods. For example, suppose we have a low season, a regular season, and a high season. We let W_{low} be the set of weeks (in the calendar year) associated with the low season, and

$$
x_{\text{low}} = \sum_{j=1}^{x} \mathbf{1}_{\lceil t_j \rceil \in W_{\text{low}}} .
$$

We add an additional parameter η_{low} and let

$$
E(t) = \sum_{i=1}^{\lfloor t \rfloor} \eta_{\text{low}}^{\mathbf{1}_{\lceil i \rceil \in W_{\text{low}}}} \eta_{\text{high}}^{\mathbf{1}_{\lceil i \rceil \in W_{\text{high}}}} + (t - \lfloor t \rfloor) \eta_{\text{low}}^{\mathbf{1}_{\lceil t \rceil \in W_{\text{low}}}} \eta_{\text{high}}^{\mathbf{1}_{\lceil t \rceil \in W_{\text{high}}}} ,
$$

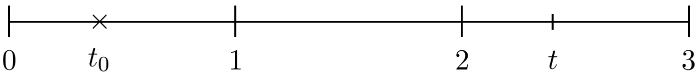
etc., which leads to

$$
\begin{aligned}
&L(\lambda, p, \eta_{\text{low}}, \eta_{\text{high}} \mid W_{\text{low}}, W_{\text{high}}, t_0, t_x, x, x_{\text{low}}, x_{\text{high}}, T) \\
&\quad = \eta_{\text{low}}^{x_{\text{low}}} \eta_{\text{high}}^{x_{\text{high}}} \Big\{ (1-p)^x \lambda^x e^{-\lambda E(t_0,T)} \\
&\qquad\qquad + \mathbf{1}_{x>0}\, p(1-p)^{x-1} \lambda^x e^{-\lambda E(t_0,t_x)} \Big\} ,
\end{aligned}
$$

and so on.

Appendix

In this appendix, we document the logic for solving the integral associated with (7) via the following example. In the timeline below, we see that the customer is acquired in the first time period and t falls in the third time period:



For this specific example, we can rewrite

$$\begin{aligned} E(X(t_0,t)\,|\,\lambda,p,\boldsymbol{\beta};\mathbf{Z}(t)) &= \lambda C(t_0,t)P(\Omega > t\,|\,\lambda,p,\boldsymbol{\beta};\mathbf{Z}(\omega),t_0) \\ &\qquad + \int_{t_0}^{t} \lambda C(t_0,\omega) g(\omega\,|\,\lambda,p,\boldsymbol{\beta};\mathbf{Z}(\omega),t_0)d\omega \end{aligned}$$

as

$$E(X(t_0,t)\,|\,\lambda,p,\boldsymbol{\beta};\mathbf{Z}(t)) = \mathsf{A}+\mathsf{B}+\mathsf{C}+\mathsf{D} \tag{A1}$$

where

$$\mathsf{A} = \lambda C(t_0,t)e^{\lambda p C(t_0,t)}\,, \tag{A2}$$

$$\mathsf{B} = \int_{t_0}^{1} \lambda C(t_0,\omega)\lambda p A(\lceil\omega\rceil)e^{-\lambda p C(t_0,\omega)}d\omega\,, \tag{A3}$$

$$\mathsf{C} = \int_{1}^{2} \lambda C(t_0,\omega)\lambda p A(\lceil\omega\rceil)e^{-\lambda p C(t_0,\omega)}d\omega\,,$$

$$\mathsf{D} = \int_{2}^{t} \lambda C(t_0,\omega)\lambda p A(\lceil\omega\rceil)e^{-\lambda p C(t_0,\omega)}d\omega\,.$$

