Technical Appendix†

"Customer-Base Valuation in a Contractual Setting:
The Perils of Ignoring Heterogeneity"

November 2008
Updated January 2011

This appendix contains step-by-step derivations of the key results presented in Fader and Hardie (2010), along with additional analyses not included in the paper.

1. Deriving an Expression for Discounted Expected (Residual) Lifetime

As we seek to derive an expression for a customer's discounted expected (residual) lifetime, we must be very careful to identify exactly where in that customer's relationship with the firm that the estimate is being made. To illustrate this, consider the timeline presented in Figure 1, where 0 corresponds to the point in time at which the individual becomes a customer of the firm.

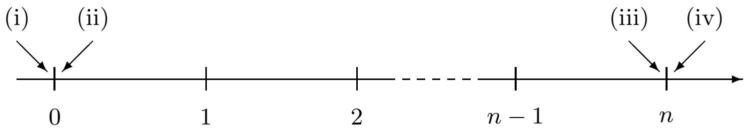


Figure 1: Different points at which a customer's discounted expected (residual) lifetime can be computed.

Assuming that the constant expected net cashflow associated with the contractual relationship is "booked" at the beginning of the contract period,

- (i) corresponds to the point in time *just before* we receive the first payment that signals the start of the customer's relationship with the firm,
- (ii) corresponds to the point in time *immediately after* we receive the first payment that signals the start of the customer's relationship with the firm,
- (iii) corresponds to the point in time *just before* we find out whether or not a customer who has renewed his contract $n-1$ times will make an nth contract renewal, and
- (iv) corresponds to the point in time *immediately after* we have received the payment associated with the customer's nth contract renewal.

(Case (iii) is the situation considered in the paper.)

Case (i)

The discounted expected lifetime of an as-yet-to-be-acquired customer is computed as

$$DEL(d) = \sum_{t=0}^{\infty} \frac{S(t)}{(1+d)^t}\,,$$

† This document can be found at <`http://brucehardie.com/notes/020/`>.

where d is the appropriate discount factor used to represent the time value of money. (When there are m contract periods per year, an annual discount rate of $(100 \times r)\%$ maps to a period discount rate of $d = (1+r)^{1/m} - 1$.)

Noting that the sum of an infinite geometric series is

$$\sum_{n=0}^{\infty} ak^n = \frac{a}{1-k},$$

it follows that when lifetimes are distributed shifted-geometric with parameter θ,

$$\begin{aligned} DEL(d \,|\, \theta) &= \sum_{t=0}^{\infty} \Big(\frac{1-\theta}{1+d}\Big)^t \\ &= \frac{1+d}{d+\theta}\,. \end{aligned} \tag{B1}$$

As θ is unobserved, we take the expectation of (B1) over the beta distribution of θ:

$$DEL(d \,|\, \alpha, \beta) = \int_0^1 \frac{1+d}{d+\theta} \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}\, d\theta$$

letting $s = 1-\theta$

$$= \frac{1}{B(\alpha,\beta)} \int_0^1 s^{\beta-1}(1-s)^{\alpha-1}\big(1 - \big(\tfrac{1}{1+d}\big)s\big)^{-1}\, ds$$

which, recalling the integral representation of the Gaussian hypergeometric function (Wolfram Research 2001a),

$$= {}_2F_1\big(1, \beta; \alpha+\beta; \tfrac{1}{1+d}\big)\,. \tag{B2}$$

Multiplying this quantity by the expected net value associated with each contract payment would give us the expected customer lifetime value of an as-yet-to-be-acquired customer, and could be thought of as an upper bound on the amount that can be spent to acquire a customer.

Noting that ${}_2F_1(a, b; c; 1) = \Gamma(c)\Gamma(c-a-b)/\Gamma(c-a)\Gamma(c-b)$, we see that

$$\begin{aligned} DEL(d=0 \,|\, \alpha, \beta) &= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha-1)}{\Gamma(\alpha+\beta-1)\Gamma(\alpha)} \\ &= \frac{\alpha+\beta-1}{\alpha-1} \end{aligned}$$

which is the mean of the sBG distribution (i.e., the discounted expected lifetime becomes the expected lifetime when the discount rate is set to 0).

Case (ii)

The discounted expected residual lifetime of a just-acquired customer is computed as

$$DERL(d) = \sum_{t=1}^{\infty} \frac{S(t)}{(1+d)^t}\,.$$

(Note that this equals $DEL(d)-1$, since it does not count the first-ever purchase by the customer.)

When survival times are distributed shifted-geometric with parameter θ we have

$$
\begin{aligned}
DERL(d \mid \theta) &= \sum_{t=1}^{\infty} \Big(\frac{1-\theta}{1+d}\Big)^{t} \\
&= \frac{1-\theta}{d+\theta}\,,
\end{aligned} \tag{B3}
$$

which is equal to $DEL(d \mid \theta) - 1$.

