University of Zurich[UZH]

SocialNetworks
UNIVERSITY RESEARCH PRIORITY PROGRAM

The Role of Time-Varying Contextual Factors in Latent Attrition Models for Customer Base Analysis

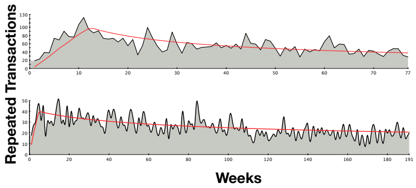
Patrick Bachmann[1], Markus Meierer[1]

[1]Department of Business Administration, URPP Social Networks, University of Zurich

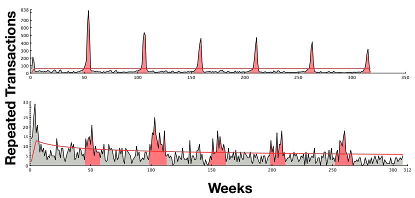
1. Background

Nowadays, **modeling customer purchases and attrition in non-contractual businesses is a straight-forward task**, but this simplicity comes at a price. Having access to recency and frequency data of customers' past transactions allows marketers to apply the Pareto/NBD model (Schmittlein, Morrison, and Colombo 1987). However, predictions are said to represent an educated guess rather than a precise value (Fader 2012; Malthouse 2009; Wübben and Wangenheim 2008).

- In general, prediction of transactions in non-contractual business settings works well.



- But purchasing patterns in some industries lead to less accurate predictions of the Pareto/NBD model:



2. Research Question

Efforts to improve the original Pareto/NBD model are manifold. However, **there exists no generalization that allows modeling time-varying contextual factors in a continuous non-contractual setting.**

Study	Setting continuous	Time-invariant factors attrition	Time-invariant factors purchase	Time-variant factors attrition	Time-variant factors purchase	Endogeneity Control	Maximum Likelihood	Gamma Heterogeneity	DERT/ DECT
Schmittlein, Morrison & Columbo (1987)	x						x	x	
Gupta (1991)	x				x		x	x	
Fader & Hardie (2007)	x	x	x				x	x	
Abe (2009)	x	x	x					x	
Singh, Borle & Jain (2009)	x	x	x						
Neslin & Rhoads (2009)	x	x	x						
van Oest &Knox (2012)	x						x	x	
Schweidel & Knox (2013)		x	x	x	x	x	x		
Knox & van Oest (2014)	x					x	x	x	
Braun & Schweidel & Stein (2015)	x	x		x			x	x	x
Platzer & Reutterer (2016)	x							x	
Gopalakrishnan, Bradlow & Fader (2017)	x	x	x	x	x				
Dew & Ansari (2018)				x	x				
McCarthy & Fader (2018)	x		x		x				
Xia, Chatterje & May (2019)	x		x	x					
This paper	x	x	x	x	x	x	x	x	x

3. Model & Contribution

In this paper, **we propose a latent attrition model that allows time-varying contextual factors to be modeled in continuous non-contractual settings**. Distinct to previous literature, we combine all following characteristics within the proposed approach:

(1) continuous nature of both, the purchase or attrition process

(2) possibility to include multiple time-varying and -invariant contextual factors which can separately influence both processes, only one or none,

(3) gamma heterogeneity for both processes,

(4) reduction to the standard Pareto-NBD model when estimated without any contextual factors,

(5) closed-form maximum-likelihood solution,

(6) derivation of managerial relevant expressions.

Relying on two simulation studies and the empirical analysis of three retailing datasets we assess the performance of the proposed approach and benchmark it against state-of-the-art Pareto- and Non-Pareto-type models. The results provide evidence about the inferential and predictive ability of our approach. In detail, we find:

(a) An **increased predictive accuracy** when including contextual factors,

(b) differences in the **increase of predictive accuracy depending on the scope of the modeled contextual factors**, i.e. individual level contextual factors increase predictive accuracy more than aggregated-level contextual factors,

(c) reliable **identification of the impact of exogenous factors** on both processes,

(d) reliable **identification of the impact of endogenous factors** on both processes when relying on (latent) IV approaches,

(e) a **key role of controlling for endogeneity to derive reliable parameter estimates**, but neglectable importance for predictive accuracy.

Managerial applications of a latent attrition model with such characteristics and performance are widespread: Combined with a Gamma/Gamma model (Colombo and Jiang 1999; Fader, Hardie, and Lee 2005), it enables academics and managers improves the identification of best future customers more accurately. In addition, the possibility to control for endogenous contextual factors allows to rigorously identify and quantify drivers of two important pillars of customer lifetime value (CLV), i.e. customers' purchases and attrition process.

4. Data: Three Retailing Datasets

- Multichannel retailer & Sporting goods retailer with **individual level** time-varying contextual factors
- Electronic retailer with **aggregate level** time-varying contextual factors

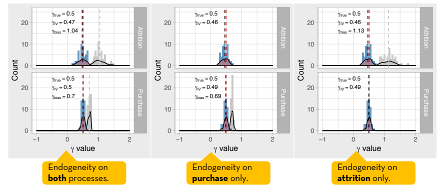
	Multichannel	Electronic	Sporting
Sample size (#cust.)	1'402	1'522	1'071
Time frame	7.7 years	6 years	6 years
Number of purchases	5'012	18'968	1'155
Zero repeaters	850	265	629
Estimation period	1 year	2 years	2 years
Cohort length	3 months	3 months	3 months
Time-varying cont. factors	D. Marketing Seasonal Pat.	Seasonal Pat.	D. Marketing (D. Marketing)² Seasonal Pat.
Time-invariant cont. factors	Acq. Channel	Gender Income	-

5. How to Handle Endogenous Contextual Factors?

Most of the contextual factors are non-random.
→ i.e. direct marketing decision are mostly based on a customer's purchase history.
We **propose the following methods to control for endogeneity:**

- 2-step IV approach for linear parameters for the endogenous variable
- Control function approach for non-linear parameters for the endogenous variable (Petrin and Train 2010)
- Copula correction method for discrete and binary endogenous variables or if IV is not available (Park and Gupta 2012)

The proposed model is able to recover endogenous contextual factors (here for the example of the 2-step IV approach):



6. Results

The presented results are from the multichannel retailer. Performance of the other two datasets is comparable.

