CrossMark

RESEARCH ARTICLE

In Pursuit of Enhanced Customer Retention Management: Review, Key Issues, and Future Directions

Eva Ascarza[1] · Scott A. Neslin[2] · Oded Netzer[1] · Zachery Anderson[3] · Peter S. Fader[4] ·
Sunil Gupta[5] · Bruce G. S. Hardie[6] · Aurélie Lemmens[7] · Barak Libai[8] · David Neal[9] ·
Foster Provost[10] · Rom Schrift[4]

Abstract In today's turbulent business environment, customer retention presents a significant challenge for many service companies. Academics have generated a large body of research that addresses part of that challenge—with a particular focus on predicting customer churn. However, several other equally important aspects of managing retention have not received similar level of attention, leaving many managerial problems not completely solved, and a program of academic research not completely aligned with managerial needs. Therefore, our goal is to draw on previous research and current practice to provide insights on managing retention and identify areas for future research. This examination leads us to advocate a broad perspective on customer retention. We propose a definition that extends the concept beyond the traditional binary retain/not retain view of retention. We discuss a variety of metrics to measure and monitor retention. We present an integrated framework for managing retention that leverages emerging opportunities offered by new data sources and new methodologies such as machine learning. We highlight the importance of distinguishing between which

✉ Eva Ascarza
  Ascarza@columbia.edu

Scott A. Neslin
  scott.a.neslin@dartmouth.edu

Oded Netzer
  onetzer@gsb.columbia.edu

Zachery Anderson
  zanderson@ea.com

Peter S. Fader
  faderp@wharton.upenn.edu

Sunil Gupta
  sgupta@hbs.edu

Bruce G. S. Hardie
  bhardie@london.edu

Aurélie Lemmens
  a.lemmens@tilburguniversity.edu

Barak Libai
  libai@idc.ac.il

David Neal
  david@catalystbehavioral.com

Foster Provost
  fprovost@stern.nyu.edu

Rom Schrift
  roms@wharton.upenn.edu

[1] Columbia Business School, New York, NY, USA

[2] Tuck School of Business, Hanover, NH, USA

[3] Electronic Arts, Redwood City, CA, USA

[4] The Wharton School, Philadelphia, PA, USA

[5] Harvard Business School, Boston, MA, USA

[6] London Business School, London, UK

[7] Tilburg School of Economics and Management, Tilburg, Netherlands

[8] Arison School of Business, Hertsliya, Israel

[9] Catalyst Behavioral Sciences and Duke University, Durham, NC, USA

[10] New York University, New York, NY, USA

∅ Springer

customers are *at risk* and which *should be targeted*—as they are not necessarily the same customers. We identify trade-offs between reactive and proactive retention programs, between short- and long-term remedies, and between discrete campaigns and continuous processes for managing retention. We identify several areas of research where further investigation will significantly enhance retention management.

Keywords Customer retention · Churn · Customer relationship management (CRM)

1 In Pursuit of Enhanced Customer Retention Management

The central purpose of managing customer relationships is for the enterprise to focus on increasing the overall value of its customer base—and customer retention is critical to its success.
([66], p. 15)

Customer relationship management (CRM) managers and academics have long recognized the centrality and, in fact, the imperative of retaining customers. Customer retention is a cornerstone of broader CRM concepts such as customer equity [14, 76] and is arguably the most important component of the customer lifetime value (CLV) framework [37].

Yet, there are indications that companies have problems managing customer retention. From the customer point of view, 85% of customers report that companies could do more to retain them [40]. From the firm's point of view, while the vast majority of top executives report that customer retention is a priority within their organization, 49% of them admit to being unhappy with their ability to support their retention goals [33]. Furthermore, there is growing evidence that retention campaigns can be futile or even harmful [2, 4, 12].

Retention is especially important in the digital environment. For example, MusicWatch reported that only 48% of those who tried Apple Music, a music streaming service, still were using the service 2 months after it was launched [82, 100]. Similarly, for mobile apps, it is reported that, across all categories, 75% of all app users churn within 90 days [68].

Many researchers have studied customer retention (e.g., [49, 73, 92, 96]). Our review of the literature suggests an inordinate amount of effort has been devoted to predicting customer churn. Predicting churn is important. However, academics have afforded less attention to elements of campaign design such as whom to target, when to target, and with what incentives, as well the broader issues of managing multiple campaigns and integrating retention programs with the firm's marketing activities and strategy.

We therefore see managers struggling with how to manage retention appropriately and an academic community not focused on all the important issues. Therein motivates the purpose of this paper, to draw on previous research and current practice to distill insights and key issues regarding customer retention and to propose avenues for future research. In particular, we propose:

- Insights on how to measure retention. These range from 0:1 measures to recency/frequency calculations to metrics imputed from models designed to measure unobserved churn.
- An integrated framework for retention management. The framework takes a holistic view of customer retention, starting with a foundation of data and methods, and continuing with single campaign design and management, coordination across multiple campaigns, and integration with marketing strategy. The heart of this framework is the development of individual retention campaigns. We expand upon current practice in several ways, for example recognizing that the customers at highest risk of not being retained do not necessarily overlap 100% with those who should be targeted. We also highlight new data sources and emerging tools such as machine learning.
- A set of key directions for future research: Using the above framework, we identify opportunities for future research to enhance retention management.

A central theme of our work is that researchers and managers need to take a *broad view of retention management*. This starts with a flexible definition of retention and a leveraging variety of data and methodologies to measure and manage retention. It continues with a view of individual campaigns that goes beyond the measurement of customer churn and concludes with an integration of individual campaigns and consideration of their place in the firm's overall marketing strategy. We close each section with key issues to consider and build on these at the conclusion of the paper with a summary of the research opportunities.

2 Defining and Measuring Customer Retention

2.1 A Definition of Customer Retention

Our proposed definition is designed to capture the following concepts. First, the central idea that customer retention is continuity—the customer continues to interact with the firm. Second, that customer retention is a form of customer behavior—a behavior that firms intend to manage. Accordingly, we propose that "Customer retention is the customer continuing to transact with the firm."

A few things are worth noting in this deceptively simple definition. It emphasizes retention as something the *customer does* (which possibly could be affected by the firm). We use the word "transact" because the interactions between customers and firms can take multiple forms. For some businesses, the transaction is monetary—e.g., continuing to pay a subscription fee, make a purchase on a website. For other services, often in digital settings, transaction may not involve monetary exchange (e.g., using an e-mail account or using the free part of a freemium service).

The proposed definition encompasses both contractual and non-contractual relationships. Following Schmittlein et al. [77], customer-firm relationships can be categorized into *contractual* settings, in which firms directly observe when a customer terminates the relationship (e.g., a newspaper subscription), and *non-contractual* settings, where a customer makes no formal declaration of the termination of such a relationship (e.g., an online retailer). Even in contractual settings, where retention has been commonly defined by the customer's decision to renew, the focus on transaction in our definition goes beyond the renewal transaction and includes situations in which the customer ceases to transact and leaves the company long before she actually informs the firm. That is, the customer may be "walking dead" for a while [3]. Moreover, there are an increasing number of "hybrid" settings, in which customers can formally or informally cease their relationship with the firm [7]. This is the case for many online services where customers can either formally cancel their account or simply ignore the firm. Examples include daily deal sites (e.g., Groupon), social networks (e.g., LinkedIn), Web retailers (e.g., eBay), Web/e-mail services (e.g., Gmail), or travel services (e.g., Travelocity). The proposed definition encompasses those settings as well.