Taking the expectation of this over the beta distribution of θ gives us

$$
DERL(d \mid \alpha, \beta) = \int_0^1 \frac{1-\theta}{d+\theta} \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}\, d\theta
$$

letting $s = 1-\theta$

$$
\begin{aligned}
&= \frac{1}{B(\alpha,\beta)(1+d)} \int_0^1 s^{\beta}(1-s)^{\alpha-1}\big(1-\big(\tfrac{1}{1+d}\big)s\big)^{-1}\, ds \\
&= \frac{\beta}{(\alpha+\beta)(1+d)}\, {}_2F_1\big(1, \beta+1; \alpha+\beta+1; \tfrac{1}{1+d}\big)\,.
\end{aligned} \tag{B4}
$$

Case (iii)

Standing at the end of period n, just prior to the point in time at which the contract renewal decision is made (i.e., the customer has renewed his contract $n-1$ times and we have yet to learn whether or not the nth contract renewal will be made), the discounted expected residual lifetime is computed as

$$
\begin{aligned}
DERL(d \mid \text{active for } n \text{ periods}) &= \sum_{t=n}^{\infty} S(t \mid t > n-1)\Big(\frac{1}{1+d}\Big)^{t-n} \\
&= \sum_{t=n}^{\infty} \frac{S(t)}{S(n-1)}\Big(\frac{1}{1+d}\Big)^{t-n}.
\end{aligned} \tag{B5}
$$

When survival times are distributed shifted-geometric with parameter θ we have

$$
\begin{aligned}
DERL(d \mid \theta, \text{active for } n \text{ periods}) &= \sum_{t=n}^{\infty} \frac{(1-\theta)^{t-n+1}}{(1+d)^{t-n}} \\
&= (1-\theta)\sum_{s=0}^{\infty}\Big(\frac{1-\theta}{1+d}\Big)^{s} \\
&= \frac{(1-\theta)(1+d)}{d+\theta}\,.
\end{aligned} \tag{B6}
$$

When $n > 1$ we cannot take the expectation of (B6) using the original beta distribution $f(\theta \mid \alpha, \beta)$, since the fact that the person has survived to period n (i.e., has made $n-1$ renewals) means that she is more likely to have a lower-than-average value of θ. We therefore use the posterior distribution of θ for an individual who is still a customer in period n. By Bayes' theorem,

$$
\begin{aligned}
f(\theta \mid \alpha, \beta, \text{active for } n \text{ periods}) &= \frac{S(n-1 \mid \theta) f(\theta \mid \alpha, \beta)}{S(n-1 \mid \alpha, \beta)} \\
&= \frac{\theta^{\alpha-1}(1-\theta)^{\beta+n-2}}{B(\alpha, \beta+n-1)} \,.
\end{aligned} \tag{B7}
$$

It follows that

$$
\begin{aligned}
&DERL(d \mid \alpha, \beta, \text{active for } n \text{ periods}) \\
&\qquad = \int_0^1 \frac{(1-\theta)(1+d)}{d+\theta} \frac{\theta^{\alpha-1}(1-\theta)^{\beta+n-2}}{B(\alpha, \beta+n-1)} \, d\theta
\end{aligned}
$$

letting $s = 1 - \theta$

$$
\begin{aligned}
&= \frac{1}{B(\alpha, \beta+n-1)} \int_0^1 s^{\beta+n-1}(1-s)^{\alpha-1}\big(1 - \big(\tfrac{1}{1+d}\big)s\big)^{-1} ds \\
&= \left(\frac{\beta+n-1}{\alpha+\beta+n-1}\right) {}_2F_1\big(1, \beta+n; \alpha+\beta+n; \tfrac{1}{1+d}\big) \,.
\end{aligned} \tag{B8}
$$

Case (iv)

The discounted expected residual lifetime of a customer evaluated *immediately after* we have received the payment associated with her nth contract renewal is computed as

$$
\begin{aligned}
DERL(d \mid n \text{ contract renewals})] &= \sum_{t=n+1}^{\infty} S(t \mid t > n) \Big(\frac{1}{1+d}\Big)^{t-n} \\
&= \sum_{t=n+1}^{\infty} \frac{S(t)}{S(n)} \Big(\frac{1}{1+d}\Big)^{t-n} \,.
\end{aligned}
$$

When survival times are distributed shifted-geometric with parameter θ we have

$$
\begin{aligned}
DERL(d \mid \theta, n) &= \sum_{t=n+1}^{\infty} \Big(\frac{1-\theta}{1+d}\Big)^{t-n} \\
&= \frac{1-\theta}{d+\theta} \,.
\end{aligned} \tag{B9}
$$

