The Mean and Variance of Customer Lifetime Value in Contractual Settings

Peter S. Fader
www.petefader.com

Bruce G. S. Hardie[†]
www.brucehardie.com

January 2018

In the note "Exploring the Distribution of Customer Lifetime Value (in Contractual Settings)" (Fader and Hardie 2017), we show how questions about the distribution of the value of a cohort of customers can be answered easily if we know the mean and variance of CLV. In that note we computed these quantities from the distribution of CLV, assuming the distribution of customer lifetimes is characterized by the beta-geometric (BG) distribution (Fader and Hardie 2007). In this note, we derive closed-form expressions for these quantities, thus simplifying the process of answering questions about the distribution of cohort value. In presenting these derivations, we assume familiarity with the derivations presented in Fader and Hardie (2010).

1. Set-up

- We assume a discrete-time contractual setting. Let the random variable L, with realizations $l = 1, 2, 3, \ldots$, denote the lifetime of a customer.
- We assume that the distribution of lifetimes is characterized by the BG. In other words, conditional on θ, $L \sim$ geometric(θ), and θ is a realization of Θ, which is distributed beta(γ, δ).
- Given a discount rate of $d \times 100\%$, a lifetime of l periods, and period-by-period value $w_1, w_2, \ldots, w_l$, lifetime value is computed as

$$LV(d \mid L = l, w_1, w_2, \ldots, w_l) = w_1 + \frac{w_2}{1+d} + \frac{w_3}{(1+d)^2} + \cdots + \frac{w_l}{(1+d)^{l-1}}\,.$$

These quantities are only known with certainty ex post; ex ante, lifetime value a random variable.

- Suppose $w_i = \bar{w}\ \forall\, i$ (which is known with certainty). We have

$$\begin{aligned} LV(d \mid L = l) &= \bar{w}\left\{1 + \frac{1}{1+d} + \frac{1}{(1+d)^2} + \cdots + \frac{1}{(1+d)^{l-1}}\right\} \\ &= \bar{w} DL(d \mid L = l)\,, \end{aligned}$$

where the *discounted lifetime,* given a lifetime of l periods, is computed as

$$DL(d \mid L = l) = 1 + \frac{1}{1+d} + \frac{1}{(1+d)^2} + \cdots + \frac{1}{(1+d)^{l-1}}\,.$$

- So as to make our derivations a little tidier, let

$$\rho = \frac{1}{1+d}\,, \tag{1}$$

in which case

$$\begin{aligned} LV(d \mid L = l, w_1, w_2, \ldots, w_l) &= \sum_{i=1}^{l} w_i \rho^{i-1} \\ DL(d \mid L = l) &= \sum_{i=0}^{l-1} \rho^i \\ &= \frac{1-\rho^l}{1-\rho}\,. \end{aligned} \tag{2}$$

In the following derivations, we will use d and ρ in the same equation without comment.

Note that it follows from (1) that

$$\frac{1}{1-\rho} = \frac{1+d}{d}\,.$$

2. Discounted Lifetime

Conditional on $L = l$, DL is not a random variable. Let us remove that conditioning under the assumption of BG-distributed lifetimes.

- Conditional on θ, the expected discounted lifetime is

$$E[DL(d) \mid \theta] = \sum_{l=1}^{\infty} DL(d \mid L = l) P(L = l \mid \theta)$$

$$
\begin{aligned}
&= \frac{1}{1-\rho}\sum_{l=1}^{\infty}(1-\rho^l)\theta(1-\theta)^{l-1} \\
&= \left(\frac{1+d}{d}\right)\bigl[1-J(\theta,d)\bigr]\,, \qquad (3)
\end{aligned}
$$

where

$$
\begin{aligned}
J(\theta,d) &= \sum_{l=1}^{\infty}\rho^l\theta(1-\theta)^{l-1} \qquad (4)\\
&= \rho\theta\sum_{m=0}^{\infty}[\rho(1-\theta)]^m \\
&= \frac{\rho\theta}{1-\rho(1-\theta)} \\
&= \frac{\theta}{d+\theta}\,.
\end{aligned}
$$

