MARKETING SCIENCE
Vol. 29, No. 1, January–February 2010, pp. 85–93
ISSN 0732-2399 | EISSN 1526-548X | 10 | 2901 | 0085

informs®

DOI 10.1287/mksc.1080.0482

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

Peter S. Fader
Department of Marketing, The Wharton School of the University of Pennsylvania, Philadelphia, Pennsylvania 19104, faderp@wharton.upenn.edu

Bruce G. S. Hardie
London Business School, Regent's Park, London NW1 4SA, United Kingdom, bhardie@london.edu

The past few years have seen increasing interest in taking the notion of customer lifetime value (CLV) and extending it to value a customer base (with subsequent links to corporate valuation). The application of standard textbook discussions of CLV leads to calculations based on a single retention rate. However, at the cohort level, retention rates typically increase over time. It has been suggested that these observed dynamics are due, in large part, to a sorting effect in a heterogeneous population. We show that failing to recognize these dynamics yields a downward-biased estimate of the residual value of the customer base (compared to an aggregate analysis that ignores these dynamics). We also explore the implications of failing to account for retention dynamics when computing retention elasticities and find that the frequently reported values underestimate the true effect of increases in underlying retention rates in a heterogeneous world.

Key words: customer lifetime value; shifted-beta-geometric; retention elasticity
History: Received: June 15, 2006; accepted: December 5, 2008; processed by Scott Neslin. Published online in *Articles in Advance* April 8, 2009.

1. Introduction

The past few years have seen increasing interest in taking the notion of customer lifetime value (CLV) and extending it to value a customer base, with subsequent linkages to the value of a firm (Bauer et al. 2003, Gupta et al. 2004). As we undertake such valuation exercises, it is essential to ensure that our underlying calculations are correct, i.e., fully reflective of the factors that underlie the observed behavior.

Consider a hypothetical contractual business setting where the firm's customer base is described as in Table 1.[1] The firm acquired 10,000 customers in 2003, 6,334 of whom renewed their contracts at the beginning of 2004. By 2007, 2,604 members of this cohort were still customers of the firm. Another cohort of 10,000 customers was acquired in 2004; by 2007, 3,264 of them were still customers of the firm. For this hypothetical firm, the average net cashflow per customer is \$100/year, which is "booked" at the beginning of the contract period.

Suppose it is December 31, 2007, and we ask the question: "What is the expected residual value of the 26,569 customers currently on the firm's books?" (For the sake of simplicity, let us assume that each contract is annual, starting on January 1 and expires at 11:59 P.M. on December 31. Furthermore, we assume a 10% discount rate.)

Any student who has been taught the basic principles of customer lifetime value (e.g., Kotler and Keller 2009, Lehmann and Winer 2005, Ofek 2002) would likely take the following approach: First, he would calculate the *aggregate* retention rate as

$$(2{,}604+3{,}264+4{,}367+6{,}334)/(3{,}264+4{,}367+6{,}334+10{,}000)=0.691.^{2}$$

[1] As noted in Fader and Hardie (2009), a defining characteristic of a contractual (or subscription) business setting is that the loss of a customer is observed. For example, the customer has to contact the firm to cancel her cable TV contract; similarly, a magazine publisher can observe that a subscriber has not renewed his annual subscription. This is in contrast to a noncontractual setting, a defining characteristic of which is that the loss of a customer is not observed by the firm.

[2] Farris et al. (2006, p. 132) define the retention rate as "the ratio of the number of retained customers to the number at risk." As noted in PricewaterhouseCoopers (2002), the two standard ways of determining the "number at risk" are (i) the number of customers at the beginning of the period of interest and (ii) the average of the opening and closing balance of customers for the period of interest. The first approach is clearly the more appropriate choice given the discrete-time nature of the setting we are examining.

Table 1 Number of Active Customers Each Year by Year-of-Acquisition Cohort

2003	2004	2005	2006	2007
10,000	6,334	4,367	3,264	2,604
	10,000	6,334	4,367	3,264
		10,000	6,334	4,367
			10,000	6,334
				10,000
10,000	16,334	20,701	23,965	26,569

Using the logic of standard textbook expressions for CLV, he would then calculate the expected residual value of this group of customers as

$$26{,}569 \times \sum_{t=1}^{\infty} \$100 \times 0.691^t/(1+0.1)^{t-1} = \$4{,}945{,}049.$$

Most professors would pat this student on the back, encouraged that he had been able to apply this important course concept. (This is consistent with the methods used by Bauer et al. 2003, Gupta et al. 2004, and Wiesel et al. 2008, among others.) The only problem is that this answer is wrong; as we shall soon see, this calculation underestimates the value of the customer base by 38%!

So what is wrong with this approach? Table 2 reports the annual *cohort-level* retention rates as computed off the numbers presented in Table 1 (e.g., $6{,}334/10{,}000 = 0.633$). We note that although the aggregate numbers (bottom row) are fairly constant, the cohort-level numbers are increasing quite substantially over time. (Note that the growth pattern in this stylized example is the same for each cohort.)

Clearly, any attempt to compute the expected residual value of this customer base needs to recognize the *intercohort differences* at any point in time (e.g., the different 2006/2007 retention rates for each cohort). Furthermore, it is necessary to *project* each cohort's (increasing) retention numbers beyond the set of observed retention rates (i.e., to 2007/2008, 2008/2009, and beyond).