By Bayes' theorem, the posterior distribution of θ for an individual who has made n contract renewals is

$$
\begin{aligned}
f(\theta \mid \alpha, \beta, n \text{ contract renewals}) &= \frac{S(n \mid \theta) f(\theta \mid \alpha, \beta)}{S(n \mid \alpha, \beta)} \\
&= \frac{\theta^{\alpha-1}(1-\theta)^{\beta+n-1}}{B(\alpha, \beta+n)} \,.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&DERL(d \mid \alpha, \beta, n \text{ contract renewals}) \\
&\qquad = \int_0^1 \frac{(1-\theta)}{d+\theta} \frac{\theta^{\alpha-1}(1-\theta)^{\beta+n-1}}{B(\alpha, \beta+n)} \, d\theta
\end{aligned}
$$

letting $s = 1 - \theta$

$$= \frac{1}{B(\alpha, \beta + n)(1 + d)} \int_0^1 s^{\beta+n}(1 - s)^{\alpha-1}\big(1 - \big(\tfrac{1}{1+d}\big)s\big)^{-1} ds$$
$$= \frac{\beta + n}{(\alpha + \beta + n)(1 + d)} \, {}_2F_1\big(1, \beta + n + 1; \alpha + \beta + n + 1; \tfrac{1}{1+d}\big)\,. \tag{B10}$$

(We note that this expression for DERL collapses to that given in (B4) when $n = 0$.)

Observation

While these different expressions for *DEL*/*DERL* may appear rather overwhelming, they demonstrate the need to be quite specific about the exact point in time at which we are computing an estimate of a customer's discounted expected (residual) lifetime value (as noted in Figure 1). Furthermore, it is very important that any calculations of *residual* lifetime value take into consideration a customer's tenure when our underlying model of contract duration assumes heterogeneity in individual retention rates.

2. An Alternative Derivation of the Key Expression for DERL

An alternative derivation of the paper's key expression for DERL follows from substituting the expression for the sBG survivor function in the general expression given in (B5):

$$DERL(d \,|\, \alpha, \beta, \text{active for } n \text{ periods}) = \sum_{t=n}^{\infty} \frac{B(\alpha, \beta + t)}{B(\alpha, \beta + n - 1)} \Big(\frac{1}{1 + d}\Big)^{t-n}$$

which, letting $s = t - n$ and expressing the beta functions in terms of gamma functions,

$$= \frac{\Gamma(\alpha + \beta + n - 1)}{\Gamma(\beta + n - 1)} \sum_{s=0}^{\infty} \frac{\Gamma(\beta + n + s)}{\Gamma(\alpha + \beta + n + s)} \Big(\frac{1}{1 + d}\Big)^{s}$$

and recalling that $\Gamma(1 + s) = s!$,

$$= \frac{\Gamma(\alpha + \beta + n - 1)}{\Gamma(\beta + n - 1)} \sum_{s=0}^{\infty} \frac{\Gamma(1 + s)\Gamma(\beta + n + s)}{\Gamma(\alpha + \beta + n + s)} \Big(\frac{1}{1 + d}\Big)^{s} \frac{1}{s!}\,.$$

Noting that the infinite sum is of the form associated with the primary definition of the Gaussian hypergeometric function (Wolfram Research 2001b), it is easy to show that this equals the expression given in (B8).

3. Deriving the Expression for Retention Elasticity

Recall that when individual survival times are distributed shifted-geometric with parameter θ, the associated retention elasticity (interpreted as the percentage increase in the customer's expected residual lifetime value for a given percentage increase in their underlying retention rate) is given by

$$\epsilon_{\text{ret}}(d \,|\, \theta, \text{active for } n \text{ periods}) = \frac{1 + d}{\theta + d}\,. \tag{B11}$$

Since θ is unobserved, we need to take the expectation of (B11) over the posterior distribution of θ for an individual who is still a customer in period n, (B7). It follows that

$$\epsilon_{\text{ret}}(d \mid \alpha, \beta, \text{active for } n \text{ periods}) = \int_0^1 \left(\frac{1+d}{\theta+d}\right) \frac{\theta^{\alpha-1}(1-\theta)^{\beta+n-2}}{B(\alpha, \beta+n-1)}\, d\theta$$

letting $s = 1 - \theta$

$$\begin{aligned} &= \frac{1}{B(\alpha, \beta+n-1)} \int_0^1 s^{\beta+n-2}(1-s)^{\alpha-1}\big(1 - \big(\tfrac{1}{1+d}\big)s\big)^{-1} ds \\ &= {}_2F_1\big(1, \beta+n-1; \alpha+\beta+n-1; \tfrac{1}{1+d}\big)\,. \end{aligned} \tag{B12}$$

This is the retention elasticity for a randomly chosen individual calculated for case (iii) in Section 1 of this note. The expression for the retention elasticity will be different for the other three cases.