- Let us now remove the conditioning on θ. By definition,

$$
\begin{aligned}
E[DL(d)\,|\,\gamma,\delta] &= \int_0^1 E[DL(d)\,|\,\theta]\,g(\theta\,|\,\gamma,\delta)d\theta \\
&= \left(\frac{1+d}{d}\right)\bigl[1-J(\gamma,\delta,d)\bigr]\,,
\end{aligned}
$$

where

$$
\begin{aligned}
J(\gamma,\delta,d) &= \int_0^1 J(\theta,d)g(\theta\,|\,\gamma,\delta)d\theta \\
&= \int_0^1 \frac{\theta}{d+\theta}\frac{\theta^{\gamma-1}(1-\theta)^{\delta-1}}{B(\gamma,\delta)}d\theta \\
&= \frac{1}{B(\gamma,\delta)}\int_0^1 \theta^{\gamma}(1-\theta)^{\delta-1}(d+\theta)^{-1}d\theta
\end{aligned}
$$

letting $s = 1-\theta$

$$
\begin{aligned}
&= \frac{1}{B(\gamma,\delta)}\int_0^1 s^{\delta-1}(1-s)^{\gamma}(1+d-s)^{-1}ds \\
&= \frac{1}{B(\gamma,\delta)(1+d)}\int_0^1 s^{\delta-1}(1-s)^{\gamma}(1-\tfrac{1}{1+d}s)^{-1}ds \\
&= \frac{1}{1+d}\frac{B(\gamma+1,\delta)}{B(\gamma,\delta)}\,{}_2F_1\bigl(1,\delta;\gamma+\delta+1;\tfrac{1}{1+d}\bigr) \\
&= \frac{1}{1+d}\left(\frac{\gamma}{\gamma+\delta}\right){}_2F_1\bigl(1,\delta;\gamma+\delta+1;\tfrac{1}{1+d}\bigr)\,. \qquad (5)
\end{aligned}
$$

- We now turn our attention to $\text{var}[DL(d) \mid \gamma, \delta]$, the variance of DL. As an intermediate step, we derive an expression for $E[DL(d)^2 \mid \gamma, \delta)]$.

- By definition,

$$\begin{aligned} E[DL(d)^2 \mid \theta] &= \sum_{l=1}^{\infty} DL(d \mid L = l)^2 P(L = l \mid \theta) \\ &= \frac{1}{(1-\rho)^2} \sum_{l=1}^{\infty} (1-\rho^l)^2 \theta (1-\theta)^{l-1} \\ &= \frac{1}{(1-\rho)^2} \sum_{n=1}^{\infty} (1 - 2\rho^l + \rho^{2l}) \theta (1-\theta)^{l-1} \end{aligned}$$

which, letting $\rho' = \rho^2 \;\Leftrightarrow\; d' = d(d+2)$, and recalling our definition of $J(\theta, d)$, (4),

$$\begin{aligned} &= \frac{1 - 2J(\theta, d) + J(\theta, d')}{(1-\rho)^2} \\ &= \left(\frac{1+d}{d}\right)^2 \big[1 - 2J(\theta, d) + J(\theta, d')\big] \,. \end{aligned}$$

We remove the conditioning on θ using the result given in (5).

- To summarize,

$$E[DL(d) \mid \gamma, \delta)] = \left(\frac{1+d}{d}\right) \big[1 - J(\gamma, \delta, d)\big] \,, \tag{6}$$

$$E[DL(d)^2 \mid \gamma, \delta] = \left(\frac{1+d}{d}\right)^2 \big[1 - 2J(\gamma, \delta, d) + J(\gamma, \delta, d')\big] \,, \tag{7}$$

$$\text{var}[DL(d)^2 \mid \gamma, \delta] = E[DL(d)^2 \mid \gamma, \delta] - E[DL(d) \mid \gamma, \delta]^2 \,, \tag{8}$$

where $d' = d(d+2)$ and

$$J(\gamma, \delta, d) = \frac{1}{1+d} \left(\frac{\gamma}{\gamma+\delta}\right) {}_2F_1\big(1, \delta; \gamma+\delta+1; \tfrac{1}{1+d}\big) \,.$$

- We present an example of these calculations in Appendix A.

3. Lifetime Value

- We now turn our attention to the derivation of expressions for the mean and variance of LV.

- Recall

$$LV(d \mid L = l, w_1, w_2, \ldots, w_l) = \sum_{i=1}^{l} w_i \rho^{i-1} \,.$$

For a given individual, we assume the w_i are realizations of W_i, which are iid with pdf $f(w \mid \eta)$. We assume that η varies across individuals with distribution $g(\eta \mid \xi)$ (which is independent of the distribution of Θ).