We therefore need a model of relationship duration that can be used to project the growth of the retention rates beyond 2007, from which we can then compute a customer's expected residual lifetime value (and therefore the expected residual value of the whole customer base). In developing this model, we need to provide a valid explanation for the phenomenon of increasing retention rates at the cohort level. In §2, we present an exploratory analysis in which this phenomenon is simply an artifact of cross-sectional heterogeneity in individual retention probabilities (e.g., Gupta and Lehmann 2005, pp. 29–31); this is examined more formally and systematically in §3 using the shifted-beta-geometric model of contract duration (Fader and Hardie 2007a). We then explore the implications of failing to account for retention dynamics on the true magnitude of "retention elasticities." We conclude with a discussion of several issues that arise from this work.

Table 2 Annual Retention Rates by Cohort

2003	2004	2005	2006	2007
—	0.633	0.689	0.747	0.798
	—	0.633	0.689	0.747
		—	0.633	0.689
			—	0.633
				—
—	0.633	0.655	0.675	0.691

2. An Initial Exploration

Suppose we track a cohort of 10,000 customers, consisting of two underlying segments: Segment 1 comprises one-third of the customers, each with an unobserved (and unobservable) time-invariant annual retention probability of 0.9, and Segment 2 comprises two-thirds of the customers, each with a time-invariant annual retention probability of 0.5.

Looking at Table 3, we see that, as time progresses, the number of people in the pool of active customers with a (constant) high-churn probability (Segment 2) declines at a faster rate than the number of people with a (constant) low-churn probability (Segment 1). As a result, the cohort-level retention rate increases until members of Segment 2 effectively disappear, and therefore, it levels off to the rate of 0.9 associated with Segment 1 alone.

We see that the cohort-level retention rates are exactly the same as those reported for the cohort of customers acquired in 2003 (as presented in Table 2), as well as the subsequent cohorts. The key observation is that *the observed phenomenon of increasing cohort-level retention rates is, in this case, purely an artifact of cross-sectional heterogeneity*.

This observation should not come as a surprise to any serious student of customer retention (e.g., Gupta and Lehmann 2005, pp. 29–31) as it is a well-known

Table 3 Changes in the Number of Active Customers Over Time

	No. of active customers			Retention rate		
Year	Seg. 1	Seg. 2	Total	Seg. 1	Seg. 2	Agg.
1	3,333	6,667	10,000			
2	3,000	3,334	6,334	0.900	0.500	0.633
3	2,700	1,667	4,367	0.900	0.500	0.689
4	2,430	834	3,264	0.900	0.500	0.747
5	2,187	417	2,604	0.900	0.500	0.798

"ruse of heterogeneity" (Vaupel and Yashin 1985).[3] Yet it is generally ignored in subsequent calculations of customer lifetime value and ignored in any customer-base valuation exercise. As we shall see, the impact of ignoring the phenomenon of increasing cohort-level retention rates can be quite dramatic.

Let us consider the expected residual lifetime value of one of the 2,604 members of the 2003 cohort who is still active in 2007. If this person belongs to Segment 1, her expected residual lifetime value, $E(RLV)$, is given by

$$E[RLV(r=0.9, d=10\%)] = \sum_{t=1}^{\infty} \$100 \times \frac{0.9^t}{(1+0.1)^{t-1}} = \$495,$$

whereas if she belongs to Segment 2,

$$E[RLV(r=0.5, d=10\%)] = \sum_{t=1}^{\infty} \$100 \times \frac{0.5^t}{(1+0.1)^{t-1}} = \$92.$$

However, the segment to which this individual belongs is unobserved.

According to Bayes' theorem, the probability that a member of the first cohort, still active in 2007, belongs to Segment 1 is

$$\frac{P(\text{renewed contract four times} \mid \text{Segment 1}) \times P(\text{Segment 1})}{P(\text{renewed contract four times})}$$
$$= \frac{0.9^4 \times 0.333}{0.9^4 \times 0.333 + 0.5^4 \times 0.667} = 0.84.$$

This means that the expected residual lifetime value for a randomly chosen member of the 2003 cohort who is still active in 2007 is $0.84 \times \$495 + (1-0.84) \times \$92 = \$430$.

The corresponding numbers for the 2004–2007 cohorts are \$392, \$341, \$283, and \$226, respectively. Therefore, the expected residual value of the firm's customer base is $2{,}604 \times \$430 + 3{,}264 \times \$392 + 4{,}367 \times \$341 + 6{,}334 \times \$283 + 10{,}000 \times \$226 = \$7{,}940{,}992$. Thus the naïve calculation based on the aggregate retention rate (\$4,945,049) underestimates the expected residual value of the customer base by 38%.

This informal analysis suggests that any calculation of the residual value of a customer base performed using a single aggregate retention rate will underestimate the true value of the customer base whenever the cohort-level retention rates are increasing over time. As we typically observe increasing cohort-level retention rates (e.g., Fader and Hardie 2007a, Kumar and Reinartz 2006, Reichheld 1996, Schweidel et al. 2008), it is crucial that our calculations incorporate this phenomenon.

[3] This insight about heterogeneity is generally seen as a purely static problem, but it is compounded in the customer-base valuation context.

3. A Systematic Examination of the Effects of Heterogeneity

We have demonstrated that failing to account for heterogeneity in cohort-level retention rates can lead to an underestimation of the residual lifetime value of a customer (and therefore the value of the customer base). We now wish to undertake a more systematic study of this problem, still assuming that the dynamics in cohort-level retention rates stem entirely from cross-sectional heterogeneity. We will undertake this analysis using the shifted-beta-geometric (sBG) model for contract duration proposed and tested by Fader and Hardie (2007a).