Furthermore, for firms with multiple products or services, ceasing to transact with one product or service may not mean that the individual stops transacting with the firm. For example, an online game player who got bored with game A but continues to play game B (provided by the same company) has stopped playing game A but still retained her/his relationship with the firm. Our definition indeed emphasizes that retention is about continued transactions with *the firm*, rather than merely with the product/service. This concept was investigated by Schweidel et al. [81] who show how customer retention for individual services offered by a firm relates to the customer retention with the firm.

Lastly, "churn" is the counterpart of retention. If the customer has decided to stop transacting with the firm, the customer has churned. In that sense, churn is inferred by the cessation the customers' transactions with the firm.

In summary, our definition of retention centers on continued transactions and applies to a wider range of businesses, regardless of the existence of a customer-firm contract or the type of transactions. These issues play out in the measurement of retention, which we discuss next.

2.2 Measurement

Measuring retention is important for several reasons. First, predicting retention lies at the heart of any attempt to calculate lifetime value (CLV) [11, 30, 37, 93] and customer equity (CE) [14, 27, 75]. Second, retention drives firm profitability and value. Across the firms analyzed by Gupta et al. [37], average retention elasticity is 4.9. That is, a 1% increase in retention rate produces nearly a 5% increase in CE. Third, measuring retention rates over time can provide a key metric of the firm's health. Fourth, as attributed to Peter Drucker, "If you can't measure it, you can't manage it."

There are several dimensions to consider when measuring retention. The main ones are articulated in Table 1 by considering the 2 × 2 table formed by contractual vs. non-contractual settings and individual customer- vs. aggregate-level measures.

Most of the retention metrics relevant for contractual firms are also relevant for non-contractual firms. A simple 0:1 indicator of transaction, and a measure of recency—how long it has been since the customer last transacted—are appropriate for both types of firms. However, recency on its own might not be a good indicator of customer retention. For example, customers A and B last purchased 6 months ago. On the one hand, customer A typically purchases once a year (i.e., her inter-purchase time is 12 months), thus a recency of 6 months should not be taken as an indication of churn because it is well within the customer's purchase cycle. On the other hand, customer B usually purchases every month, in which case a recency of 6 months should be worrisome for the firm. We thus recommend calculating a recency/inter-purchase-time ratio, where a ratio larger than one is an indication of a retention problem. Regarding the examples above, customer A's recency/inter-purchase-time is 6:12 = 0.5, whereas customer B's ratio is 6:1 = 6.

A few limitations are worth noting with respect to the recency/inter-purchase-time ratio. First, it requires observing a reasonable number of transactions to reliably calculate the average (or median) inter-purchase time.[1] Second, even for those customers for whom one observes a sufficient number of transactions, the inter-purchase time measure might be biased due to the right censoring nature of the data—the time between the last purchase and the last observation is ignored. To address these issues, statistical models have been developed to infer whether the customer is no longer retained, i.e., whether there has been "latent attrition" [31, 32, 77].

---

[1] We recommend to use the median in the recency/inter-purchase-time ratio when the distribution of inter-purchase time is skewed or when the number of observations is small.

Table 1   Examples of alternative retention metrics

		Type of business	
		Contractual	Non-contractual
Level of aggregation	Individual customer level	• 0:1 indicator of whether customer is still under contract at end of the period • 0:1 indicator of whether customer transacted this period. • Recency − number of periods since previous transaction • Recency/inter-purchase-time ratio	• Latent attrition, i.e., P(Alive) for a specific customer, inferred from a statistical model • 0:1 indicator of whether customer transacted this period • Recency − number of periods since previous transaction • Recency/inter-purchase-time ratio
	Aggregate level	• Retention rate − Number of customers who renewed in a particular period divided by the total number of customers who were up for renewal in that same period • Percentage of customers transacting during period • Distribution and summary statistics of recency across customers • Distribution and summary statistics of recency/inter-purchase-time ratio across customers	• Distribution and summary statistics of latent attrition, i.e., of P(Alive) by cohort • Percentage of customers transacting during period • Distribution and summary statistics of recency across customers • Distribution and summary statistics of recency/inter-purchase-time ratio across customers

Any measure calculated at the individual level can be aggregated across customers. The distribution and summary statistics of these aggregates are important, and particularly useful when monitoring retention over time [3, 47].

Perhaps the most common measure of retention is the retention rate, quantified as a zero-one indicator at the individual level and the percentage of customers retained at the aggregate level. Gupta et al. [37] provide a good example of using aggregate retention rate to calculate firm value. McCarthy et al. [55] expand on this work to show how to incorporate aggregate publicly disclosed customer data to value subscription-based firms, taking into account non-constant retention rates.

Caution needs to be taken when calculating aggregate retention rates because customer heterogeneity in individual retention rates can cause "survivor bias" [29]. Following Fader and Hardie [29], we demonstrate this bias using simple numerical scenarios with illustrative but realistic data (see Tables 2 and 3).

Table 2 illustrates Fader and Hardie's fundamental insight. We have two types of customers, "good" and "bad," with retention rates of 70 and 20%, respectively. We begin with 500 customers of each type. The 500 good customers experience a 70% retention rate; hence they go from 500 to $0.70 \times 500 = 350$ in the next year, etc. The 500 bad customers, with their 20% retention rate, go from 500 to $0.20 \times 500 = 100$ customers, etc. This produces a survivor bias with the good customers dominating the aggregate retention rate over time as they are more likely to stay within the customer base. The aggregate retention rate goes from 45 to 70% over 5 years, even though the underlying retention rate for each customer

has not changed. The company might be misled into thinking it is doing a wonderful job on customer retention, although in fact, no single customer has increased her retention propensity; only the customer mix has changed. The situation is more complicated if we incorporate newly acquired customers and retention rates differ across customers (see Table 3).

The above is an example of the general problem of aggregation bias, applied to the context of measuring aggregate retention. In order to overcome this aggregation problem, we suggest that firms report retention rates by acquisition cohort and/or by other observable characteristics that are thought to be related to retention rate. For example, if retention rate varies systematically by acquisition channel or customer gender, the measures/metrics should be reported by cohort, acquisition channel, or gender. See Fader and Hardie [29] for further discussion on how to develop models that account for heterogeneous retention rates, the root cause of the aggregation bias.

Of particular challenge is measuring retention for non-contractual services where transaction frequency is inherently low. For example, many repair services (for kitchen appliances for example) are purchased say every 8 years. Such a long purchase cycle does not allow managers to reliably measure latent attrition or the recency/inter-purchase-time ratio. One suggestion for such cases is to obtain attitudinal or other behavioral measures (e.g., usage) within the purchase cycle.