Note that the *retention* elasticity should not be confused with the *churn* elasticity. Recall that the expected residual lifetime value of a customer acquired in period 1 who is still active in period n is $\bar{v}DERL(d \mid \text{active for } n \text{ periods})$. When individual survival times are distributed shifted-geometric with parameter θ, the associated *churn* elasticity (interpreted as the percentage increase in the customer's residual lifetime value for a given percentage increase in their underlying churn rate) is given by

$$\begin{aligned} \epsilon_{\text{churn}}(d \mid \theta, \text{active for } n \text{ periods}) &= \frac{\partial DERL(d \mid \theta, \text{active for } n \text{ periods})}{\partial \theta} \\ &\quad \times \frac{\theta}{DERL(d \mid \theta, \text{active for } n \text{ periods})} \\ &= -\frac{\theta}{1-\theta}\left(\frac{1+d}{\theta+d}\right). \end{aligned} \tag{B13}$$

Taking the expectation of this over the posterior distribution of θ gives us

$$\epsilon_{\text{churn}}(d \mid \alpha, \beta, \text{active for } n \text{ periods}) = -\left(\frac{\alpha}{\beta+n-2}\right) {}_2F_1\big(1, \beta+n-2; \alpha+\beta+n-1; \tfrac{1}{1+d}\big)\,. \tag{B14}$$

4. Understanding the Underestimation of Expected Residual Value

The following analysis expands on the paper's examination of the effects of failing to account for (aggregate) retention-rate dynamics when estimating the expected residual value of the customer base.

For a fine grid of points in the (μ, ϕ) space, we compute the implied aggregate 2006/07 churn rate — see Figure 2. This figure illustrates the fact that the mean of the underlying churn probability distribution does not perfectly reflect the observed churn rates. (For instance, Cases 1 and 2 from the paper both have $\mu = 0.20$ but lead to observed 2006/07 churn rates of 0.19 and 0.08, respectively.) In general, the observed churn rate equals the underlying mean only in the limiting case of no heterogeneity (i.e,. $\phi = 0$). As heterogeneity increases (holding μ constant), the observed churn level will always decrease, although these changes are not especially dramatic, as seen in the figure.

Given these aggregate churn rates, our naïve estimate of the expected residual value of the customer base is obtained using (B6) and the corresponding number of active customers in 2007;

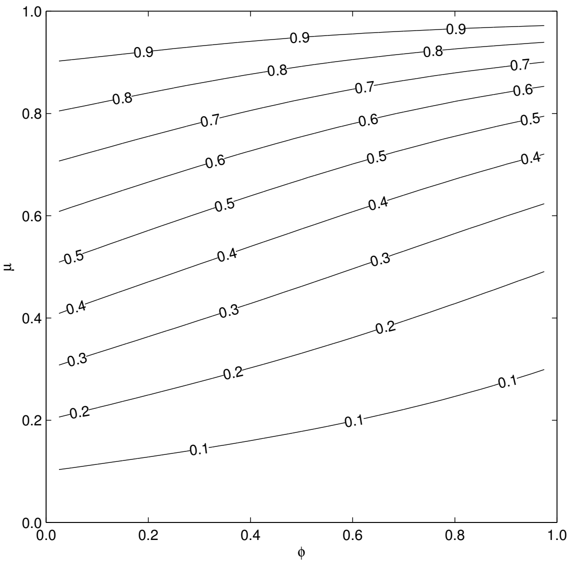


Figure 2: Aggregate 2006/07 churn rate as a function of μ and ϕ.

these valuations are plotted in Figure 3. The relatively flat contours show that large changes in heterogeneity do not necessarily lead to correspondingly large changes in the estimated value of the customer base. For instance, Cases 1 and 2 (as discussed in the paper) have very different degrees of polarization ($\phi = 0.05$ and $\phi = 0.75$, respectively), yet their naïve valuations are quite similar ($105,845 and $120,543, respectively). This small deviation largely reflects the (relatively flat) aggregate churn rate contours shown in Figure 2.

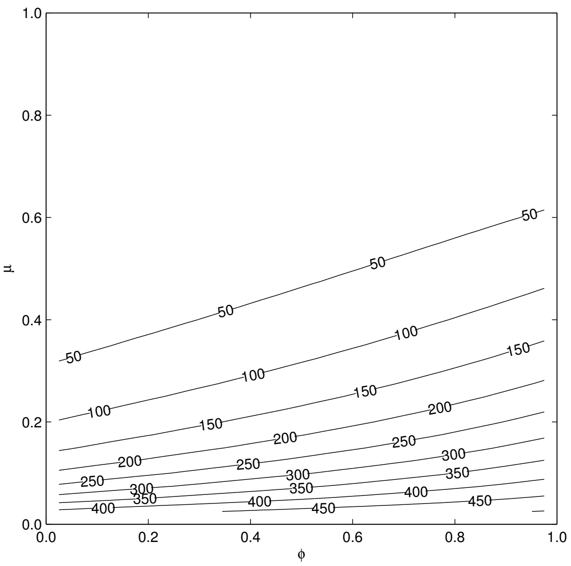


Figure 3: Expected residual lifetime value (in $000) of the customer base computed using the aggregate 2006/07 churn rate, as a function of μ and ϕ.