The sBG model is based on the following two assumptions:

(1) An individual remains a customer of the firm with constant retention probability $1-\theta$; this implies that the duration of the customer's relationship with the firm is characterized by the shifted-geometric distribution with survivor function:

$$S(t \mid \theta) = (1-\theta)^t, \quad t = 1, 2, 3, \ldots.$$

(2) Heterogeneity in θ is captured by a beta distribution with probability density function:

$$f(\theta \mid \alpha, \beta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)}.$$

It follows that the probability that a randomly chosen customer's tenure is greater than t periods (i.e., the survivor function) is

$$S(t \mid \alpha, \beta) = \frac{B(\alpha, \beta+t)}{B(\alpha, \beta)}, \quad t = 1, 2, \ldots. \tag{1}$$

The cohort-level retention rate for period t (r_t) is defined as the proportion of customers active at the end of period $t-1$ who are still active during t, which is simply the ratio of sequential values of the survivor function, $r_t = S(t)/S(t-1)$. Given (1), it follows that the retention rate associated with the sBG model is

$$r_t = \frac{\beta + t - 1}{\alpha + \beta + t - 1}. \tag{2}$$

Fader and Hardie (2007a) note that under this model, the retention rate is always an increasing function of time, even though the underlying (unobserved) individual-level retention probabilities are constant. Because there are no underlying time dynamics at the level of the individual customer, the observed phenomenon of retention rates increasing over time is simply a sorting effect in a heterogeneous population (i.e., the high-churn customers drop out early in the observation period, with the remaining customers having lower churn probabilities).

Fader and Hardie (2007a) stop with the development and empirical testing of the sBG model. Now we take it to the next step by showing how to generate estimates of CLV given the model's parameters. This is, in essence, equivalent to the contribution of Fader et al. (2005), who performed a similar task for the Pareto/NBD model (Schmittlein et al. 1987) in a noncontractual setting.

Standing at the end of a customer's nth contract period, just prior to the point in time at which she makes her contract renewal decision, the expected residual lifetime value of a customer acquired at the beginning of period 1 is

$$E[RLV(d \mid \text{active for } n \text{ periods})] = \sum_{t=n}^{\infty} E[v(t)] \frac{S(t \mid t > n-1)}{(1+d)^{t-n}},$$

where $E[v(t)]$ is the expected net cashflow in period t, assumed to be "booked" at the beginning of the contract period. (The summation index runs from $t = n$, not $t = n+1$, since a customer active in period n has made only $n-1$ contract renewals and the survivor function is $P(T > t)$.)

Assuming a constant expected net cashflow per period (i.e., $E[v(t)] = \bar{v}$), we can factor it out of the calculation:

$$E[RLV(d \mid \text{active for } n \text{ periods})] = \bar{v} DERL(d \mid \text{active for } n \text{ periods}), \quad (3)$$

where the *discounted expected residual lifetime* is defined as

$$DERL(d \mid \text{active for } n \text{ periods}) = \sum_{t=n}^{\infty} \frac{S(t \mid t > n-1)}{(1+d)^{t-n}} = \sum_{t=n}^{\infty} \frac{S(t)}{S(n-1)} \left(\frac{1}{1+d}\right)^{t-n}. \quad (4)$$

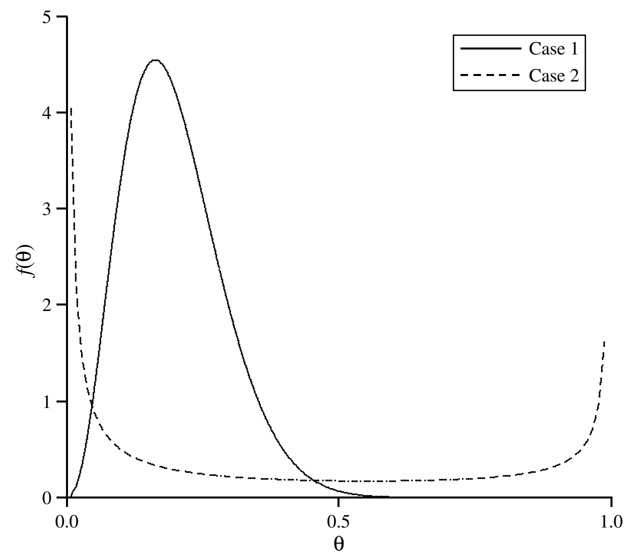
When individual survival times are distributed shifted-geometric with parameter θ, we have

$$DERL(d \mid \theta, \text{active for } n \text{ periods}) = \frac{(1-\theta)(1+d)}{d+\theta}. \quad (5)$$

However, θ is unobserved; we therefore need to take the expectation of (5) over the distribution of θ. But when $n > 1$, we cannot use 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. Recalling Bayes' theorem, it follows that

$$\begin{aligned} &DERL(d \mid \alpha, \beta, \text{active for } n \text{ periods}) \\ &= \int_0^1 \frac{(1-\theta)(1+d)}{d+\theta} \frac{S(n-1 \mid \theta) f(\theta \mid \alpha, \beta)}{S(n-1 \mid \alpha, \beta)} \, d\theta \\ &= \left(\frac{\beta+n-1}{\alpha+\beta+n-1}\right) {}_2F_1\left(1, \beta+n; \alpha+\beta+n; \frac{1}{1+d}\right), \end{aligned} \quad (6)$$

Figure 1 Shape of the Beta Distribution for Cases 1 and 2



where ${}_2F_1(\cdot)$ denotes the Gaussian hypergeometric function.[4] (See the appendix for details of how to compute this quantity when using a modeling environment that does not have a built-in routine for evaluating the Gaussian hypergeometric function.)