- Given $L = l$ and η, LV is the random variable

$$LV(d \mid L = l, \eta) = \sum_{i=1}^{l} W_i \rho^{i-1}$$

with mean

$$\begin{aligned} E[LV(d \mid L = l, \eta)] &= \sum_{i=1}^{l} E(W_i \mid \eta)\rho^{i-1} \\ &= E(W \mid \eta) \sum_{i=1}^{l} \rho^{i-1} \\ &= E(W \mid \eta) DL(d \mid L = l) \,. \end{aligned}$$

- Removing the conditioning on $L = l$, we have

$$E[LV(d \mid \eta) \mid \gamma, \delta] = E(W \mid \eta) E[DL(d) \mid \gamma, \delta] \,.$$

- Removing the conditioning on η, we have

$$E[LV(d) \mid \gamma, \delta, \xi] = E(W \mid \xi) E[DL(d) \mid \gamma, \delta] \,. \tag{9}$$

- In order to compute the variance of LV, we need an expression for $E[LV(d)^2 \mid \gamma, \delta, \xi]$. As a first step, we derive an expression for $E[LV(d \mid L = l, \eta)^2]$.

- Approach 1: Since the W_i are iid,

$$\begin{aligned} \mathrm{var}[LV(d \mid L = l, \eta)] &= \sum_{i=1}^{l} \mathrm{var}(W_i \mid \eta)(\rho^{i-1})^2 \\ &= \mathrm{var}(W \mid \eta) \sum_{i=1}^{l} (\rho^2)^{i-1} \\ &= \frac{1 - \rho^{2l}}{1 - \rho^2} \left[E(W^2 \mid \eta) - E(W \mid \eta)^2 \right] . \end{aligned}$$

It follows that

$$
\begin{aligned}
&E[LV(d \mid L=l,\eta)^2] \\
&= \mathrm{var}[LV(d \mid L=l,\eta)] + E[LV(d \mid L=l,\eta)]^2 \\
&= \frac{1-\rho^{2l}}{1-\rho^2}E(W^2 \mid \eta) - \frac{1-\rho^{2l}}{1-\rho^2}E(W \mid \eta)^2 + \left\{\frac{1-\rho^l}{1-\rho}\right\}^2 E(W \mid \eta)^2 . \qquad (10)
\end{aligned}
$$

- Approach 2: By definition,

$$
\begin{aligned}
E[LV(d \mid L=l,\eta)^2] &= E[(W_1+\rho W_2+\cdots+\rho^{l-1}W_l)^2] \\
&= \sum_{i=1}^{l}\sum_{j=1}^{l}\rho^{i-1}\rho^{j-1}E(W_iW_j \mid \eta) .
\end{aligned}
$$

Noting that $E(W_iW_j \mid \eta) = E(W^2 \mid \eta)$ if $i=j$, $E(W \mid \eta)^2$ otherwise, we have

$$
\begin{aligned}
E[LV(d \mid L=l,\eta)^2] &= \sum_{i=1}^{l}(\rho^{i-1})^2E(W^2 \mid \eta) + 2\sum_{i=1}^{l}\sum_{j=i+1}^{l}\rho^{i-1}\rho^{j-1}E(W \mid \eta)^2 \\
&= E(W^2 \mid \eta)\sum_{i=1}^{l}(\rho^2)^{i-1} \\
&\qquad + E(W \mid \eta)^2 \underbrace{\left\{\sum_{i=1}^{l}\sum_{j=1}^{l}\rho^{i-1}\rho^{j-1} - \sum_{i=1}^{l}(\rho^{i-1})^2\right\}}_{\text{sum of all terms minus sum of diagonal}}
\end{aligned}
$$

which, noting that the finite double series can be written as a product of series[1]

$$
\begin{aligned}
&= E(W^2 \mid \eta)\sum_{i=1}^{l}(\rho^2)^{i-1} \\
&\qquad + E(W \mid \eta)^2\left\{\left(\sum_{i=1}^{l}\rho^{i-1}\right)^2 - \sum_{i=1}^{l}(\rho^2)^{i-1}\right\} \\
&= \frac{1-\rho^{2l}}{1-\rho^2}E(W^2 \mid \eta) + \left\{\frac{1-\rho^l}{1-\rho}\right\}^2 E(W \mid \eta)^2 \\
&\qquad - \frac{1-\rho^{2l}}{1-\rho^2}E(W \mid \eta)^2 ,
\end{aligned}
$$

which is the same as (10).

