

Reflections on the Measurement of Customer Lifetime Value

Bruce G. S. Hardie
London Business School

AMA Sheth Foundation Doctoral Consortium
May 16-20, 2007

Measurement 101

- Define the construct
 - Theoretical definition
 - Operational definition
- Develop and test the measurement instrument
 - Validity
 - Reliability
- Use the measurement instrument for theory testing or applied work

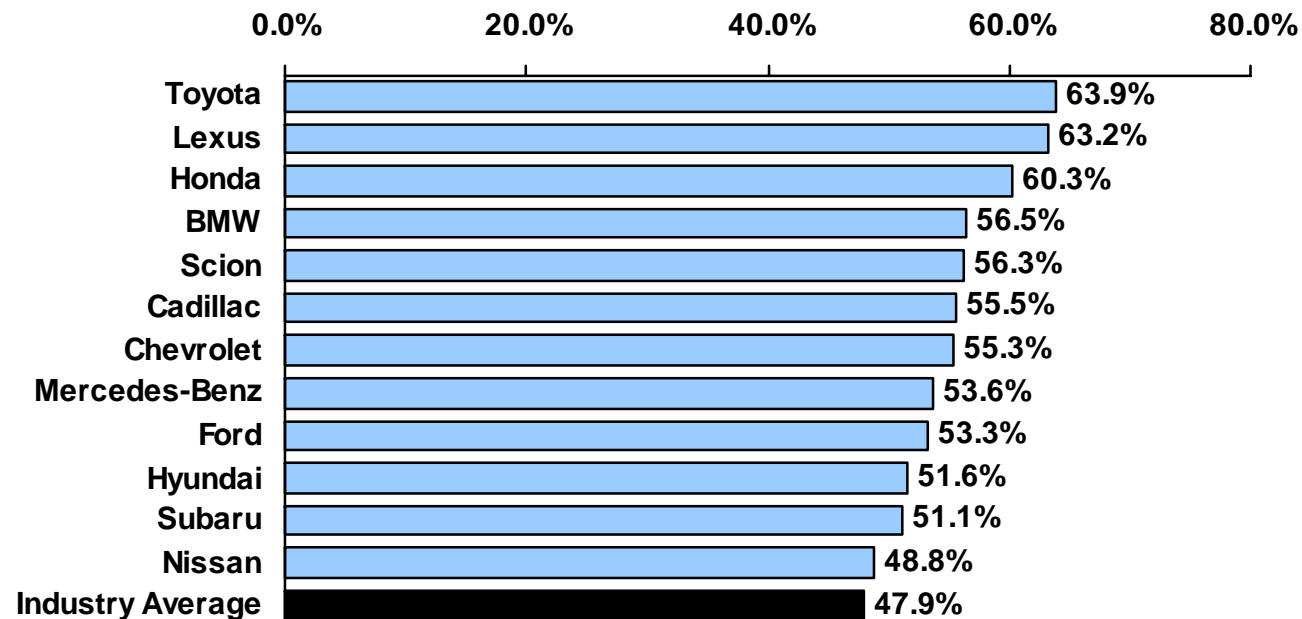
Why Care?

Invalid measurement can rob a research study of correspondence to the phenomena of interest, resulting in findings that have minimal use and add precious little to our knowledge of the world around us.

Alford and Engelland (*Journal of Business Research*, 2004, p. 95)

J.D. Power and Associates 2006 Customer Retention StudySM

2006 Make Retention Rates



“[T]he study measures the percentage of new-vehicle buyers and lessees who replace a previously purchased new vehicle with another from the same nameplate.”

Retention rate:

The ratio of customers retained to the number at risk.

Repeat rate:

The percentage of brand customers in a given period who are also brand customers in the subsequent period.

Repurchase rate:

The percentage of brand customers for a brand who repurchase that brand on their next purchase occasion.

Farris et al. (2006), *Marketing Metrics*.

Customer Lifetime Value

Customer lifetime value is

- *the present value of the future cash flows associated with the customer* (Pfeifer et al. 2005).
- *the present value of all future profits obtained from a customer over his or her life of relationship with a firm* (Gupta et al. 2006).

→ “customer lifetime value over the next two years” is not CLV

Calculating CLV

Using the language of math,

$$E(CLV) = \int_0^{\infty} E[v(t)]S(t)d(t)dt$$

- where $E[v(t)]$ = expected net cashflow/profit of the customer at time t (if active)
- $S(t)$ = the probability that the customer has remained active to at least time t
- $d(t)$ = discount factor that reflects the present value of money received at time t

Calculating $E(CLV)$

- Definitional — of little use by itself.
- We must operationalize $E[v(t)]$, $S(t)$, and $d(t)$ in a specific business setting ... then solve the integral.
- Important distinctions:
 - $E(CLV)$ of an as-yet-be-acquired customer
 - $E(CLV)$ of a just-acquired customer
 - $E(CLV)$ of an existing customer (expected *residual* lifetime value)