Equation (6) is the focal expression for this paper and is derived here for the first time. Multiplied by $\bar{v}$, it is the natural extension of the standard textbook CLV formula that recognizes (i) heterogeneity in churn propensities and (ii) the need to condition on the length of the customer's relationship with the firm when computing her residual value. It also makes explicit the fact that we are creating an *expectation*. This acknowledges the fact that CLV calculations are inherently probabilistic, not deterministic, as implicit in most discussions of the topic.

Suppose a firm acquires a new cohort of 10,000 customers each year from 2003 to 2007. For different values of α and β, we will compare the expected residual value of the customer base at the end of 2007 as calculated using the naïve approach (a constant aggregate retention rate) with the true value that takes the (heterogeneity-induced) dynamics in the retention rates into account. We will assume that $d = 0.1$ (i.e., a 10% discount rate) and $\bar{v} = \$1$.

We start by considering two scenarios: Case 1 ($\alpha = 3.80$ and $\beta = 15.20$) and Case 2 ($\alpha = 0.067$ and $\beta = 0.267$). The shapes of the corresponding beta distributions are illustrated in Figure 1. In Case 1, the distribution of θ is relatively homogeneous with an interior mode. In Case 2, there is quite a bit of

[4] This expression for DERL is based on specific assumptions about the exact point in time at which the residual lifetime value is being evaluated. See the electronic companion, available as part of the online version that can be found at http://mktsci.pubs.informs.org, for details of the DERL expressions associated with alternative assumptions.

Table 4 Number of Active Customers Each Year by Year-of-Acquisition

2003	2004	2005	2006	2007	*DERL*
			Case 1		
10,000	8,000	6,480	5,307	4,391	3.84
	10,000	8,000	6,480	5,307	3.72
		10,000	8,000	6,480	3.59
			10,000	8,000	3.45
				10,000	3.31
10,000	18,000	24,480	29,787	34,178	
			Case 2		
10,000	8,000	7,600	7,383	7,235	10.19
	10,000	8,000	7,600	7,383	10.06
		10,000	8,000	7,600	9.86
			10,000	8,000	9.46
				10,000	7.68
10,000	18,000	25,600	32,983	40,218	

heterogeneity in the distribution of θ; this U-shaped distribution indicates that some of the acquired customers have a high value of θ (i.e., a low retention rate), while a larger number of customers have a small value of θ. For both cases, $E(\theta) = 0.20$.

Table 4 reports the number of active customers each year by year-of-acquisition cohort for both cases. (These numbers are computed using the sBG survivor function (1) and the customer acquisition numbers given above.) For Case 1, the aggregate 2006/2007 retention rate is $24{,}178/29{,}787 = 0.81$; for Case 2, the aggregate 2006/2007 retention rate is $30{,}218/32{,}983 = 0.92$. It follows that the naïve calculation of the expected residual value of the Case 1 customer base is $34{,}178 \times \$1 \times \sum_{t=1}^{\infty} 0.81^t/(1+0.1)^{t-1} = \$105{,}845$; for Case 2, we have $40{,}218 \times \$1 \times \sum_{t=1}^{\infty} 0.92^t/(1+0.1)^{t-1} = \$220{,}488$.

The rightmost column of each subtable in Table 4 reports the discounted expected residual lifetime (given a 10% discount rate), as computed using (6), by year of acquisition. It follows that for Case 1, the correct value of the expected residual lifetime value of the customer base is $4{,}391 \times 3.84 + 5{,}307 \times 3.72 + 6{,}480 \times 3.59 + 8{,}000 \times 3.45 + 10{,}000 \times 3.31 = \$120{,}543$. Thus the naïve calculation underestimates the true value by 12%. The equivalent correct value for Case 2 is \$375,437, which means the naïve calculation underestimates the true valuation by 41%.

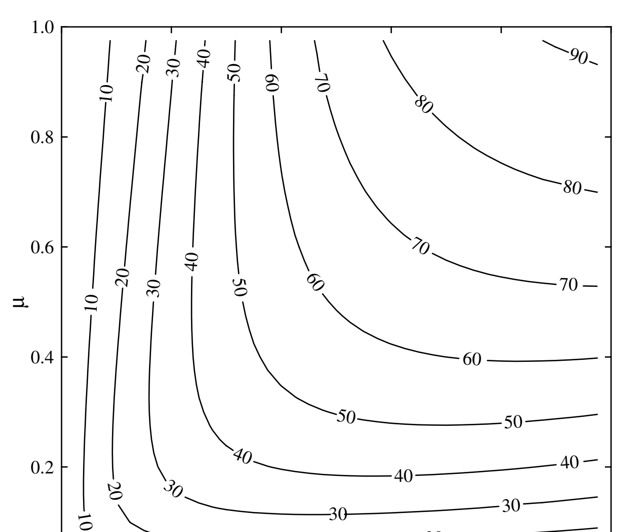
Referring back to Figure 1, the (relatively) high degree of homogeneity in the distribution of θ for Case 1 means that the error associated with calculating the expected residual value of the customer base using the aggregate 2006/2007 retention rate is relatively small. However, for Case 2, where the distribution of θ is U-shaped, the error is much greater; a large number of customers have small values of θ and are therefore expected to have long relationships with the firm. The aggregate retention number cannot take this into account. Failure to take this heterogeneity into consideration will necessarily result in a highly biased estimate of the residual value of the customer base.