[1] `http://mathworld.wolfram.com/DoubleSeries.html`

- Recalling (2) and our definition of d',

$$\frac{1-\rho^{2l}}{1-\rho^2} = DL(d' \,|\, L=l)\,.$$

Therefore,

$$\begin{aligned} E[LV(d \,|\, L=l,\eta)^2] = \big\{E(W^2 \,|\, \eta) - E(W \,|\, \eta)^2\big\} DL(d' \,|\, L=l) \\ + E(W \,|\, \eta)^2 DL(d \,|\, L=l)^2\,. \end{aligned}$$

- Removing the conditioning on $L=l$, we get

$$\begin{aligned} E[LV(d \,|\, \eta)^2 \,|\, \gamma,\delta] = \big\{E(W^2 \,|\, \eta) - E(W \,|\, \eta)^2\big\} E[DL(d') \,|\, \gamma,\delta] \\ + E(W \,|\, \eta)^2 E[DL(d)^2 \,|\, \gamma,\delta]\,. \end{aligned} \tag{11}$$

- Removing the conditioning on η, we get

$$\begin{aligned} E[LV(d)^2 \,|\, \gamma,\delta,\xi] = \big\{E(W^2 \,|\, \xi) - E[E(W \,|\, \eta)^2 \,|\, \xi]\big\} E[DL(d') \,|\, \gamma,\delta] \\ + E[E(W \,|\, \eta)^2 \,|\, \xi] E[DL(d)^2 \,|\, \gamma,\delta]\,, \end{aligned} \tag{12}$$

where

$$E(W^2 \,|\, \xi) = \int_0^\infty E(W^2 \,|\, \eta) g(\eta \,|\, \xi) d\eta\,, \text{ and} \tag{13}$$

$$E[E(W \,|\, \eta)^2 \,|\, \xi] = \int_0^\infty E(W \,|\, \eta)^2 g(\eta \,|\, \xi) d\eta\,. \tag{14}$$

- See Appendix B for a specific example of $E(W \,|\, \xi)$, $E(W^2 \,|\, \xi)$, and $E[E(W \,|\, \eta)^2 \,|\, \xi]$ under the assumption of value being distributed according to the gamma-gamma model (Fader and Hardie 2013).

- It follows that

$$\text{var}[LV(d) \,|\, \gamma,\delta,\xi] = E[LV(d)^2 \,|\, \gamma,\delta,\xi] - E[LV(d) \,|\, \gamma,\delta,\xi]^2\,.$$

- Aside: We can write (11) as

$$E[LV(d)^2 \,|\, \gamma,\delta,\eta] = \text{var}(W \,|\, \eta) E[DL(d') \,|\, \gamma,\delta] + E(W \,|\, \eta)^2 E[DL(d)^2 \,|\, \gamma,\delta]\,.$$

Suppose there is no within-individual variance in per-period value; in other words, variation in value is purely cross-sectional. This means $\text{var}(W \,|\, \eta) = 0$ and therefore

$$E[LV(d)^2 \,|\, \gamma,\delta,\eta] = E(W \,|\, \eta)^2 E[DL(d)^2 \,|\, \gamma,\delta]\,.$$

It follows that

$$\begin{aligned}\operatorname{var}[LV(d) \mid \gamma, \delta, \eta] &= E[LV(d)^2 \mid \gamma, \delta, \eta] - E[LV(d) \mid \gamma, \delta, \eta]^2 \\ &= E(W \mid \eta)^2 E[DL(d)^2 \mid \gamma, \delta] - E(W \mid \eta)^2 E[DL(d) \mid \gamma, \delta]^2 \\ &= E(W \mid \eta)^2 \big\{E[DL(d)^2 \mid \gamma, \delta] - E[DL(d) \mid \gamma, \delta]^2\big\} \\ &= E(W \mid \eta)^2 \operatorname{var}[DL(d) \mid \gamma, \delta]\,.\end{aligned}$$

This is what we would expect given the basic property of variances: $\operatorname{var}(cX) = c^2\operatorname{var}(X)$ for constant c.

- Pushing this a little further, suppose there is no variation in within-individual value (i.e., $w_i = \bar{w}\ \forall\, i$) but there is between-individual variation captured by $f(\bar{w} \mid \xi)$. Conditional on $\bar{W} = \bar{w}$,

$$\begin{aligned}E[LV(d \mid \bar{w}) \mid \gamma, \delta] &= \bar{w} E[DL(d) \mid \gamma, \delta]\,, \\ E[LV(d \mid \bar{w})^2 \mid \gamma, \delta] &= \bar{w}^2 E[DL(d)^2 \mid \gamma, \delta]\,.\end{aligned}$$