Another issue is the time period for calculating retention. Churn rates in the telecom industry are often reported at a monthly level [85], which might be misleading because some customers can churn, technically, at any time, whereas others are locked in longer contracts that are not due to renew on a

Table 2 Survivor bias in cohort-level retention rate calculations

| | | Retention rate of good customers = 70% | | | |
| | | Retention rate of bad customers = 20% | | | |
Years after acquisition	Customers (n)	Good (n)	Bad (n)	Customers left at end of year	Retention rate (%)
0	1000	500	500	450*	45**
1	450	350	100	265	59
2	265	245	20	176	66
3	176	172	4	121	69
4	121	120	1	84	70
5	84	84	0	59	70
6	59	59	0	41	70
7	41	41	0	29	70
8	29	29	0	20	70

*e.g., 450 = 500 × 70% + 500 × 20%

**e.g., 45% = 450/1000

monthly basis. One solution is to calculate the percentage who renewed from among those whose contract ran out this month.

- Key issues: Which retention measure(s) should a manager monitor? Which ones are more diagnostic for the particular managerial context? How to calculate aggregate retention measures?

3 Managing Customer Retention

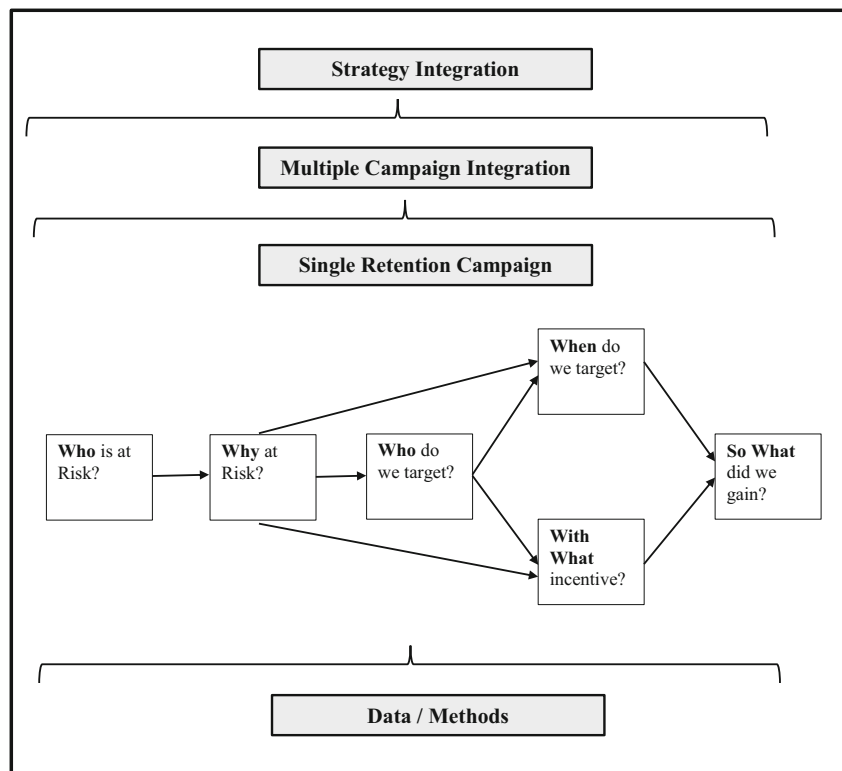Figure 1 proposes a framework for managing customer retention. It starts with a retention infrastructure—data and methods—and proceeds to the design of individual campaigns, coordinating multiple campaigns, and integrating retention efforts with marketing strategy. A key distinction is between *reactive* campaigns, where the firm waits for the customer to churn and then tries to "win back" the customer, typically with a financial incentive, and *proactive* campaigns, where the firm takes actions to solve in advance the problem that generates churn. Both of these approaches carry their own challenges. Firms typically implement multiple campaigns. For example, an insurance company may implement five discrete proactive campaigns during the year, with a reactive program constantly in effect throughout the year. The reactive and proactive programs need to be coordinated. Finally, all these efforts need to be integrated with the firm's marketing

Table 3 Different customers acquired first 2 years

Year	Custs. acquired (n)	Year 0 custs.	Year 1 custs.	Total custs. beg. of year	Year 0 custs. end of year	Year 1 custs. end of year	Total custs. end of year	Retention rate (%)
		Year 0 customer retention rate = 70%						
		Year 1 customer retention rate = 20%						
0	1000	1000	0	1000	700	0	700	70
1	2000	700	2000	2700	490	400	890	33
2		490	400	890	343	80	423	48
3		343	80	423	240	16	256	61
4		240	16	256	168	3	171	67
5		168	3	171	118	1	118	69
6		118	1	118	82	0	82	70
7		82	0	82	58	0	58	70
8		58	0	58	40	0	40	70
9		40	0	40	28	0	28	70

Fig. 1 Managing customer retention

strategy. For example, the target group for a Hulu retention campaign might be customers with no existing cable subscription, because they are more likely to respond positively to the campaign. However, this may be inconsistent with Hulu's overall segmentation and targeting strategy, which is to attract cable users to switch to streaming services.

3.1 Designing Single Retention Campaigns

As depicted in Fig. 1, to develop and evaluate a single retention campaign, the firm needs to: First, identify customers who are at risk of not being retained. Second, diagnose why each customer is at risk. Third, decide which customers to target with the campaign. Next, decide when to target these customers and with what incentive and/or action. Finally, implement the campaign and evaluate it.

These steps are applicable to both proactive and reactive campaigns. However, some difference between the two campaigns are worth noting. On the one hand, reactive campaigns are simpler because the firm does not need to identify who is at risk—the customer who calls to cancel self-identifies. "Rescue rates" can readily be calculated to evaluate the program, and subsequent behavior can be monitored. On the other hand, reactive campaigns tend to be more challenging because not all customers can be rescued, and customers learn that informing the firm about their intention to churn can be copiously rewarded by the firm, endangering the long-run sustainability of reactive churn management [50]. Hence, the incentive offered to win back the customer in a reactive campaign is typically high value, relative to a proactive campaign because the firm is certain the customer will churn [83].

Proactive campaigns are more challenging starting from the basic task of identifying who is at risk. State-of-the-art analytics are required to balance the costs of false positives (targeting a customer who had no intention to leave) and false negatives (failing to identify a customer who was truly at risk) ([15], chapter 27). Additionally, such campaigns must consider the probability of being able to retain those customers identified as would-be churners [2, 48, 69]. Our discussion of developing individual retention campaigns will thus largely focus on proactive campaigns.

Who is at risk? This entails using a predictive model to identify customers at risk of not being retained or in general of generating lower retention metrics. The dependent variable could be 0:1 churn or any measure of retention. Table 4 summarizes variables that have been found to predict churn in contractual settings. These include well-researched predictors like customer satisfaction, usage behavior, switching costs, customer characteristics, and marketing efforts, as well as more recently explored factors such as emotions and social connectivity.