The trick to solving these integrals is integration by parts, which states that

$$\int_a^b u(x)v'(x)dx = u(x)v(x)\Big|_a^b - \int_a^b u'(x)v(x)dx\,.$$

Let us consider (A2):

$$\begin{aligned} \mathsf{B} &= \lambda^2 pA(1)\int_{t_0}^{1} \underbrace{C(t_0,\omega)}_{u}\underbrace{e^{-\lambda p C(t_0,\omega)}}_{v'}d\omega \\ &= \lambda^2 pA(1)\left[\frac{-C(t_0,\omega)e^{-\lambda p C(t_0,\omega)}}{\lambda p A(\lceil\omega\rceil)}\Bigg|_{t_0}^{1}\right] \\ &\qquad - \lambda^2 pA(1)\int_{t_0}^{1} A(\lceil\omega\rceil)\Big(\frac{-1}{\lambda p A(\lceil\omega\rceil)}\Big)e^{-\lambda p C(t_0,\omega)}d\omega \end{aligned}$$

which noting that $C(t_0, t_0) = 0$,

$$= -\lambda C(t_0, 1)e^{-\lambda p C(t_0,1)} - \lambda^2 p A(1) \left[\frac{e^{-\lambda p C(t_0,\omega)}}{\lambda^2 p^2 A(\lceil \omega \rceil)} \Bigg|_{t_0}^{1} \right]$$

$$= -\lambda C(t_0, 1)e^{-\lambda p C(t_0,1)} + \frac{1}{p} - \frac{1}{p} e^{-\lambda p C(t_0,1)} . \tag{A4}$$

By similar logic,

$$\mathsf{C} = \lambda C(t_0, 1)e^{-\lambda p C(t_0,1)} - \lambda C(t_0, 2)e^{-\lambda p C(t_0,2)} + \frac{1}{p} e^{-\lambda p C(t_0,1)} - \frac{1}{p} e^{-\lambda p C(t_0,2)} \tag{A5}$$

and

$$\mathsf{D} = \lambda C(t_0, 2)e^{-\lambda p C(t_0,2)} - \lambda C(t_0, t)e^{-\lambda p C(t_0,t)} + \frac{1}{p} e^{-\lambda p C(t_0,2)} - \frac{1}{p} e^{-\lambda p C(t_0,t)} . \tag{A6}$$

Substituting (A2) and (A4)–(A6) in (A1) gives us (7).

References

Batislam, Emine Persentili, Meltem Denizel, and Alpay Filiztekin (2007), "Empirical Validation and Comparison of Models for Customer Base Analysis," *International Journal of Research in Marketing*, **24** (September), 201–209.

Batislam, E. P. M. Denizela and A. Filiztekin (2008), "Formal response to "Erratum on the MBG/NBD Model"," *International Journal of Research in Marketing* **25** (September), 227.

Braun, Michael, David A. Schweidel, and Eli Stein (2015), "Transaction Attributes and Customer Valuation," *Journal of Marketing Research*, **52** (December), 848–864.

Fader, Peter S. and Bruce G. S. Hardie (2007), "Incorporating Time-Invariant Covariates into the Pareto/NBD and BG/NBD Models." <`http://brucehardie.com/notes/019/`>

Fader, Peter S. and Bruce G. S. Hardie (2020), "Incorporating Time-Varying Covariates in a Simple Mixture Model for Discrete-Time Duration Data." <`http://brucehardie.com/notes/037/`>

Fader, Peter S., Bruce G. S. Hardie, and Chun-Yao Huang (2004), "A Dynamic Changepoint Model for New Product Sales Forecasting," *Marketing Science*, **23** (Winter), 50–65.

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2005), ""Counting Your Customers" the Easy Way: An Alternative to the Pareto/NBD Model," *Marketing Science*, **24** (Spring), 275–284.

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2019), "A Step-by-Step Derivation of the BG/NBD Model." <`http://brucehardie.com/notes/039/`>

Harman, David M. (2016), *Stochastic Process Customer Lifetime Value Models with Time-varying Covariates*, unpublished PhD thesis, University of Iowa.

Hoppe, Daniel and Udo Wagner (2007), "Customer Base Analysis: The Case for a Central Variant of the Betageometric/NBD Model," *Marketing — Journal of Research and Management*, **3** (2), 75–90.

Huang, Chun-Yao (2002), *An Integrated Stochastic Model for the Analysis of Web Site Visiting Behaviors*, unpublished PhD thesis, University of London.

Wagner, Udo and Daniel Hoppe (2008), "Erratum on the MBG/NBD Model," *International Journal of Research in Marketing* **25** (September), 225–226.