In contrast, Figure 4 shows the valuations derived using (B8), and the differences are immediately apparent. Holding μ fixed, there is a dramatic lift in valuation as heterogeneity increases. For example, the large difference in the valuations of Case 1 versus Case 2 ($220,488 and $375,437, respectively) reflects their very different degrees of polarization. This is a clear manifestation of the "sorting effect" of heterogeneity emphasized throughout the paper.

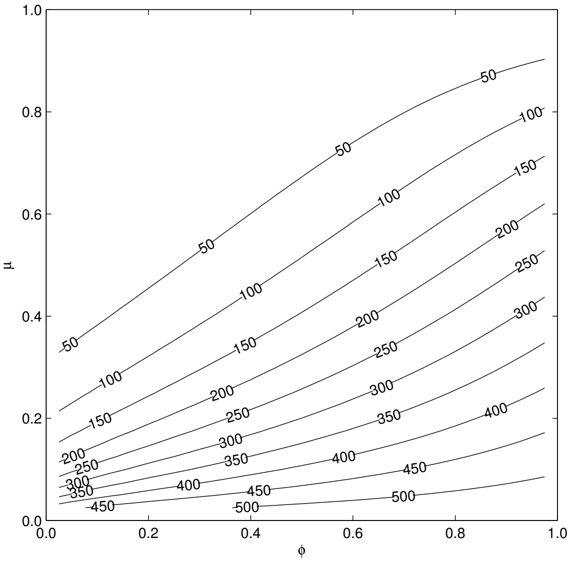


Figure 4: Expected residual lifetime value (in \$000) of the customer base computed using the sBG model, as a function of μ and ϕ.

The true impact of heterogeneity is better appreciated when we draw the comparisons across the two methods. For any point in the (μ, ϕ) space, the contour plot presented in the paper tells us the extent to which the naïve calculation method (from Figure 3) underestimates the true value of the customer base (from Figure 4).

5. Sensitivity Analysis

The analysis presented in the paper assumes that i) every cohort is identical in size, and ii) the mean churn rates across the cohorts are also identical. We now investigate the sensitivity of our conclusions to relaxing each of them.

Our initial analysis assumed that the same number of customers was acquired each period. To explore the impact of relaxing this assumption, we allow the size of each new cohort to increase over time ($10{,}000e^{0.15(j-1)}$ for the jth cohort, $j = 1, \ldots, 5$), decrease over time ($10{,}000e^{-0.15(j-1)}$), or remain constant ($10{,}000\ \forall\, j$). This means the 2007 cohort is approximately double (half) the size of the 2003 cohort in the increasing (decreasing) case.

Our initial analysis also assumed that the distribution of within-cohort churn rates was the same for each new cohort. In some situations we would expect new cohorts to have a churn-rate distribution with a lower mean. One example of this is where learning on the part of the firm means that it is better at targeting its acquisition activities to those potential customers whose needs are better met by the firm's offerings (and are therefore less prone to switching behavior). In other situations we would expect new cohorts to have a churn-rate distribution with a higher mean (e.g., as a market saturates and thus the pool of new potential customers is increasingly composed of people who have "churned" from other providers).

We examine the effect of nonstationarity in the mean churn rates *across* cohorts in the following manner. Let μ_j be the underlying mean churn probability for the jth cohort, $j =$

1, . . . , 5 (i.e., the mean of its beta distribution). For the case of between-cohort churn rates decreasing over time, we have $\mu_j = \mu(0.975)^{j-1}$, while for the case of increasing churn rates, we have $\mu_j = 1 - (1 - \mu)(0.975)^{j-1}$. (For the "constant" case considered in the paper, $\mu_j = \mu \, \forall \, j$.)

This results in nine "mean churn rate across cohorts" by "inflow of customers" scenarios. For each scenario, we compute the percentage underestimation of the expected residual value of the customer base as a function of μ and ϕ in the manner outlined in the previous section. The associated plots are presented in Figure 5. The middle plot (constant/constant) is exactly the same as that presented in the paper.)