Social connectivity factors can predict churn. In the context of telecommunications, it has been shown that high "social embeddedness," the extent to which the customer is connected to other customers within the network, is negatively correlated with churn [10]. Furthermore, the behavior of a

Table 4 Predictors of churn in contractual settings

Factors	Example	Method	References and industries
Customer satisfaction	Emotion in e-mails	Logistic, SVM, random forests	[21] (newspaper)
	Customer service calls	SVM + ALBA	[94] (telecom)
	Usage trends	Logistic, NN, SVM, genetic	[43] (telecom)
	Complaints	Logistic, NN, SVM, genetic	[43] (telecom)
	Previous non-renewal	Logistic, SVM, random forests	[21] (newspaper)
Usage behavior	Usage level	SVM with ALBA	[94] (telecom)
	Usage level	Logistic, NN, SVM, genetic	[43] (telecom)
Switching costs	Add-on services	Logistic, NN, SVM, genetic	[43] (telecom)
	Pricing plan	Dec Tree, Naïve Bayes, Logistic, NN, SVM	[91] (telecom)
	Ease of switching	Graphical comparison	[15] (p. 613) (telecom)
Customer characteristics	Psychographic segment	Logistic, NN, SVM, genetic	[43] (telecom)
	Demographics	Logistic, NN, SVM, genetic	[43] (telecom)
	Customer tenure	Logistic, decision tree	[1] (banking)
Marketing	Mail responders	Bagging and boosting	[47] (telecom)
	Response to direct mail	Logistic, SVM, random forests	[21] (newspaper)
	Previous marketing campaigns	Decision rules	[10] (telecom)
	Acquisition method	Probit	[23] (interactive TV)
	Acquisition channel	Logistic	[97] (financial services)
Social connectivity	Neighbor churn	Hazard	[61] (telecom)
	Social network connections	Random forests, Bayesian Networks	[95] (telecom)
	Social embeddedness	Decision rules	[10] (telecom)
	Neighbor/connections usage	Logistic	[4] (telecom)

customer's connections also affects her own retention. For example, a customer is more likely to churn from a service/company if her contacts or friends within the company churn [61, 95]. Relatedly, customers are less likely to churn if their contacts increase their use of the service [5]. Due to network externalities, the service becomes less valuable to the customer if her friends are not using it. These social-related factors are more likely to relate to churn for network-oriented services such as multi-player gaming, communications, and shared services, because customers exert an externality by using the service. If one goes beyond the predictive power of social connections and toward understanding the social effect, it is imperative to consider the similarity among connected customers (homophily), correlated random shocks among connected customers (e.g., a marketing campaign that is geographically targeted and hence affects connected customers), and true social contagious effects.

As noted above, Table 4 highlights predictors of churn in contractual settings, the dominant domain for churn prediction research. Much less work has been devoted to predicting attrition in non-contractual settings, where attrition is latent. More work is needed to evaluate the accuracy of latent attrition measures and its predictors (e.g., [80]).

An important question is whether churn prediction can be improved using ultra-fine-grained "big data." These are actions consumers take such as visiting a Web page, visiting a specific location, "Liking" something on Facebook, etc. These data often have very high dimensionality (millions to billions of potential actions) and extreme sparsity, as individuals only have so much "behavioral capital" to spend [24]. Thus far, there is no direct evidence of this sort of data improving retention prediction. However, it has been shown to improve prediction of customer acquisition [67, 70] and cross-selling [53]. Thus, we might postulate that ultra-fine-grained data should be considered for churn prediction as well.

Regarding methodology, a large body of research has focused on finding the best methods for predicting churn. These range from simple RFM models [20] to ensemble machine learning methods such as random forests [87]. While to the best of our knowledge, a formal meta-analysis has not been undertaken, studies often find that machine learning methods outperform traditional statistical methods [47].

Despite the plethora of work devoted to predicting churn, by no means is prediction perfect. Neslin et al. [59] report of a churn prediction tournament, the average top decile lift was about 2.1 to 1, meaning customers in the top decile were twice as likely to churn compared to average. However, churn rate was about 2%, which means that customers in the top decile had roughly a 4% chance of churning, so 96% were non-churners. These low accuracy rates endanger the profitability

of proactive retention. At the same time, they can be less costly for firms that it appears. Lemmens and Gupta [48] show that a profit-based loss function can reduce the risk of mis-prediction for the customers that matter the most and lead to profitable campaigns.

- Key issues: Are there yet-to-be-tested variables (e.g., ultra-fine-grade data) that could significantly enhance churn prediction? Is predictive accuracy high enough to be managerially useful? How can we predict retention in a non-contractual setting (where lack of activity is the proxy for retention)? Can advanced machine learning tools such as deep learning enhance the ability to predict churn?

Why at risk? The goal of a retention program is to *prevent* churn, hence understanding the causes of such behavior is necessary to design effective retention programs. There is a difference between determining the best predictors of churn and understanding *why* the customer is at risk of churning. For example, demographic variables might predict churn, but these variables rarely cause customers to leave the company. This distinction becomes less clear when we consider factors like past consumption or related behaviors. Are heavy users more likely to be retained *because* they consume more or is it that satisfaction with the service is driving both behaviors? More work is needed to isolate causality in churn behavior.

Furthermore, identifying specific causes for an individual customer to churn is quite different from identifying general causes in a population. For instance, in the context of telecommunication, a logistic churn prediction model may find that overage charges are an important predictor of churn. But a customer at risk may not have incurred any overage charges. Hence, overage is not the cause of risk for *that* particular customer. To identify the potential causes of churn for an individual customer, the researcher needs to find variables or combinations of variables that are both viable causes and for which the customer exhibits a risky behavior. One could use a competing risk hazard model [71], to predict which of the possible reasons of churn are most likely to cause churn at any point in time.

Once causes of churn are identified, one needs to isolate those that are controllable by the firm from those that are not [17]. For example, if a customer cancels her gym membership because she is moving to a different country, this is uncontrollable because there is nothing the gym can do about it. Analyzing a telecommunications service, Braun and Schweidel [17] find that accounting for uncontrollable factors can produce more profitable targeting of retention campaigns.

- Key issues: We need to identify not only *correlates* of low retention but also *causes* of it. This requires deeper theory and possibly "soft" measures such as customer satisfaction. We need to separate controllable (by the firm) causes of churn from non-controllable ones. How can we incorporate these causes into retention campaign design?

Whom do we target? At first blush, it seems we should target customers who are at the highest risk of not being retained. However, this may not be the best approach. The highest-risk customers may not be receptive to retention efforts [2]. They might be so turned off by the company that nothing can retain them. Instead, we propose that the best targets are customers who are at the risk of leaving *and* are likely to change their minds and stay if targeted.

Research on customer *response* to retention programs is scarce. A notable exception is recent work by Ascarza [2] who advocates the use of "uplift" models. Uplift models attempt to model directly the *incremental* impact of a campaign on *individual* customers. In the jargon of customer-level decision models, uplift modeling recognizes customer *heterogeneity* with respect to the incremental impact of the test. There are many ways to model incremental impact, ranging from interactions models to machine learning methods [2, 8, 36]. More research is needed to determine which methods work best under which circumstances.

Due to the predictive inaccuracies, some customers (in fact, often *most*) targeted in a proactive campaign will be those whom the company would have retained anyway. The targeting decision needs to take this into account ([15], chapter 27). Two examples highlight why targeting non-churners may be risky. Berson et al. [12] found that customers targeted by a retention campaign who did not accept the retention offer ended up with a higher churn rate than average. One possibility is that the offer triggered these customers to examine whether they wanted to stay with the company, and the answer turned out to be "no." This study is provocative but the data do not permit rejection of the alternative explanation that the retention offer bifurcated non-churners (satisfied customers who therefore accepted the offer) and churners (dissatisfied customer who therefore churned without accepting the offer).