- Removing the conditioning on $\bar{w}$ gives us

$$E[LV(d) \mid \gamma, \delta, \xi] = E(\bar{W} \mid \xi) E[DL(d) \mid \gamma, \delta]\,, \tag{15}$$

$$E[LV(d)^2 \mid \gamma, \delta, \xi] = E(\bar{W}^2 \mid \xi) E[DL(d)^2 \mid \gamma, \delta]\,. \tag{16}$$

- When there is no variation in $\bar{w}$ across customers (i.e., $\operatorname{var}(\bar{W}^2 \mid \xi) = 0$, it is obviously the case that $E(\bar{W}^2 \mid \xi) = E(\bar{W} \mid \xi)^2 = \bar{w}^2$, and therefore

$$\begin{aligned}E[LV(d) \mid \gamma, \delta, \bar{w}] &= \bar{w} E[DL(d) \mid \gamma, \delta]\,, \\ E[LV(d)^2 \mid \gamma, \delta, \bar{w}] &= \bar{w}^2 E[DL(d)^2 \mid \gamma, \delta]\,, \\ \operatorname{var}[LV(d) \mid \gamma, \delta, \bar{w}] &= \bar{w}^2 \operatorname{var}[DL(d) \mid \gamma, \delta]\,.\end{aligned}$$

4. Residual Lifetime Value

Suppose a customer acquired at the beginning of period 1 is still a customer in period n. Standing at the end of period n, what are the mean and variance of the *residual* lifetime value of this customer?

- Given a residual lifetime of $l\,(> 0)$ periods and period-by-period value w_{n+1}, $w_{n+2}, \ldots, w_{n+l}$, residual lifetime value is computed as

$$\begin{aligned}&RLV(d, \text{active for } n \text{ periods} \mid RL = l, w_{n+1}, w_{n+2}, \ldots, w_{n+l}) \\ &\qquad = w_{n+1} + \frac{w_{n+2}}{1+d} + \frac{w_{n+3}}{(1+d)^2} + \cdots + \frac{w_{n+l}}{(1+d)^{l-1}}\,.\end{aligned}$$

(If $RL = 0$, the residual lifetime value is obviously zero.) These quantities are only known with certainty ex post; ex ante, residual lifetime value is a random variable.

- Suppose $w_i = \bar{w}\ \forall\, i$ (which is known with certainty). We have

$$
\begin{aligned}
&RLV(d, \text{active for } n \text{ periods} \,|\, RL = l) \\
&\qquad = \bar{w}\left\{1 + \frac{1}{1+d} + \frac{1}{(1+d)^2} + \cdots + \frac{1}{(1+d)^{l-1}}\right\} \\
&\qquad = \bar{w} DRL(d \,|\, L = l)\,,
\end{aligned}
$$

where the *discounted residual lifetime,* given a residual lifetime of $l\,(> 0)$ periods, is computed as

$$
\begin{aligned}
&DRL(d, \text{active for } n \text{ periods} \,|\, RL = l) \\
&\qquad = 1 + \frac{1}{1+d} + \frac{1}{(1+d)^2} + \cdots + \frac{1}{(1+d)^{l-1}} \\
&\qquad = \frac{1-\rho^l}{1-\rho}\,.
\end{aligned}
$$

If $RL = 0$, the discounted residual lifetime is zero.

- Conditional on θ, the probability of a residual lifetime of l periods ($l = 0, 1, 2, \ldots$) is

$$
\begin{aligned}
P(RL = l \,|\, \theta, \text{active for } n \text{ periods}) &= \frac{P(L = n + l \,|\, \theta)}{S(n-1 \,|\, \theta)} \\
&= \frac{\theta(1-\theta)^{n+l-1}}{(1-\theta)^{n-1}} \\
&= \theta(1-\theta)^l\,.
\end{aligned}
$$

- It follows that, conditional on θ, the expected discounted residual lifetime is

$$
\begin{aligned}
&E[DRL(d, \text{active for } n \text{ periods}) \,|\, \theta] \\
&\qquad = \sum_{l=0}^{\infty} \big\{DRL(d, \text{active for } n \text{ periods} \,|\, L = l) \\
&\qquad\qquad \times P(RL = l \,|\, \theta, \text{active for } n \text{ periods})\big\} \\
&\qquad = \frac{1}{1-\rho}\sum_{l=1}^{\infty}(1-\rho^l)\theta(1-\theta)^l \\
&\qquad = \left(\frac{1+d}{d}\right)\big[(1-\theta) - K(\theta, d)\big]
\end{aligned}
$$

where

$$
K(\theta, d) = \sum_{l=1}^{\infty} \rho^l \theta (1-\theta)^l \tag{17}
$$

$$
\begin{aligned}
&= \rho\theta(1-\theta)\sum_{m=0}^{\infty}[\rho(1-\theta)]^m \\
&= \frac{\rho\theta(1-\theta)}{1-\rho(1-\theta)} \\
&= \frac{\theta(1-\theta)}{d+\theta}\,.
\end{aligned}
$$