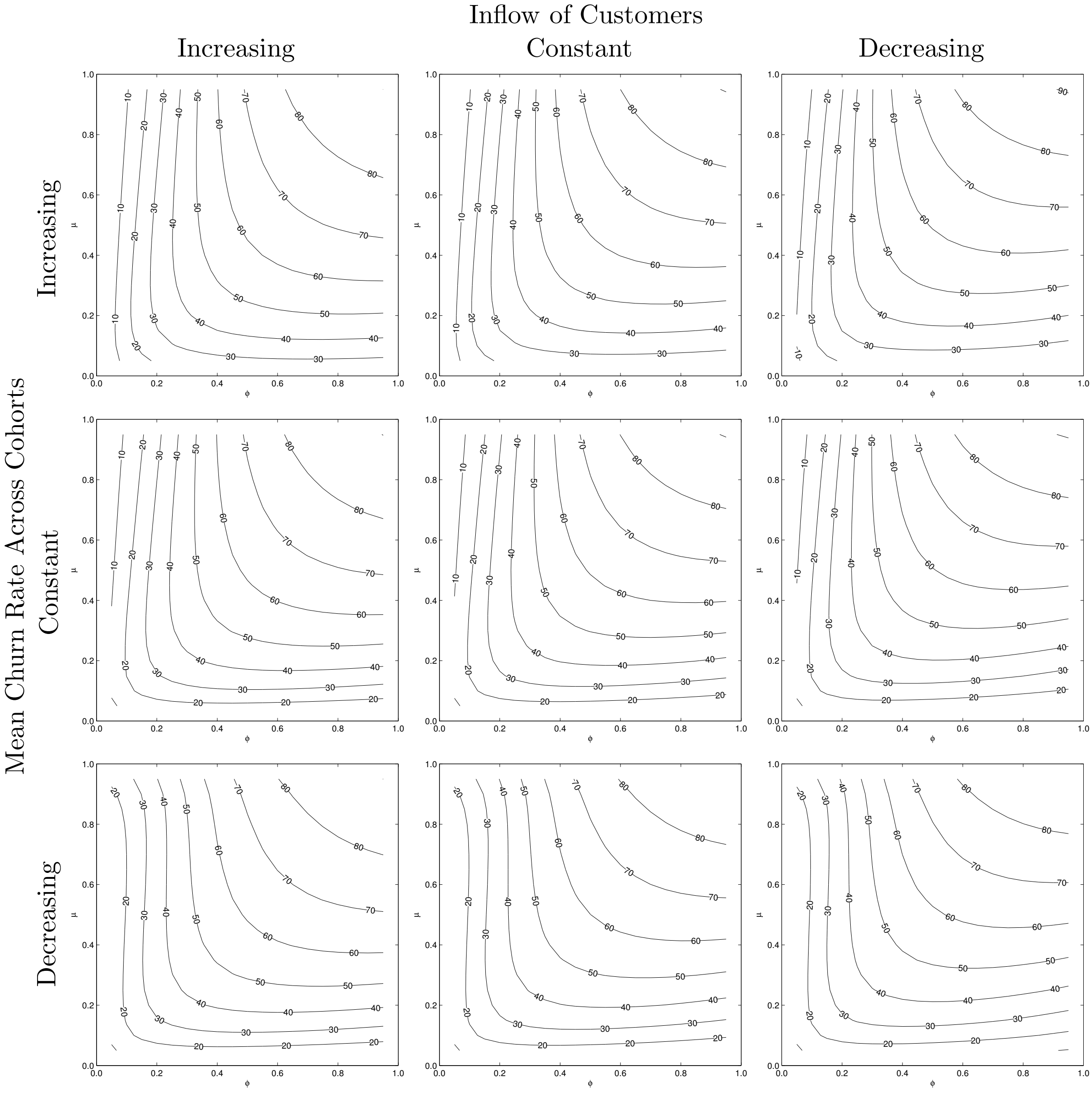


Figure 5: Percentage underestimation of the expected residual value of the customer base as a function of μ and ϕ under conditions of different customer inflow dynamics and different cross-cohort churn patterns.

What immediately jumps out of this analysis is that the basic shapes (and qualitative conclusions) associated with the constant/constant case are quite robust; the shape of the contours and their corresponding values are relatively insensitive to changes in the size of each new cohort of customers and the mean churn rate of each new cohort. For example, the percentage underestimation associated with the (0.20,0.20) point in the (μ, ϕ) space only varies from 31.5% (top left scenario) to 32.4% (bottom left scenario). Similarly, for the (0.40,0.40) point in the (μ, ϕ) space, the percentage underestimation varies from 50.9% (bottom right scenario) to 52.9% (top left scenario).

These extended analyses provide clear evidence that our main result (i.e., the systematic underestimation of the residual value of the customer-base when using an aggregate retention rate) is insensitive to varying types of cross-cohort differences. While these observations admittedly fall short of a formal proof that the bias will always exist (at least when the true, underlying behavior is governed by an sBG process), they demonstrate that this phenomenon has very broad applicability.

6. Reanalysis Using (Within-Cohort) Retention Rates for Periods 1 and 2

The analyses presented in the paper show the percentage underestimation of the expected residual value of the customer base and retention elasticities as a function of μ and ϕ, which are not quantities that most managers or analysts think about. On the other hand, retention rates are easy to comprehend.