This effect of heterogeneity is also reflected in the $DERL(d \mid \text{active for } n \text{ periods})$ numbers reported in Table 4. For Case 1, we see that there is very little variability in the numbers across cohorts, with the expected residual value of a customer who has made four renewals only being 16% higher than that of a customer who has made no renewals. On the other hand, in Case 2, there is a 33% difference.

Of course, these are but two possible sets of values for α and β. As we seek to undertake a more systematic investigation of the effects of heterogeneity on valuation error, we note that the parameters of the beta distribution can be characterized in terms of the mean $\mu = E(\theta) = \alpha/(\alpha+\beta)$ and the polarization index $\phi = 1/(\alpha+\beta+1)$. The logic behind the polarization index is as follows: as $\alpha, \beta \to 0$ (thus $\phi \to 1$), the values of θ are concentrated near $\theta = 0$ and $\theta = 1$, and we can think of the values of θ as being very different, or "highly polarized." As $\alpha, \beta \to \infty$ (thus $\phi \to 0$), the beta distribution becomes a spike at its mean; there is no "polarization" in the values of θ. (For Case 1 above, $\mu = 0.20$ and $\phi = 0.05$; for Case 2, $\mu = 0.20$ and $\phi = 0.75$.) For a fine grid of points in the (μ, ϕ) space, we compute the corresponding values of (α, β) and then compute the expected residual value of the customer base using both the naïve calculation and the model-based methods.

Figure 2 presents a contour plot of the percentage by which the naïve calculation method underestimates the true value of the customer base (as implied by the sBG model). For instance, as noted earlier, Case 1 has

Figure 2 Percentage Underestimation of the Expected Residual Value of the Customer Base, as a Function of μ and ϕ



an underestimation of 12% while Case 2 has an underestimation of 41%. As expected, we see that when the individual retention rates are relatively homogeneous (low values of ϕ), the error associated with calculating the expected residual value of the customer base using the naïve method is small. In the limit as $\phi \to 0$, there is no heterogeneity in θ and the sBG model tends to a pure shifted-geometric model (i.e., everyone has the same retention rate), and therefore there is no error associated with using the naïve method to value a customer base. However, as the level of heterogeneity increases (moving to the right of the figure), the magnitude of the error quickly becomes unacceptable. In such cases, any attempt to estimate the expected residual value of a customer base should be based on a formal model of contract duration. The only possible case where it is still acceptable to use the naïve method is when the mean underlying retention probability is *very* high (i.e., μ is very low).

It is worth noting that we would rarely find a business setting with very high churn probabilities (say, $\mu > 0.4$). As a result, the extreme levels of underestimation of residual value (70+%) will rarely be observed. However, even levels of 20% are very significant to the financial community, and it is easy to see such discrepancies arising for very realistic values of μ and ϕ. For instance, Fader and Hardie (2007a) fit the sBG model to two actual (and typical) data sets; the estimated values of the mean and polarization index are $\hat{\mu} = 0.15$, $\hat{\phi} = 0.18$ and $\hat{\mu} = 0.37$, $\hat{\phi} = 0.35$. Applying these patterns of customer contract duration to the customer acquisition patterns used in the above analysis (i.e., 10,000 new customers each year), the naïve calculation of the residual value of the firm's customer base would underestimate the value computed using the sBG model by 28% and 48%, respectively.

The preceding analysis assumes that (i) every cohort is identical in size and (ii) the mean churn rates across the cohorts are also identical. In the electronic companion, we investigate the sensitivity of our conclusions to relaxing each of them. In particular, we allow (i) the size of each new cohort to be larger, smaller, or the same as that of the preceding cohort, and (ii) the mean churn rate of each new cohort to be higher, lower, or the same as that of the preceding cohort. These extended analyses provide clear evidence that our main result (i.e., the systematic underestimation of customer-base value when using a naïve calculation based on an aggregate retention rate) is insensitive to varying types of cross-cohort differences.

4. Implications for Retention Elasticities

There is a widely held belief that improvements in customer retention can have a major impact on customer (and therefore firm) value. Such an idea received a great deal of attention from work by Frederick Reichheld (e.g., Reichheld 1996) and has been the focus of subsequent studies by researchers such as Gupta et al. (2004) and Pfeifer and Farris (2004). For instance, Gupta et al. (2004) report retention elasticities ranging from 2.45 to 6.75, with an average of 4.88 (i.e., a 1% increase in the retention rate results in roughly a 5% increase in the value of the customer base). But these analyses are based on a formula that uses a single aggregate retention rate. Pfeifer and Farris (2004) and Reichheld (1996) go one step further and allow for increasing retention patterns in some of their analyses, but neither one gives any consideration to the role (or even the existence) of heterogeneity and the need to condition on the length of the customer's relationship with the firm. In light of the analyses presented above, it makes sense to investigate how these elasticities change when we account for heterogeneity in retention rates.