- The posterior distribution of Θ for a customer active for n periods is

$$
\begin{aligned}
g(\theta \mid \gamma, \delta, \text{active for } n \text{ periods}) &= \frac{S(n-1 \mid \theta) g(\theta \mid \gamma, \delta)}{S(n-1 \mid \gamma, \delta)} \\
&= \frac{\theta^{\gamma-1}(1-\theta)^{\delta+n-2}}{B(\gamma, \delta+n-1)}\,.
\end{aligned}
$$

- Let us now remove the conditioning on θ. By definition,

$$
\begin{aligned}
&E[DRL(d, \text{active for } n \text{ periods}) \mid \gamma, \delta] \\
&\quad = \int_0^1 E[DRL(d, \text{active for } n \text{ periods}) \mid \theta]\, g(\theta \mid \gamma, \delta, \text{active for } n \text{ periods}) d\theta \\
&\quad = \left(\frac{1+d}{d}\right)\left[\left(\frac{\delta+n-1}{\gamma+\delta+n-1}\right) - K(\gamma, \delta, d, n)\right], \qquad (18)
\end{aligned}
$$

where

$$
\begin{aligned}
K(\gamma, \delta, d, n) &= \int_0^1 K(\theta, d)\, g(\theta \mid \gamma, \delta, \text{active for } n \text{ periods}) d\theta \\
&= \int_0^1 \frac{\theta(1-\theta)}{d+\theta} \frac{\theta^{\gamma-1}(1-\theta)^{\delta+n-2}}{B(\gamma, \delta+n-1)} d\theta \\
&= \frac{1}{B(\gamma, \delta+n-1)} \int_0^1 \theta^{\gamma}(1-\theta)^{\delta+n-1}(d+\theta)^{-1} d\theta \\
&= \frac{1}{B(\gamma, \delta+n-1)} \int_0^1 s^{\delta+n-1}(1-s)^{\gamma}(1+d-s)^{-1} ds \\
&= \frac{1}{B(\gamma, \delta+n-1)(1+d)} \int_0^1 s^{\delta+n-1}(1-s)^{\gamma}(1-\tfrac{1}{1+d}s)^{-1} ds \\
&= \frac{1}{1+d} \frac{B(\gamma+1, \delta+n)}{B(\gamma, \delta+n-1)} {}_2F_1\big(1, \delta+n; \gamma+\delta+n+1; \tfrac{1}{1+d}\big) \\
&= \left(\frac{1}{1+d}\right) \frac{\gamma(\delta+n-1)}{(\gamma+\delta+n-1)(\gamma+\delta+n)} \\
&\qquad \times {}_2F_1\big(1, \delta+n; \gamma+\delta+n+1; \tfrac{1}{1+d}\big)\,. \qquad (19)
\end{aligned}
$$

- We now turn to var$[DRL(d, \text{active for } n \text{ periods}) \,|\, \gamma, \delta]$, the variance of DRL for a customer active for n periods. As for DL, we first derive an expression for $E[DRL(d, \text{active for } n \text{ periods})^2 \,|\, \gamma, \delta,]$:

$$\begin{aligned}
&E[DRL(d, \text{active for } n \text{ periods})^2 \,|\theta] \\
&\qquad\qquad = \sum_{l=0}^{\infty} \big\{DRL(d, \text{active for } n \text{ periods} \,|\, L = l)^2 \\
&\qquad\qquad\qquad\qquad \times P(RL = l \,|\, \theta, \text{active for } n \text{ periods})\big\} \\
&= \frac{1}{(1-\rho)^2} \sum_{l=1}^{\infty} (1-\rho^l)^2 \theta (1-\theta)^l \\
&= \frac{1}{(1-\rho)^2} \sum_{n=1}^{\infty} (1 - 2\rho^l + \rho^{2l}) \theta (1-\theta)^l
\end{aligned}$$

which, letting $\rho' = \rho^2 \;\Leftrightarrow\; d' = d(d+2)$, and recalling our definition of $K(\theta, d)$, (17),

$$\begin{aligned}
&= \frac{(1-\theta) - 2K(\theta, d) + K(\theta, d')}{(1-\rho)^2} \\
&= \left(\frac{1+d}{d}\right)^2 \big[(1-\theta) - 2K(\theta, d) + K(\theta, d')\big] \,.
\end{aligned}$$

We remove the conditioning on θ using the results given in (18) and (19).