Ascarza et al. [4] demonstrated that some *would-be non-churners* can be provoked by the retention effort to churn. Non-churners may be continuing to transact with the firm partly out of habit or inertia. Consumer habits have been found to keep consumers mindlessly locked into a relationship, but only insofar as they are not triggered to deliberate and actively weigh the pros and cons of switching [46, 101]. Thus, when directed to "habitual non-churners," retention efforts may inadvertently disrupt renewal habits, make people realize they are not happy with the status quo, and paradoxically cause churn. Of course, the impact could also go the other way—habitual non-churners may be so "delighted" by the retention offer that they like the company even more, their habits may be further strengthened, and their retention rates may increase. To date, the field has yet to systematically unpack these

dynamics, although the findings of Ascarza et al. [4] suggest that habits play a significant, under-studied, role in determining consumer response to retention efforts (see also [1]).

Another factor to consider in deciding whom to target is the position of the customer in the firm's social network. Historically, marketing applications of social network analysis have entailed customer acquisition (e.g., [41]), yet an emerging body of research is considering the effect of social connectivity on retention [5, 60]. Taking a social perspective, a customer with many contacts, or highly connected with customers who themselves are highly connected, can be very valuable because her/his defection could cause others to churn. The extent to which this applies may depend on the connectivity among network members [38]. Individuals who are central in a network may have lower risk of churning due to high social cost of leaving [35]. Because of the tendency of individuals to be in social networks with others like them, high-profitability customers may have a strong effect on each other, which further increases the value of targeting high CLV customers [39]. The socially related monetary loss due to customer churn would be much higher early in the product life cycle, when social influence is critical in driving product growth [42].

Finally, the decision of whom to target needs to take into account who is worth targeting. Because the goal of retention management is to maximize CLV, it may be that the customers who are easiest to retain are not the most valuable customers. The one-time cost of retaining a customer might exceed the customer's CLV. Neslin et al. [59] consider this in their calculation of retention campaign profitability. However, in addition, retention may entail long-term costs such as price reductions that decrease CLV going forward. Lemmens and Gupta [48] emphasize the importance of considering CLV in the decision of whom to target. Indeed, it may turn out that the customer who could be retained is not worth the cost it would take to retain her/him and thus the firm should let that customer churn. We call for further work on how to incorporate the change in future CLV vs. the required investment (that is, $\Delta CLV/\Delta Investment$).

- Key issues: Targeting high-risk customers is not always optimal. Firms need to quantify the incremental effect of their retention actions. How can firms employ uplift modeling? How can firms incorporate social connectivity into retention targeting? How can firms incorporate changes in CLV in targeting decisions?

With what incentive? One approach is to build on why customers churn to develop and target appropriate retention efforts to each customer. For example, in the telecom industry, age of the customer's smart phone and overage costs may be important predictors and drivers of churn. Consider two customers both in the high-risk pool. Customer A might have a 2-year-old phone but no overage charges, while customer B might have a 3-month-old phone but substantial overage charges. We would target customer A with an incentive for a new phone, but customer B would be targeted with a phone call aiming to put the customer on a more economical plan.

Advances in data availability and machine learning tools allow companies to personalize incentives to different customers now. In order to apply such tools to the domain of retention, firms can collect real-time information on the customers, including structured information such as usage of the product and browsing behavior as well as unstructured information from, for example, online chat sessions and call center conversations. Natural language processing can be used to convert the unstructured data (e.g., voice or text) to meaningful insights such as the underlying reason for the customer likelihood of churning. Combining these with structured data using machine learning tools can enable firms to customize and "optimize" the incentives they offer to the customers.

Other considerations need to be evaluated in selecting the best action to take to prevent churn. For example, a key one is whether to use price or non-price incentives. Price incentives perhaps are effective in the short run, but are easily copied by competition (cf. today's telecom industry) and imbue the customer with a "cherry-picking" mindset [9]. Thus, non-price incentives, such as product improvements (e.g., a gaming company adds additional levels in a game) may work better in the long term. This is one of the reasons for the success of smart phone subsidization in the telecom industry. It focused the customer on the service rather than on price, which is better for long-term attitudes toward the company (cf. [26]).

A promising approach is to let customers choose the incentive among a set of options. Research has shown that including in that set the option of no choice (i.e., do nothing) increases persistence among customers, which likely results in higher retention [78]. The firm can also design the retention effort in a way that it will mainly affect customers at high risk of churn and/or in a way that all customers, targeted or not, would appreciate it (e.g., product or service improvement). Another factor is the element of surprise. This is rooted in the concept of customer delight prevalent in services marketing [74]. Research has shown that surprising positive events strengthen the relationship between the customer to the company [63]. This is especially important because, as discussed earlier, the retention campaign will probably target many non-would-be churners. Ideally, these customers will be delighted by the offer and this will enhance retention in the long run even among those who were not going to churn at the time of the retention campaign.

- Key issues: What is the best way to personalize retention efforts? How can we use information on *why* customers churn to prescribe retention efforts? What is the value of monetary incentives for different types of retention

campaigns? What efforts are most likely to delight the customer?

When do we target? There is a trade-off between targeting a retention campaign to a customer too late vs. too early. In the extreme, a too-late proactive campaign becomes a reactive campaign—the customer is already out the door—and may either not retain the customer or cost too much to do so. However, a too-early campaign runs the risk of at best being irrelevant to the customer, at worst starting to get the customer thinking about churning.

One way to balance the too-late vs. too-early dilemma is to utilize data from previous campaigns to estimate models for churn and rescue probabilities as a function of time. Figure 2 shows how such approach could identify the right time to target a customer. In this example, starting from the current period, churn probability increases over time while rescue probability decreases. This yields period four as the optimal time to target because it balances the two trends and maximizes the probability of rescuing a would-be churner.

More broadly, a way to conceptualize when-to-target is to consider the different type of marketing campaign throughout the customer's lifecycle: acquisition → pre-emptive → proactive → reactive → win-back → post win-back (Fig. 3). It should not be necessary to target immediately after the customer has been acquired. However, even right after acquisition, the firm should be thinking about retention. For example, a telecom company could make sure the customer is on the right data plan from the very beginning. Pre-emptive timing would be to target the customer before the customer shows any sign of diminished retention. For example, a gaming company would offer incentives for the customer to start a new game *before* the usage rate of the current game starts to diminish. Proactive timing would be to launch a campaign targeted at customers who are identified as a retention risk but have not churned yet. Reactive timing is when the firm tries to prevent the customer from churning, while that customer literally is in the act of churning. Win-back is when the customer has churned and the company attempts to re-acquire or the customer [45, 89]. Stauss and Friege [86] provide a conceptual foundation for win-back strategies. They propose an in-depth dialog with the customer to determine the reasons for churn and design customized incentives. Post win-back actions refer to contacts initiated after the customer has rejected a win-back offer. These are frequent in industries where the customer must return some device to the company after defecting (e.g., a cable box).