Recall from (3) 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 retention elasticity (interpreted as the percentage increase in the customer's residual lifetime value for a given percentage increase in their underlying retention rate)[5] is given by

$$\begin{aligned}\epsilon_{\text{ret}}&(d \mid \theta, \text{active for } n \text{ periods})\\ &= \frac{\partial DERL(d \mid \theta, \text{active for } n \text{ periods})}{\partial \rho}\\ &\quad \times \frac{\rho}{DERL(d \mid \theta, \text{active for } n \text{ periods})},\end{aligned}$$

where $\rho = 1 - \theta$ is the individual's (unobserved) retention probability and the expression for $DERL(d \mid \theta$, active for n periods) is that given in (5),

$$= \frac{1+d}{\theta + d}. \tag{7}$$

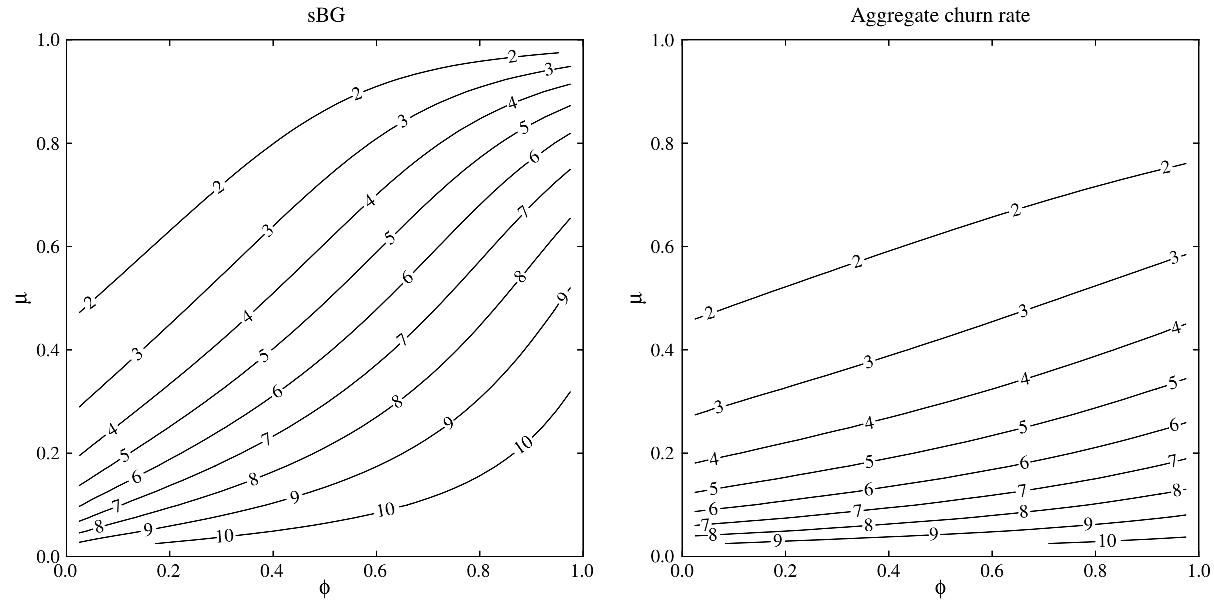
Following the logic associated with the derivation of (6), we take the expectation of (7) over the posterior distribution of θ for a customer who has been active for n periods since acquisition, giving us

$$\begin{aligned}\epsilon_{\text{ret}}&(d \mid \alpha, \beta, \text{active for } n \text{ periods})\\ &= {}_2F_1\left(1, \beta + n - 1; \alpha + \beta + n - 1; \frac{1}{1+d}\right).\end{aligned} \tag{8}$$

This is the retention elasticity for a randomly chosen individual acquired in period 1 who is still a customer in period n.

[5] This should not be confused with the *churn* elasticity, which would lead to different results.

Figure 3 Retention Elasticities Computed Using the sBG Model and Aggregate (2006/2007) Churn Rate, as a Function of μ and ϕ



We again assume a constant pattern of new customer acquisitions (i.e., 10,000 new customers each year) and no between-cohort variation in the mean churn rate. For each point on a fine grid of points in the (μ, ϕ) space, we compute the retention elasticity for the corresponding customer base at the end of 2007. The resulting contour plot of retention elasticities is presented on the left-hand side of Figure 3.

At first glance, these numbers seem consistent with those reported by other researchers. For example, the estimated values of the mean and polarization index for the first data set considered in Fader and Hardie (2007a) are $\hat{\mu} = 0.15$, $\hat{\phi} = 0.18$, which corresponds to an estimated retention elasticity of 6.5. Using the parameter estimates from the second data set ($\hat{\mu} = 0.37$, $\hat{\phi} = 0.35$), the estimated retention elasticity is 4.9.

Looking more closely at this plot, the fact that the contour lines are not flat indicates that accounting for heterogeneity clearly matters when computing retention elasticities. To fully understand the importance of heterogeneity in this context, we examine how these elasticity estimates compare to those computed using the aggregate retention rate. We compute the corresponding naïve estimates of the retention elasticity by evaluating (7) using the aggregate churn rate associated with each point on the fine grid of points in the (μ, ϕ) space; the resulting contour plot of naïve retention elasticities is presented on the right-hand side of Figure 3.

Comparing these two plots, we see that while the range of elasticities is the same, the pattern is quite different. We note that the elasticity estimates computed using the aggregate retention rate always underestimate the "true" values associated with the sBG model. To get a sense of the magnitude of this error, we again consider the parameter estimates associated with the two data sets considered in Fader and Hardie (2007a). For the first data set, the naïve elasticity estimate (5.0) is 23% lower than the model-based estimate (6.5). Likewise for the second data set, $\hat{\mu} = 0.37$, $\hat{\phi} = 0.35$, the naïve elasticity estimate (3.0) is 39% lower than the model-based value (4.9). Thus we find that failing to account for heterogeneity when computing the retention elasticity for a customer base underestimates the true value, just as we observed when computing the residual value of that same customer base. We can therefore conclude that accounting for heterogeneity makes a substantial difference when computing retention elasticities.