- We therefore have

$$\begin{aligned}
&E[DRL(d, \text{active for } n \text{ periods}) \,|\, \gamma, \delta] \\
&\quad = \left(\frac{1+d}{d}\right)\left[\left(\frac{\delta+n-1}{\gamma+\delta+n-1}\right) - K(\gamma, \delta, d, n)\right],
\end{aligned} \tag{20}$$

$$\begin{aligned}
&E[DRL(d, \text{active for } n \text{ periods})^2 \,|\, \gamma, \delta] \\
&\quad = \left(\frac{1+d}{d}\right)^2\left[\left(\frac{\delta+n-1}{\gamma+\delta+n-1}\right) - 2K(\gamma, \delta, d, n) + K(\gamma, \delta, d', n)\right],
\end{aligned} \tag{21}$$

$$\begin{aligned}
&\text{var}[DRL(d, \text{active for } n \text{ periods}) \,|\, \gamma, \delta] \\
&\quad = E[DRL(d, \text{active for } n \text{ periods})^2 \,|\, \gamma, \delta] \\
&\qquad\quad - E[DRL(d, \text{active for } n \text{ periods}) \,|\, \gamma, \delta]^2 \,,
\end{aligned} \tag{22}$$

where $d' = d(d+2)$ and

$$\begin{aligned}
K(\gamma, \delta, d, n) = \left(\frac{1}{1+d}\right) &\frac{\gamma(\delta+n-1)}{(\gamma+\delta+n-1)(\gamma+\delta+n)} \\
&\times {}_2F_1\big(1, \delta+n; \gamma+\delta+n+1; \tfrac{1}{1+d}\big) \,.
\end{aligned}$$

- We present an example of these calculations in Appendix A.
- Drawing on the derivation of our expressions for the mean and variance of LV, it follows that

$$\begin{aligned}
&E[RLV(d, \text{active for } n \text{ periods}) \,|\, \gamma, \delta, \xi] \\
&\qquad = E(W \,|\, \xi) E[DRL(d, \text{active for } n \text{ periods}) \,|\, \gamma, \delta] \,,
\end{aligned} \tag{23}$$

$$\begin{aligned}
&E[RLV(d, \text{active for } n \text{ periods})^2 \,|\, \gamma, \delta, \xi] \\
&\qquad = \left\{E(W^2 \,|\, \xi) - E[E(W \,|\, \eta)^2 \,|\, \xi]\right\} E[DRL(d', \text{active for } n \text{ periods}) \,|\, \gamma, \delta] \\
&\qquad\qquad + E[E(W \,|\, \eta)^2 \,|\, \xi] E[DRL(d, \text{active for } n \text{ periods})^2 \,|\, \gamma, \delta] \,,
\end{aligned} \tag{24}$$

$$\begin{aligned}
&\operatorname{var}[RLV(d, \text{active for } n \text{ periods}) \,|\, \gamma, \delta, \xi] \\
&\qquad\qquad = E[RLV(d, \text{active for } n \text{ periods})^2 \,|\, \gamma, \delta, \xi] \\
&\qquad\qquad\qquad - E[RLV(d, \text{active for } n \text{ periods}) \,|\, \gamma, \delta, \xi]^2 \,.
\end{aligned} \tag{25}$$

Appendix A

In this appendix we present an illustrative example of computing the mean and variance of *DL* and *DRL*. Our BG parameter estimates are those obtained in Fader and Hardie (2014),[A1] which we use in Fader and Hardie (2017). We assume a 10% discount rate.

We perform our calculations in MATLAB and use the `h2f1` function developed in Fader et al. (2005) to evaluate the Gaussian hypergeometric function. Executing the following script

```
gamma = 0.760;
delta = 1.286;
d = 0.1;
j1 = (1/(1+d))*(gamma/(gamma+delta))...
    *h2f1(1,delta,gamma+delta+1,1/(1+d));
EDL = (1+d)/d*(1-j1)
d1 = d*(d+2);
j2 = (1/(1+d1))*(gamma/(gamma+delta))...
    *h2f1(1,delta,gamma+delta+1,1/(1+d1));
EDL2 = ((1+d)/d)^2*(1-2*j1+j2);
VDL = EDL2 - EDL^2
```

gives us

```
EDL = 3.6221

VDL = 10.5667
```

In the spreadsheet `mean_and_variance_of_DL_and_DRL.xlsx`, we compute the distribution of *DL*. The resulting estimates of the $E[DL]$ and $\mathrm{var}[DL]$ equal those computed above.