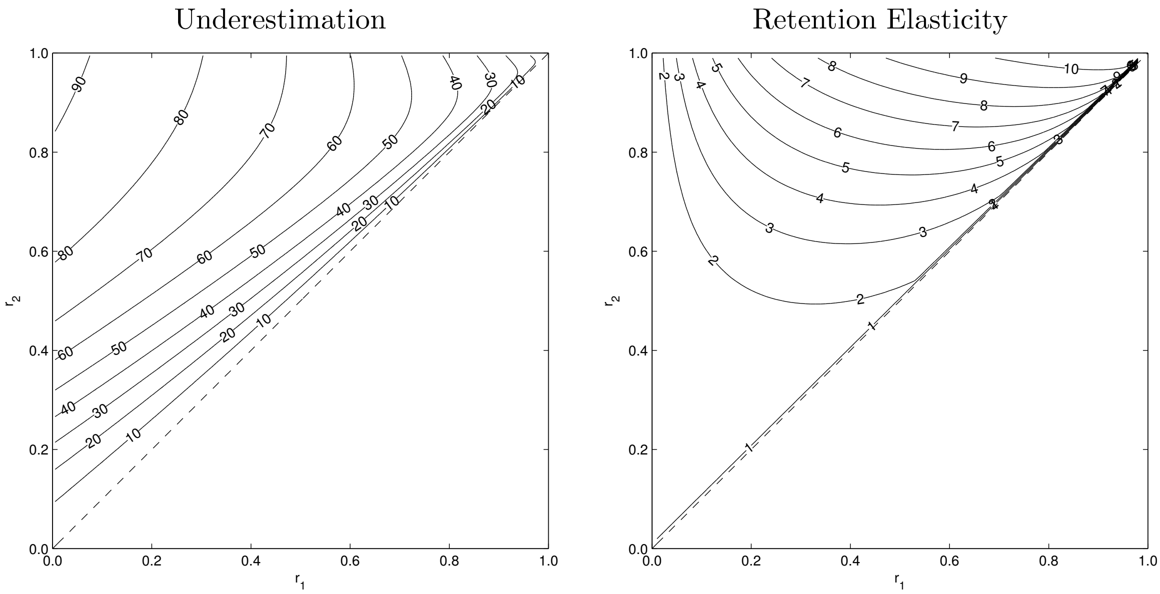
In a world where contract durations can be characterized by the sBG model, the within-cohort retention rates for the first two periods are $r_1 = \beta/(\alpha+\beta)$ and $r_2 = (\beta+1)/(\alpha+\beta+1)$. It follows that the two sBG model parameters can be expressed as functions of r_1 and r_2: $\alpha = (1-r_1)(1-r_2)/(r_2-r_1)$ and $\beta = r_1(1-r_2)/(r_2-r_1)$. For a fine grid of points in the (r_1, r_2) space, we determine the corresponding values of (α, β) and compute i) the percentage underestimation of the expected residual value of the customer base resulting from the use of a single aggregate retention rate (Figure 6a) and ii) the retention elasticity (Figure 6b). (In both figures, the space below the 45° line is empty since $r_2 > r_1$ under the sBG model.) These two figures are equivalent to the "underestimation" and elasticity plots presented in the paper; the difference is that expressing the quantities of interest as a function of the first-two retention rates makes the insights clearer to a broader population.

Looking at Figure 6a, we see when r_2 is very close to r_1, the error associated with calculating the expected residual value of the customer base using the naïve method is relatively small. For r_2 to be very close to r_1, it must at least be the case that $\beta \gg 1$; we would also expect $\alpha \gg 1$ so that $r_1 \not\approx 1$. This implies that there is little heterogeneity in θ across cohort members and therefore little error associated with using the naïve method to value a customer base. As the gap between r_1 and r_2 grows, so does the magnitude of the error.

Turning our attention to size of the retention elasticity (Figure 6b), we see that the retention elasticity is heavily determined by the period 1 retention rate. However, for a given value of r_1, the retention elasticity increases the greater the gap between the two retention rates.

7. Model "Validation"

Fader and Hardie (2007a) present a validation of the basic sBG model for a single cohort. We now explore how well the model captures the behavior of multiple cohorts so as to provide greater insight into what lies behind the valuation bias reported in the paper.



(a) (b)

Figure 6: Percentage underestimation of the expected residual value of the customer base and retention elasticity, as a function of (r_1, r_2).

The strongest test would be to have a dataset in which we track the entire history of all the customers acquired by a firm in, say, the first five years of its operation to the point in time where all of these customers have churned (which could be 10–20 years after their time of acquisition). This would enable us to compare the true (or realized) residual lifetime value of the customer base (from, say, year 6 onwards) with that derived using the sBG model and the naïve method. We do not have access to such a dataset, but we leverage a more limited real-world dataset to create and examine this kind of scenario.