5. Discussion

The main message of this paper is that failing to account for cohort-level retention-rate dynamics leads to biased estimates of the residual value of a customer (and therefore a customer base). The simple example presented in §§1 and 2 illustrates the problem under the assumption that the observed increase in cohort-level retention rates over time is an artifact of cross-sectional heterogeneity in customers' unobserved (and unobservable) individual propensities to renew their contracts. The systematic analysis of the problem presented in §3 uses the sBG model of contract duration to capture the cohort-level retention-rate dynamics and finds that valuations performed using an aggregate retention rate underestimate the true value of the customer base by the order of 25%–50% in standard settings.

Elements of the analysis presented in this paper may feel familiar to a serious student of the issues surrounding the calculation of customer lifetime value. As previously noted, the heterogeneity explanation for increasing cohort-level retention rates is a well-known "ruse of heterogeneity" (Gupta and Lehmann 2005, pp. 29–31; Vaupel and Yashin 1985). Blattberg et al. (2008, p. 180) note that "using an average retention rate can underestimate lifetime value" but do not explore the issue. Gupta and Lehmann (2005, pp. 174–177) explore the issue of calculating CLV with increasing retention rates, concluding the differences are small. However, their analysis does not make an explicit comparison with an appropriately chosen aggregate retention rate. More fundamentally, their analysis ignores the implications of increasing retention rates when computing the residual value of an existing customer (i.e., using the conditional survivor function), which lies at the heart of any customer-base valuation exercise. As such, the true magnitude of the bias is not uncovered. We therefore conclude that the central message of this paper and the new result that drives it (i.e., (6)) are previously undocumented and important findings.

The analysis presented in this paper assumed that the phenomenon of increasing cohort-level retention rates is simply an artifact of the cross-sectional heterogeneity in customers' unobserved (and unobservable) constant propensities to renew their contracts at any point in time. Although there is strong empirical support for such an explanation (e.g., Fader and Hardie 2007a, Schweidel et al. 2008), we acknowledge that another explanation for this phenomenon is that the individual-level propensities to renew contracts are themselves changing over time. In such cases we can use a distribution such as the Weibull to characterize individual-level contract durations, coupled with an appropriate distribution to characterize cross-sectional heterogeneity. The logic of the analysis presented in this paper would still hold (e.g., we would simply evaluate (4) using the new survivor function). Similarly, when the cohort-level retention numbers are computed on a monthly basis, we may observe seasonality and the effects of introductory offers expiring. In such cases, the sBG model can be replaced by a model of contract duration that includes covariates to capture such effects—see, for example, Schweidel et al. (2008). However, we typically will not be able to come up with a closed-form solution to (4), relying instead on a numerical evaluation of the expression.

A potential limitation of our analysis is that we have not presented a formal analytical proof of the result, relying instead on numerical methods alone. Any formal proof would have to be conditioned on specific, limited assumptions about the relative size of each cohort, cohort-level mean churn rates, and so on. As previously noted, the electronic companion presents an extended analysis in which we allow (i) the size of each new cohort to be larger, smaller, or the same as that of the preceding cohort; and (ii) the mean churn rate of each new cohort to be higher, lower, or the same as that of the preceding cohort. This provides clear evidence that our main result (i.e., the systematic underestimation of customer-base value when using an aggregate retention rate) is insensitive to varying types of cross-cohort differences. Although 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. We are therefore confident in concluding that any attempt to compute the residual value of a customer (and therefore a customer base) using an aggregate retention rate will lead to a biased estimate of the true value that takes cohort-level retention-rate dynamics into consideration.

Furthermore, this error is of sufficient magnitude for any careful analyst to want to proceed correctly. So how can we perform such a customer-base valuation exercise in practice? We first need to estimate the model parameters. Fader and Hardie (2007a) show how to do this when we have data from a single cohort of customers. However, in a customer-base valuation setting, we will have data from more than one cohort of customers; Fader and Hardie (2007b) describe how to estimate the sBG model parameters in such a setting. Once we have estimates of the two model parameters, we compute *DERL* using (6) (or the approach outlined in the appendix using a simple Microsoft Office Excel spreadsheet).

6. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mktsci.pubs.informs.org/.

Acknowledgments
The authors thank John Roberts for his helpful comments on an earlier version of this paper. The second author acknowledges the support of the London Business School Centre for Marketing.

Appendix

As proven in the electronic companion, an alternative derivation of (6) follows from substituting (1) in (4),

$$DERL(d \mid \alpha, \beta, \text{active in } n) = \sum_{t=n}^{\infty} \frac{S(t \mid \alpha, \beta)}{S(n-1 \mid \alpha, \beta)} \left(\frac{1}{1+d}\right)^{t-n}, \quad (9)$$

Figure A.1 **Screenshot of Excel Worksheet for Computing DERL**

	A	B	C	D	E	F
1	α	3.8	**DERL**	3.59		
2	β	15.2				
3				2 renewals (n = 3)		
4	t	$S(t)$		$S(t \mid t>n-1)$	disc.	
5	0	1.0000				
6	1	0.8000				
7	2	0.6480				
8	3	0.5307		0.8190	1.0000	
9	4	0.4391		0.6776	0.9091	
10	[illegible]	[illegible]65		0.5656	0[illegible]	
11	[illegible]	[illegible]85		0.4761	0.7513	
12	7	0.2616		0.4037	0.6830	
13	8	0.2234		0.3447	0.6209	
14	9	0.1919		0.2962	0.5645	
15	10	0.1659		0.2560	0.5132	
201	196	6.14E–05		9.48E–05	1.03E–08	
202	197	6.03E–05		9.31E–05	9.33E–09	
203	198	5.93E–05		9.15E–05	8.48E–09	
204	199	5.82E–05		8.99E–05	7.71E–09	
205	200	5.72E–05		8.83E–05	7.01E–09	