To illustrate the equivalent discounted residual lifetime calculations, let us consider an individual who has been a customer for five periods. Executing the following script

```
n = 5;
k1 = (1/(1+d))*(gamma*(delta+n-1)/((gamma+delta+n-1)...
    *(gamma+delta+n)))...
    *h2f1(1,delta+n,gamma+delta+n+1,1/(1+d));
k2 = (1/(1+d1))*(gamma*(delta+n-1)/((gamma+delta+n-1)...
    *(gamma+delta+n)))...
    *h2f1(1,delta+n,gamma+delta+n+1,1/(1+d1));
EDRL = (1+d)/d*(((delta+n-1)/(gamma+delta+n-1))-k1)
```

[A1] These are the NLS estimates based off the retention curve, not the ML estimates.

```
EDRL2 = ((1+d)/d)^2*(((delta+n-1)/(gamma+delta+n-1))-2*k1+k2);
VDRL = EDRL2 - EDRL^2
```

gives us

```
EDRL = 5.6951

VDRL = 16.3763
```

which match those calculated from the distribution of *DRL* (as computed in the spreadsheet `mean_and_variance_of_DL_and_DRL.xlsx`).

Appendix B

In order to compute the mean and variance of LV and RLV, we need $E(W \mid \xi)$, $E(W^2 \mid \xi)$, and $E[E(W \mid \eta)^2 \mid \xi]$. Note that these are generic expressions for a per-period value distribution where the w_i are realizations of W_i, which are iid with pdf $f(w \mid \eta)$, and η varies across individuals with distribution $g(\eta \mid \xi)$.

Let us consider the specific case where per-period value is characterized by the gamma-gamma model (Fader and Hardie 2013). In this case, $W \sim \text{gamma}(p, \nu)$, with $E(W \mid p, \nu) = p/\nu$ and $E(W^2 \mid p, \nu) = p(p+1)/\nu^2$, and $N \sim \text{gamma}(q, \gamma)$.[B1] It follows that

$$\begin{aligned}
E(W \mid p, q, \gamma) &= \int_0^\infty \frac{p}{\nu} \frac{\gamma^q \nu^{q-1} e^{-\gamma\nu}}{\Gamma(q)} d\nu \\
&= \frac{p\gamma}{q-1}, \\
E(W^2 \mid p, q, \gamma) &= \int_0^\infty \frac{p(p+1)}{\nu^2} \frac{\gamma^q \nu^{q-1} e^{-\gamma\nu}}{\Gamma(q)} d\nu \\
&= \frac{p(p+1)\gamma^2}{(q-1)(q-2)}, \\
E[E(W \mid \nu)^2 \mid p, q, \gamma]) &= \int_0^\infty \frac{p^2}{\nu^2} \frac{\gamma^q \nu^{q-1} e^{-\gamma\nu}}{\Gamma(q)} d\nu \\
&= \frac{p^2\gamma^2}{(q-1)(q-2)}.
\end{aligned}$$

[B1] Note the notation clash. This γ parameter is not the same as that in the main text of this note.

References

Fader, Peter S. and Bruce G.S. Hardie (2007), "How to Project Customer Retention," *Journal of Interactive Marketing*, **21** (Winter), 76–90.

Fader, Peter S. and Bruce G. S. Hardie (2010), "Customer-Base Valuation in a Contractual Setting: The Perils of Ignoring Heterogeneity," *Marketing Science*, **29** (January–February), 85–93.

Fader, Peter S. and Bruce G.S. Hardie (2013), "The Gamma-Gamma Model of Monetary Value." <`http://brucehardie.com/notes/025/`>

Fader, Peter S. and Bruce G.S. Hardie (2014), "A Spreadsheet-Literate Non-Statistician's Guide to the Beta-Geometric Model." <`http://brucehardie.com/notes/032/`>

Fader, Peter S. and Bruce G.S. Hardie (2017), "Exploring the Distribution of Customer Lifetime Value (in Contractual Settings)." <`http://brucehardie.com/notes/035/`>

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2005), "A Note on Implementing the Pareto/NBD Model in MATLAB." <`http://brucehardie.com/notes/008/`>