$E(RLV)$ as a Function of RFM

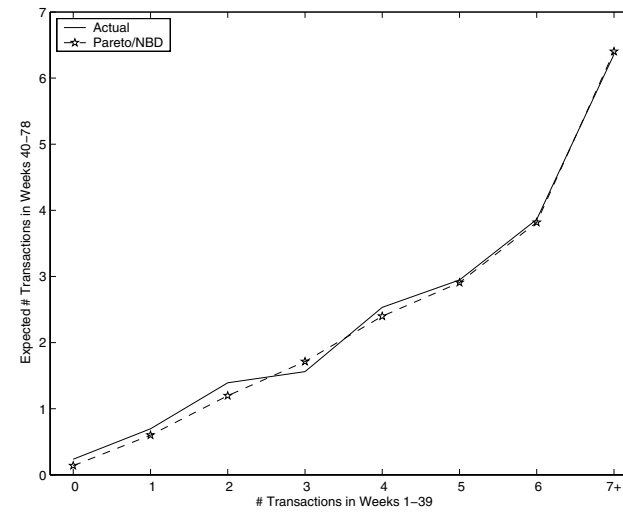
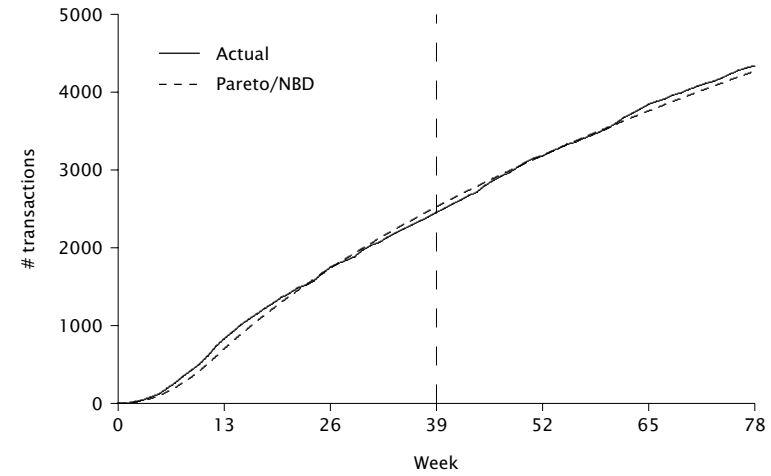
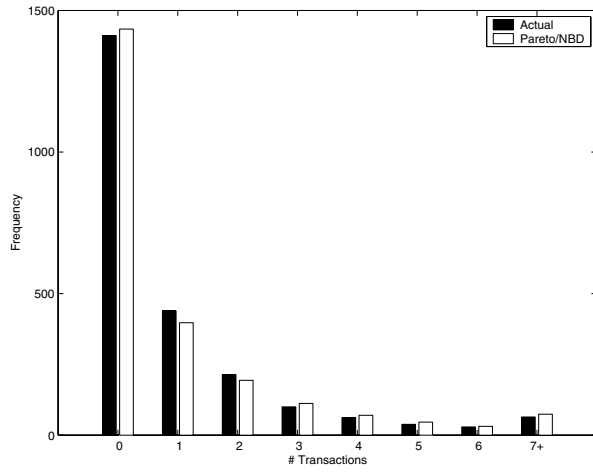
Fader et al. (*Journal of Marketing Research*, 2005):

For a customer who made x purchases in the period $(0, T]$, the last at t_x , with an average spend per transaction of $\$m_x$,

$$\begin{aligned} & E[RLV(\delta \mid r, \alpha, s, \beta, p, q, \gamma; x, t_x, m_x, T)] \\ &= \frac{\alpha^r \beta^s \delta^{s-1} \Gamma(r+x+1) \Psi(s, s; \delta(\beta+T))}{\Gamma(r)(\alpha+T)^{r+x+1} L(r, \alpha, s, \beta \mid x, t_x, T)} \\ & \quad \times \frac{p(\gamma+xm_x)}{px+q-1} \times \text{margin} \end{aligned}$$

where (r, α, s, β) are the parameters of the submodel for the flow of transactions and (p, q, γ) are the parameters of the submodel characterizing spend per transaction.

Model Validation



Classifying Customer Bases

(Adapted from Schmittlein et al. 1987)

Opportunities for Transactions	Continuous	Grocery purchases Doctor visits Hotel stays	Credit card Student mealplan Mobile phone usage
	Discrete	Event attendance Prescription refills Charity fund drives	Magazine subs Insurance policy Health club m'ship
		Noncontractual	Contractual

Type of Relationship With Customers

Opportunities for Further Research

(Gupta et al., *Journal of Service Research*, 2006)

- Moving beyond the limits of transaction data
- Moving from a customer to a portfolio of customers
- Reconciling top-down versus bottom-up measurements
- Cost allocations
- Developing incentive schemes that encourage globally optimal behavior
- Understanding the limits of CLV and CE
- Understanding the scope of Application
- Appreciating the limits of our theory-based models
- Understanding how to model rare events
- Recognizing the dangers of endogeneity
- Accounting for network effects