- Key issues: How should firms design and time their retention efforts along the continuum from acquisition to win back (Fig. 3)? When should a firm "upgrade" a customer from one product to another to enhance firm retention?

So what? Campaigns need to be evaluated, if possible, using a control group randomly selected not to be targeted. This allows top-line results to be compiled easily without formal causal modeling. Metrics to be calculated include overall profitability and various retention measures. For example, Blattberg et al. [15] suggest how to calculate the "rescue-rate" of a campaign, i.e., what percentage of would-be churners were rescued. The company should calculate the rescue rate for all its campaigns. Then, the company can undertake a meta-analysis across multiple campaigns to understand which factors influence rescue rate (incentive characteristics, characteristics of customers targeted, the match between these two, etc.). Additionally, the company could employ a *heterogeneity in treatment effect* analysis (e.g., [2]) to assess which customers were most affected by the campaign.

Another part of evaluation is to determine the long-term impact of the campaign. The customer might have been retained this time, but what impact did that have on the customer's future profitability? What is the customer's retention risk after the campaign? Did it decrease because the customer was more satisfied, or increase because now the customer expects incentives?

- Key issues: What are the characteristics of successful retention campaigns? Which part of the campaign design process (Fig. 1) is most important for enhancing success?

4 Multiple Campaign Management

The steps we advocate for single campaigns are also relevant for managing multiple campaigns, except now the questions of who is at risk, why at risk, whom to target, when to target, with what efforts, and what did we gain, will be asked in a dynamic setting across several campaigns. Blattberg et al. [15] discuss two key issues in multiple CRM campaign management: wear-in and wear-out. A campaign may take time before it reaches its maximal impact (wear-in), and then decline at some rate afterwards (wear-out). These concepts have important implications for the spacing of retention campaigns.

At the customer level, multiple campaign management is challenging because the current campaign can influence what "state" the customer is in for the next campaign. For example, the customer targeted with campaign #1 may be at a lower level of risk for campaign #2 if there is a delight effect. The customer who received a free squash racket in campaign #1 may not be very responsive to campaign #2 that offers squash balls. This set-up suggests the application of dynamic

Fig. 2 When to target a retention campaign: trading off churn probability vs. rescue probability. $P(Churn) = \alpha_c(1 - \exp^{-\beta_c t})$ and $P(Rescue) = \alpha_r \exp^{-\beta_r t}$ with $\alpha_c = 0.5$; $\beta_c = 0.1$; $\alpha_r = 0.5$; $\beta_r = 0.2$; $P(\text{Rescue would-be churner}) = P(Churn) \times P(Rescue)$
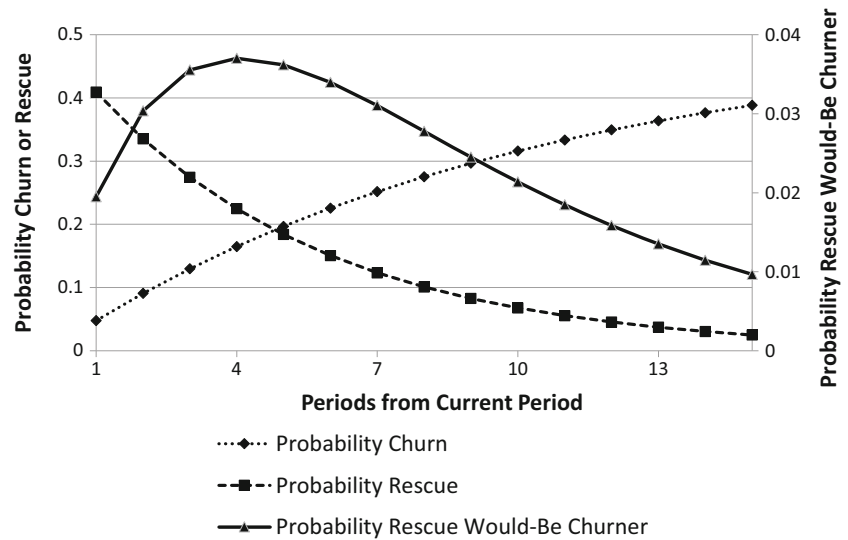
optimization. See Neslin [58] for discussion of dynamically managed customer-level CRM campaigns.

A challenging question is whether the firm should even use discrete retention campaigns or target customers as warranted on a continuing basis. Many firms in fact use a more continuous approach. For example, a telecommunication company may, on an ongoing basis, offer discounts to customers whose contract is nearing expiration. Firms need to integrate their business-as-usual retention efforts with their specially designed campaigns.

- Key issues: Firms need to coordinate proactive and ongoing retention programs. How can dynamic optimization be used to plan multiple campaigns?

5 Strategy Integration

Two important strategic integration tasks are (1) coordinating acquisition and retention and (2) aligning retention spending

Fig. 3 Alternative timing of retention campaigns
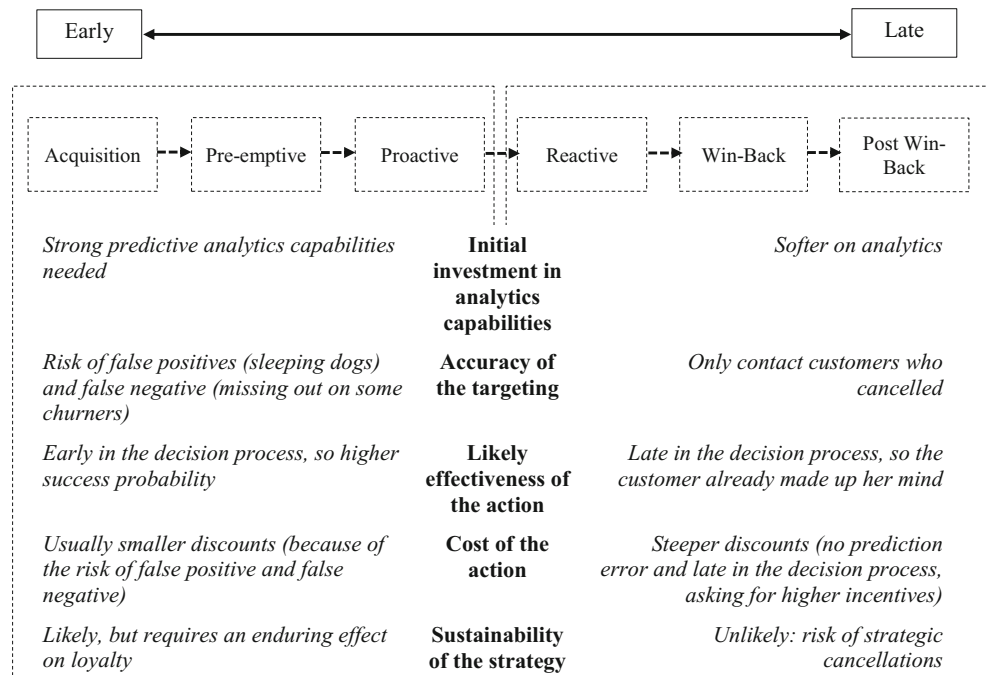
with the firm's marketing strategy and its segmentation, targeting, and positioning (STP) approach.