= SUMPRODUCT(D8:D205,E8:E205)

=(B$2+A6–1)/(B$1+B$2+A6–1)∗B5

=B8/B7

=1/1.1^($A8–3)

and solving the infinite sum. This provides us with a way of computing $DERL(d \mid \alpha, \beta$, active in $n)$ when using a modeling environment that does not include the ability to evaluate the Gaussian hypergeometric function: assuming we can compute the survivor function associated with the sBG model, we evaluate (9), terminating the series at a point where additional terms are effectively zero.

It follows from the definition of a retention rate that

$$S(t) = \prod_{i=0}^{t} r_i, \tag{10}$$

where $r_0 = 1$. Coupling this with the expression for the (aggregate) retention rate under the sBG model (2) gives us a very easy way of evaluating the sBG survivor function.

Let us consider the task of computing DERL for one of the 6,480 Case 1 ($\alpha = 3.80$ and $\beta = 15.20$) customers who has made two contract renewals (i.e., has been active for $n = 3$ periods) in an Excel spreadsheet. With reference to Figure A.1, we first compute $S(t)$ for $t = 0, 1, 2, \ldots, 200$ using (9) and (10) in column B. The conditional survivor function, the first-term of the summand in (9), is then computed in column D. The second term of the summand, the discount factor, is computed in column E. Summing up the product of the two terms (cell D1) gives us the value of DERL for a customer who has made two contract renewals (evaluated just before the point in time at which the third contract renewal decision is to be made); this is the number reported in Table 4.

Using this approach (and terminating the series after 200 terms) instead of using (6) underestimates the true value of DERL by 5.21E−12. While this error should be acceptable in most settings, it is far more convenient to use (6) when the modeling environment can evaluate the Gaussian hypergeometric function.

References

Bauer, H. H., M. Hammerschmidt, M. Braehler. 2003. The customer lifetime value concept and its contribution to corporate valuation. *Yearbook of Marketing and Consumer Research*, Vol. 1. GfK-Nürnberg, e.V., Duncker & Humblot, Berlin, 47–67.

Blattberg, R. C., B.-D. Kim, S. A. Neslin. 2008. *Database Marketing: Analyzing and Managing Customers*. Springer, New York.

Fader, P. S., B. G. S. Hardie. 2007a. How to project customer retention. *J. Interactive Marketing* **21**(Winter) 76–90.

Fader, P. S., B. G. S. Hardie. 2007b. Fitting the sBG model to multi-cohort data. Note, http://brucehardie.com/notes/017/.

Fader, P. S., B. G. S. Hardie. 2009. Probability models for customer-base analysis. *J. Interactive Marketing* **23**(1) 61–69.

Fader, P. S., B. G. S. Hardie, K. L. Lee. 2005. RFM and CLV: Using iso-value curves for customer base analysis. *J. Marketing Res.* **42**(November) 415–430.

Farris, P. W., N. T. Bendle, P. E. Pfeifer, D. J. Reibstein. 2006. *Marketing Metrics: 50+ Metrics Every Executive Should Master*. Wharton School Publishing, Upper Saddle River, NJ.

Gupta, S., D. R. Lehmann. 2005. *Managing Customers as Investments*. Wharton School Publishing, Upper Saddle River, NJ.

Gupta, S., D. R. Lehmann, J. A. Stuart. 2004. Valuing customers. *J. Marketing Res.* **41**(February) 7–18.

Kotler, P., K. L. Keller. 2009. *Marketing Management*, 13th ed. Pearson Education, Inc., Upper Saddle River, NJ.

Kumar, V., W. J. Reinartz. 2006. *Customer Relationship Management: A Databased Approach*. John Wiley & Sons, Hoboken, NJ.

Lehmann, D. R., R. S. Winer. 2005. *Analysis for Marketing Planning*, 6th ed. McGraw-Hill/Irwin, Boston.

Ofek, E. 2002. Customer profitability and lifetime value. HBS Note 9-503-019, Harvard Business School, Boston.

Pfeifer, P. E., P. W. Farris. 2004. The elasticity of customer value to retention: The duration of a customer relationship. *J. Interactive Marketing* **18**(Spring) 20–31.

PricewaterhouseCoopers. 2002. *2002 Wireless Industry Survey*. PricewaterhouseCoopers LLP, Seattle.

Reichheld, F. F. 1996. *The Loyalty Effect: The Hidden Force Behind Growth, Profits and Lasting Value*. Harvard Business School Press, Boston.

Schmittlein, D. C., D. G. Morrison, R. Colombo. 1987. Counting your customers: Who are they and what will they do next? *Management Sci.* **33**(1) 1–24.

Schweidel, D. A., P. S. Fader, E. T. Bradlow. 2008. Understanding service retention within and across cohorts using limited information. *J. Marketing* **72**(January) 82–94.

Vaupel, J. W., A. I. Yashin. 1985. Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *Amer. Statistician* **39**(August) 176–185.

Wiesel, T., B. Skiera, J. Villanueva. 2008. Customer equity: An integral part of financial reporting. *J. Marketing* **72**(March) 1–14.