Ivory Tower Conceits

- Data are easy to acquire
- Computation is not an issue

PETER S. FADER, BRUCE G.S. HARDIE, and KA LOK LEE*

The authors present a new model that links the well-known RFM (recency, frequency, and monetary value) paradigm with customer lifetime value (CLV). Although previous researchers have made a conceptual link, none has presented a formal model with a well-grounded behavioral "story." Key to this analysis is the notion of "iso-value" curves, which enable the grouping of individual customers who have different purchasing histories but similar future valuations. Iso-value curves make it easy to visualize the interactions and trade-offs among the RFM measures and CLV. The stochastic model is based on the Pareto/NBD framework to capture the flow of transactions over time and a gamma-gamma submodel for spend per transaction. The authors conduct several holdout tests to demonstrate the validity of the model's underlying components and then use it to estimate the total CLV for a cohort of new customers of the online music site CDNOW. Finally, the authors discuss broader issues and opportunities in the application of this model in actual practice.

RFM and CLV: Using Iso-Value Curves for Customer Base Analysis

MARKETING SCIENCE

Vol. 24, No. 2, Spring 2005, pp. 275–284
ISSN 0732-2399 | EISSN 1526-548X | 05 | 2402 | 0275

informs

doi 10.1287/mksc.1040.0098
© 2005 INFORMS

“Counting Your Customers” the Easy Way: An Alternative to the Pareto/NBD Model

Peter S. Fader

The Wharton School, University of Pennsylvania, 749 Huntsman Hall, 3730 Walnut Street,
Philadelphia, Pennsylvania 19104-6340, faderp@wharton.upenn.edu

Bruce G. S. Hardie

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

Ka Lok Lee

Catalina Health Resource, Blue Bell, Pennsylvania 19422, kaloklee@alumni.upenn.edu

BGNBD Estimation

	A	B	C	D	E	F	G	H	I
1	r	0.243							
2	alpha	4.414	=GAMMALN(B\$1+B8)- GAMMALN(B\$1)+B\$1*LN(B\$2)			=IF(B8>0, LN(B\$3)-LN(B\$4+B8-1)- (B\$1+B8)*LN(B\$2+C8), 0)			
3	a	0.793							
4	b	2.426							
5	LL	-9582.4				=-(B\$1+B8)*LN(B\$2+D8)			
6									
7	ID	x	t_x	T	ln(.)	ln(A_1)	ln(A_2)	ln(A_3)	ln(A_4)
8	0001	2	30.43	38.86	-9.4596	-0.8390	-0.4910	-8.4489	-9.4265
9	0002	1	1.71	38.86	-4.4711	-1.0562	-0.2828	-4.6814	-3.3709
10	=SUM(E8:E2364)		0.00	38.86	-0.5538	0.3602	0.0000	-0.9140	0.0000
11	0004	0	0.00	38.86	-0.5538	0.3602	0.0000	-0.9140	0.0000
12	0005	0	0.00	38.86	-0.5538	0.3602	0.0000	-0.9140	0.0000
13	=F8+G8+LN(EXP(H8)+(B8>0)*EXP(I8))					=GAMMALN(B\$3+B\$4)+GAMMALN(B\$4+B8)- GAMMALN(B\$4)-GAMMALN(B\$3+B\$4+B8)			
14	0007	1	5.00	38.86					
15	0008	0	0.00	38.86	-0.5538	0.3602	0.0000	-0.9140	0.0000
16	0009	2	35.71	38.86	-9.5367	-0.8390	-0.4910	-8.4489	-9.7432
17	0010	0	0.00	38.86	-0.5538	0.3602	0.0000	-0.9140	0.0000
2362	2355	0	0.00	27.00	-0.4761	0.3602	0.0000	-0.8363	0.0000
2363	2356	4	26.57	27.00	-14.1284	1.1450	-0.7922	-14.6252	-16.4902
2364	2357	0	0.00	27.00	-0.4761	0.3602	0.0000	-0.8363	0.0000

Model Development With Application in Mind

- Cost of implementation
 - Analyst time
 - Software (licenses and familiarity)
 - Computational requirements
- Access to customer-level data
 - Summaries easy; access to raw data requires IT cooperation
 - Privacy concerns; “anonymizing” not necessarily a solution
 - Partial bans on transborder data flows

Model Development With Application in Mind

- Developing models using data summaries whose size is independent of the number of customers

Fader et al. (2004) # rows in dataset a function of # periods
(discrete/noncontractual)

Fader et al. (2007a,b) repeated cross-sectional summaries
(continuous/noncontractual)

- Do I care about CLV ... or conditional expectations?

In Conclusion ...

- Remember your classes on measurement when developing marketing metrics
 - definition
 - operationalization
 - validation
- If you care about use in the “real world”, design for “application”