Consider the following simple example which highlights why acquisition and retention must be coordinated: company A may go all-out in acquiring customers. As a result, company A acquires highly risky customers who are not likely to be retained. This makes the task of designing retention campaigns more difficult. Company B may selectively acquire customers it knows to be high value (via a predictive model). This makes retention campaign design easier. Which approach is better is an empirical question, but there is no question that acquisition and retention campaigns must be coordinated.

Some work has been done in this area. Probably the earliest example is Blattberg and Deighton [13]. They use decision calculus [52] to judgmentally calibrate equations to represent customer acquisition and retention in order to maximize profits. Reinartz et al. [73] develop models for customer acquisition, lifetime (a proxy for retention), and profits. They show how the models can be used to allocate acquisition and retention funds to maximize total customer profitability. They find that suboptimal allocation of retention expenditures will have a greater impact on long-term customer profitability than will suboptimal acquisition expenditure. Several studies have examined how factors such as competitive intensity [98], responsiveness to CRM efforts [57] or supply limitations [64] affect the acquisition-retention trade-off. (See Ascarza et al. [6] for discussion of this topic.) Additionally, different marketing actions are likely to affect acquisition and retention. In the context of pharmaceutical drugs, Montoya et al. [56] find that whereas free drug samples were more useful as an acquisition tool, detailing visits were more effective as a retention tool.

Retention efforts must also be coordinated with the firm's STP and its marketing strategy because the two can easily fall out of sync. For example, the STP of a financial services company may be to segment the market by customer value and target high-value customers with premium products and services. A retention campaign that emphasizes promotional discounts would be "off strategy." Stahl et al. [84] find, for example, that advertising can increase product differentiation, but product differentiation decreases acquisition and retention rates. As a result, the customer management group may be trying to enhance acquisition and retention using targeting email, banner ads, etc., while the brand management group is undermining these efforts with its advertising.

- Key issues: How can the firm optimally manage both acquisition and retention? Are there conflicting goals between the two? Firms should coordinate retention activities with other elements of the marketing mix to ensure the STP approach of the firm is coherent.

6 Data and Methodology

Companies today have a treasure trove of structured and unstructured data that potentially offer further insights with respect to who, why, and when the customer is likely to churn, as well as whom to target and with what incentive. To date, researchers' main focus has been to predict the risk of churn, using primarily information on customer activity and demographics. They have done so by employing traditional methods such as logistic regression, probability models, and hazard models. Recent advances in data collection (e.g., clickstream behavior, social media activity, social connections, and call center/online chats) coupled with the rise of machine learning methods (e.g., classification methods and natural language processing) have opened a window of opportunities for retention research. For example, new data sources enable the field to tap into, not only the customers' activity, but also their social connections, the content of their interactions with the firm, and their emotional state. Advances in machine learning methodologies enable researchers to include a large set of predictors (often in the hundreds or thousands) in a non-linear fashion, and to extract useful information from unstructured sources of data such as audio and video conversations and images. Leveraging these data and tools will potentially allow to address questions such as why the customer is likely to churn and how to "save" him or her.

Table 5 displays the data and methods currently in use for retention planning and where they are used in the planning process. Below, we highlight some of the less commonly used data and methods.

We believe that data on emotions, unstructured customer/ firm interactions, customer states and traits [44, 54], and social connections will garner more attention. Emotions and social connections, as discussed earlier, have already proved their mettle with regards to identifying who is at risk. But we believe the most important use of these data will be in understanding why the customer is at risk and whom to target. Social connections data could be particularly important for targeting key "influencers," i.e., customers that provide network value to other customers (Ascarza et al. [5]).

Textual data can also provide key insights with respect to why customers churn. For example, analyzing the content of call centers and online chats may shed light on the causes of the customer's dissatisfaction with the product or service and can be used to identify commonly mentioned problems and competitors. Textual analysis approaches such as linguistic inquiry and word count (LIWC; [65]) and topic modeling [16] may be used to extract measures of emotions and other useful insights from consumer data.

Real-time collection of engagement metrics have potential to identify what retention actions to take. For example, measures of how a customer plays a computer game may provide a hint of what incentive (monetary or not) the customer needs to

Table 5 Data and methodologies for retention management

Data	Description	Retention question	Relevant literature
"Usual suspects"	Transactions, demographics	Who is at risk; whom to target	Neslin et al. [59]
Emotions/attitudes/traits	As inferred from text mining	Who is at risk; why at risk	Coussement and Van den Poel [21]
Detailed engagement	Detailed interactions of the customers with the product such as usage or browsing behavior	Who is at risk; why at risk	Verbeke et al. [94]
Marketing mix and retention campaigns efforts	Aggregate by region as well as per time period, and/or of individually targeted campaigns	So what; strategy integration	Verhoef [96]; Reinartz et al. [73]
Social influence/connectivity	Embeddedness, etc.	Who is at risk; whom to target	Nitzan and Libai [61]
Unstructured data on customer-firm interaction	Textual or voice data from call center or chat discussions	Why at risk; whom and when to target; with what incentive	Coussement and Van den Poel [21]

Methods	Description	Application Area	Relevant literature
Statistical methods	Regression, logistic regression, HMM	Individual campaign design and multiple campaign planning	Neslin et al. [59]; Ascarza and Hardie [3]
Probability models	BG/BB, BG, BG/NBD, Pareto-NBD	Forecasting aggregate churn patterns	Schmittlein et al. [77]; Fader and Hardie [28]
Machine learning	Classification and prediction tools such as decision trees, bagging, boosting, and random forests	Churn prediction models Proactive churn management	Lemmens and Croux [47]; Castanedo et al. [19]; Ascarza [2]
Text mining	Topic models, bag of words methods	Quantifying emotions, attitudes, and unstructured data	Coussement and Van den Poel [21]
Dynamic optimization	Dynamic programming	Multiple campaign planning	Delanote et al. [25]
Decision support systems	Decision calculus and agent-based models, optimization models	Individual and multiple campaign management; strategy integration	Blattberg and Deighton [13]; Rand and Rust [72]
Field experiment	Rescue rate, heterogeneity in treatment effect	Individual and multiple campaigns	Ascarza et al. [4]; Ascarza [2]

keep playing. Finally, knowledge management will be necessary to ensure the firm's entire experience base in retention management is codified and accessible to planners. This is especially important for planning multiple campaigns and integrating strategy, because these tasks require a broad understanding of the firm's experiences.

Turning now to methods for building models, tried and true regression-based predictive modeling is useful for constructing (1) predictive models for identifying who is at risk and who will respond to targeting, (2) meta analyses of field data that provide the insight needed for multiple campaign planning, and (3) marketing mix models that can drive strategy integration. In many settings, mostly non-contractual, churn is latent. Hence, latent space models such as hidden Markov models (HMMs) offer a natural approach to capture latent attrition ([7]; Schwartz et al. [79]). Even when churn is observed, the drivers of the customer decision to churn are generally latent. HMMs have been used to capture the dynamics in customer behavior that precedes the observed churn [3, 81].