In particular, we construct a multi-cohort dataset using a single-cohort's real survival data. The real survival data are those associated with the "Regular" segment analyzed in Fader and Hardie (2007a). That dataset summarized the behavior of those customers acquired in the first seven years of a hypothetical firm's operations, tracking their behavior over a subsequent six years. To be consistent with the paper, it is assumed that we acquire 10,000 new customers each year. The resulting dataset is reported in Table 1.

We fit the sBG model to this dataset, undertaking a "pooled" analysis in which we assume that the one set of sBG model parameters apply across all cohorts. (The associated model likelihood function is given in equation (3) of Fader and Hardie (2007b).) The maximum likelihood estimates of the model parameters are $\hat{\alpha} = 0.749$ and $\hat{\beta} = 1.269$.

Given the aggregate (Year 7) retention rate ((33,580 − 10,000(/30,959 = 0.762), the naïve prediction of the expected number of customers in the xth holdout year is 33,580 × $(0.762)^x$. (Such a projection underlies the naïve/aggregate valuation calculations reported in the paper.) These numbers, along with the "actual" numbers and the sBG predictions for the holdout period, are reported in Table 2 and plotted in Figure 7. It is very clear that the forecast of the customer-base size generated using the aggregate retention rate is bad, and this is what drives the underestimation of the residual value of the customer-base discussed in the paper.

Year												
1	2	3	4	5	6	7	8	9	10	11	12	13
10,000	6,309	4,679	3,816	3,264	2,891	2,621	2,407	2,230	2,072	1,944	1,834	1,733
	10,000	6,309	4,679	3,816	3,264	2,891	2,621	2,407	2,230	2,072	1,944	1,834
		10,000	6,309	4,679	3,816	3,264	2,891	2,621	2,407	2,230	2,072	1,944
			10,000	6,309	4,679	3,816	3,264	2,891	2,621	2,407	2,230	2,072
				10,000	6,309	4,679	3,816	3,264	2,891	2,621	2,407	2,230
					10,000	6,309	4,679	3,816	3,264	2,891	2,621	2,407
						10,000	6,309	4,679	3,816	3,264	2,891	2,621
10,000	16,309	20,988	24,804	28,068	30,959	33,580	25,987	21,908	19,301	17,429	15,999	14,841

Table 1: Number of customers by cohort for the hypothetical firm, created using actual single-cohort data.

Using the numbers presented in Table 2, we can compute the discounted residual value of the customer base over a six-year time horizon. Assuming a 10% discount rate and an average contract value of \$1 per year, this would imply an "actual" discounted residual value of \$95,092. The associated numbers computed using the sBG model and the aggregate retention rate are \$93,316 and \$73,988, which underestimate the "true" value by -2% and -22% respectively. (Computing the discounted expected residual value over an infinite time horizon for the sBG model, we find that the error associated with the use of an aggregate retention rate is -44%, which is exactly the same as that reported in Figure 2 of the paper for $\mu = 0.749/(0.749+1.269) = 0.371$ and $\phi = 1/(0.749 + 1.269 + 1) = 0.331$.)

	Year					
	8	9	10	11	12	13
Actual	25,987	21,908	19,301	17,429	15,999	14,841
sBG	25,877	21,716	18,957	16,945	15,393	14,149
Naïve	25,576	19,480	14,837	11,301	8,607	6,556

Table 2: Actual vs. predicted (residual) size of the customer base over the six-year holdout period.

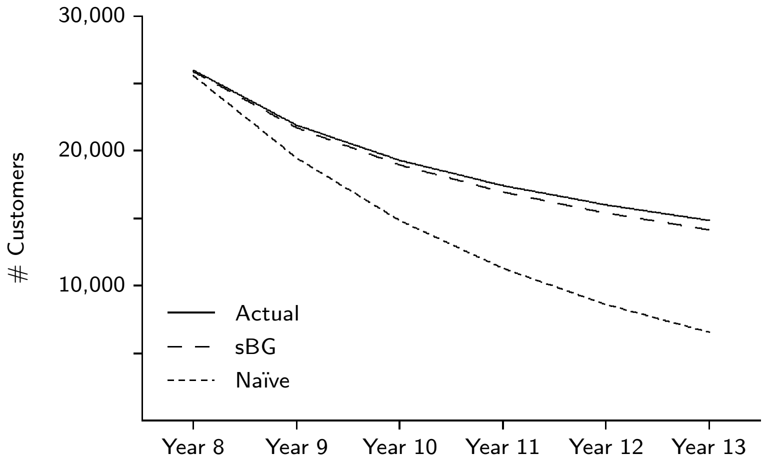


Figure 7: Actual vs. predicted (residual) size of the customer base over the six-year holdout period.

References

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.

Wolfram Research (2001a), The Wolfram Functions Site, `http://functions.wolfram.com/07.23.07.0001.01`

Wolfram Research (2001b), The Wolfram Functions Site, `http://functions.wolfram.com/07.23.02.0001.01`