We can quantify the determinants of customers' purchase and attrition process:

		Standard Pareto/NBD		Expended Pareto/NBD		
Pareto/NBD Parameters	LL-Value	-2'056		-2'020		
	BIC	4'083		3'967		
	r	1.134	*	1.610	*	Unobserved heterogeneity in transaction level is explained by contextual factors
	alpha	107.944	**	183.349	**	
	s	0.181	***	0.652	***	Unobserved heterogeneity in customer attrition is explained by contextual factors
	beta	0.373		10.374		
Contextual Factors	$gamma_{purc\ h,1}$	-		0.441	***	Direct marketing on purchase level
	$gamma_{purc\ h,2}$	-		0.729	***	Seasonality on purchase level
	$gamma_{purc\ h,3}$	-		0.088		Channel on purchase level
	$gamma_{life,1}$	-		-4.428		Direct marketing on churn
	$gamma_{life,2}$	-		-4.404		Seasonality on churn
	$gamma_{life,3}$	-		1.330	***	Channel on churn

These parameters can be interpreted as the rate elasticity of the contextual factors.

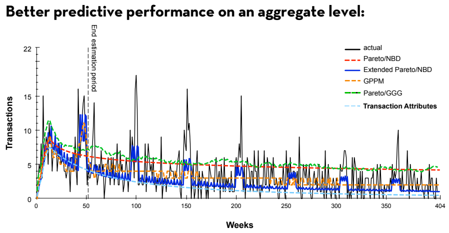
Better predictive performance on an individual level:

Prediction		Extended Pareto/NBD	Standard Pareto/NBD	Transaction Attribute	Pareto/GGG	GPPM
1 year	Individual MAE	0.263	0.306	0.326	0.301	0.346
	Aggregated MAE	2.229	2.840	2.864	3.130	1.668
2 years	Individual MAE	0.422	0.514	0.510	0.497	0.598
	Aggregated MAE	2.180	2.725	2.878	2.854	3.593
3 years	Individual MAE	0.538	0.681	0.638	0.650	0.820
	Aggregated MAE	2.013	2.635	2.655	2.692	3.936
4 years	Individual MAE	0.651	0.851	0.760	0.800	1.062
	Aggregated MAE	1.945	2.447	2.586	2.452	4.255
5 years	Individual MAE	0.757	1.025	0.871	0.950	1.347
	Aggregated MAE	1.895	2.389	2.509	2.316	4.255
6 years	Individual MAE	0.845	1.172	0.952	1.079	1.678
	Aggregated MAE	1.837	2.381	2.440	2.286	4.255
6.7 years	Individual MAE	0.904	1.277	1.022	1.171	1.951
	Aggregated MAE	1.779	2.364	2.357	2.238	4.255

Better identification the top customers:

		Extended Pareto/NBD	Standard Pareto/NBD	Transaction Attribute	Pareto/ GGG	GPPM
Best 10% Customers	High, correctly classified (%)	38.79	36.64	38.64	36.21	36.21
	Low, correctly classified (%)	87.86	87.44	87.44	87.35	87.35
	Overall correctly classified (%)	79.74	79.03	79.03	78.89	78.89
	Low but classified high (%)	12.13	12.56	12.56	12.65	12.65
	High but classified low(%)	61.21	63.36	63.36	63.79	63.79
	Overall incorrectly classified (%)	20.26	20.97	20.97	21.11	21.11
Best 20% Customers	High, correctly classified (%)	61.17	40.86	46.73	42.21	42.21
	Low, correctly classified (%)	82.06	71.53	74.66	73.20	73.31
	Overall correctly classified (%)	75.46	61.84	65.83	63.41	63.48
	Low but classified high (%)	17.94	28.47	25.33	26.80	26.69
	High but classified low(%)	38.83	59.14	53.27	57.79	57.79
	Overall incorrectly classified (%)	24.54	38.16	34.17	36.59	36.52

Better predictive performance on an aggregate level:



References

1. Colombo, Richard and Weina Jiang (1999), "A stochastic RFM model," Journal of Interactive Marketing, 13 (3), 2–12.
2. Fader, Peter S. (2012), Customer Centricity: Focus on the Right Customers for Strategic Advantage, Wharton: Wharton Digital Press.
3. ———, Bruce G.S. Hardie, and KL Lee (2005), "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," Journal of Marketing Research, 42 (4), 415–30.
4. Malthouse, Edward C. (2009), "The Results from the Lifetime Value and Customer Equity Modeling Competition," Journal of Interactive Marketing, 23 (3), 272–75.
5. Park, Sungho and Sachin Gupta (2012), "Handling Endogenous Regressors by Joint Estimation Using Copulas," Marketing Science, 31 (4), 567–86.
6. Petrin, Amil and Kenneth Train (2010), "A Control Function Approach to Endogeneity in Consumer Choice Models.," Journal of Marketing Research (JMR), 47 (1), 3–13.
7. Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987), "Counting Your Customers: Who-Are They and What Will They Do Next?," Management Science, 33 (1), 1–24.
8. Wübben, Markus and Florian Wangenheim (2008), "Instant Customer Base Analysis: Managerial Heuristics Often 'Get It Right,'" Journal of Marketing, 72 (3), 82–93.