The field of machine learning has focused on more advanced approaches using ultra-fine-grained big data modeling discussed earlier. These developments could play key roles in developing predictive models for who is at risk, why, who will respond, and when to target. Deep learning, for example, is a machine learning approach based on neural networks that combines supervised and unsupervised aspects. Deep learning has been used to learn about customer probability to defect [19], and may also be helpful in modeling response to retention offers. Similarly, boosted varying-coefficient regression models have been studied for dynamic predictions and optimization in real time [99] and have been shown to offer a major improvement over the classic stochastic gradient boosting algorithm already used for churn prediction [47]. These methods provide a promising direction for retention management.

Dimensionality reduction and variable selection techniques form another major development in the machine learning field, and may be useful for retention research and practice, which face an overflow of potential defection predictors. When building models to determine *who is at risk, whom to target*, and *when to target*, modelers should include in their toolkits modern regularization techniques such as the Lasso

[90], elastic net [102], and adaptive regularization [22]. Employing methods that allow modeling larger feature sets, also offer the potential to include interactions between churn predictors. Of particular interest to estimating the effects of churn incentives is to include interactions between the treatment variable (i.e., being targeted with a retention action) and customer or campaign design covariates [2, 51].

Feature-selecting regularization such as the Lasso need to be applied with care as interpreting the coefficients of these approaches may suffer from "selective inference" [88]. That is, searching for the best predictors from a set of 100s of potential predictors "cherry-picks" the best ones and can overstate their statistical significance. This is particularly relevant when trying to better understand *why customers are at risk*. Taylor and Tibshirani [88] discuss some potential solutions to the problem of selective inference. Dynamic optimization will be necessary for multiple campaign planning, especially if companies morph the classic discrete retention campaign into continuous-time retention management (e.g., [25, 62]). Finally, decision support systems can also be used for individual and multiple campaign design as well as strategy integration.

- Key issues: How to leverage text data (from emails, recommendations, etc.) for managing retention? What is the best way to identify a limited but valid set of predictors for retention models? How can machine learning methods be leveraged for retention management? What methods will prove scalable and implementable in real time, and what are the potential benefits of and drawbacks of each method?

7 Future Research Directions

The objective of this article is to draw on previous research and current practice to generate insights on retention management and identify key areas for future research. The main theme emerging from this analysis is an advocacy to *take a broad perspective* on retention management. We encourage researchers and practitioners to define retention in terms of customer transaction activity as well as the traditional either/or view. Retention should be measured and evaluated using a variety of metrics at both the customer and the firm levels, and consider whether the business is contractual or non-contractual. We caution against the use of unweighted aggregate metrics of retention observed over time, which can be misleading due to survivor bias (Tables 2 and 3). We encourage thinking of churn prediction as just one of several interrelated elements that comprise retention campaign

management (Fig. 1). Both academics and practitioners need to move beyond predicting which customers are at risk, and work on understanding *why* they are at risk, *whether they are retainable*, and *what incentive* will increase retention. Retention campaigns should be planned with a view toward future retention campaigns and integrated with customer acquisition as well as the firm's fundamental STP marketing strategy. All these lessons require a broad and expanded view of retention, recognizing that retention management is about keeping customers transacting with the firm.

Our analysis has yielded several areas for future research. These include:

- The design of reactive campaigns: Research is needed to determine whether reactive campaigns can be profitable, and which incentives are most effective. Do reactive campaigns undermine the customers' baseline retention rate in the same way that over-use of price-discount promotions can undermine repeat rates for consumer packaged goods [34]? Do reactive campaigns train a cadre of customers to be in chronic need of a special treat from the firm's customer care department? How can we best design reactive campaigns to complement proactive retention campaigns?
- Who is at risk: This area has received a good deal of research. It is time for a meta-analysis of churn prediction to identify which methods are best, and which variables better predict such behavior. Despite all the attention, predictions of retention risk are far from perfect. Is there in fact a ceiling in predicting churn, i.e., perhaps top decile lift for a 2% base rate simply cannot be higher than say 5 to 1? Are there overlooked powerful churn predictors (e.g., ultra-fine-grained data, the value of textual or voice data, customers' emotions, social connectivity)? Can attribution modeling be applied to churn prediction, i.e., what *sequence* of experiences makes a customer likely not to be retained? Churn-prediction models work well for well-defined retention measures such as 0:1 churn. The analogous metric for non-contractual businesses is the probability of continuous transactions − P(Alive). Further research is needed that shows the factors that predict P(Alive) (e.g., [80]). Finally, how should we measure retention for low-frequency purchased products, particularly durables?
- Why at risk: Progress has been made in uncovering deeper reasons why customers churn, particularly the role of social influence. However, further research is needed in this area. We need to separate *predictors* from *causes* (e.g., customer satisfaction and habit are likely to be causes; demographics are just predictors). Furthermore, can we build on Braun and Schweidel [17] in isolating controllable causes so that managers can utilize the insights, not just the predictions, from retention models? Another important why-at-risk question relates to acquisition channel.

Previous work has shown that acquisition channel matters, but is there a deeper theory that could guide the manager considering whether to utilize a particular channel for acquisition? Are different retention rates for different channels due to self-selection or due to the process of acquisition?

- Whom do we target? Traditional approaches rank order customers in terms of their risk of churning and "draw the line" somewhere down the list (e.g., [18]). An improved approach is to focus on incremental response (uplift) to retention efforts. Can retention risk, incentive response, and customer profitability be integrated to guide the whom-to-target decision? How effective are various methods of modeling uplift? Not only do we need to understand better whom to target, but whom *not* to target. Due to the typical low base rate of churn, retention campaigns are more likely to commit a type I error (target a customer who does not need to be retained) than a type II error (not target a customer who needs to be retained). Can we better measure these errors and balance them appropriately in deciding whom to target? What factors identify the campaign design factors that create delight rather than a churn-creating negative reaction among non-churners?

- With what incentive: When are monetary incentives advisable in contrast to non-monetary efforts? How can firms implement personalized retention incentives? Which retention efforts not only rescue the customer in the short run, but enhance CLV in the long run? How can we integrate the reasons *why* customers churn into the best action for retaining customers? We need as many studies on appropriate retention efforts as there are on predicting churn.

- When do we target? How can we trade off targeting a customer too early vs. too late? From a broader perspective and following Fig. 3, how much should a firm invest in retention along the continuum ranging from acquisition to win-back? When is it time for the firm to upsell the customer to a new product to forestall churn?

- So what? First, there is some evidence that many campaigns are futile in terms of preventing churn (e.g., [2, 12]). Can we document campaigns that show success? What are the key factors that determine success? What impact do retention campaigns have on long-term retention rate as well as customer profit contribution given the customer is retained? What are typical rescue rates, and what influences them? Finally, how does long-term success vary by industry and program design?

- Data and Methods: What are the best ways to leverage unstructured data, ultra-fine-grained behavior data, and social connectedness data to enhance retention management? How can we use attitudinal and other measures of retention in low-transaction frequency industries? Where and how can machine learning tools be applied to managing retention. How can we extend Lemmens and Gupta's [48] notion that models should be estimated using campaign profitability as an objective, not mere maximization of a "neutral" likelihood function? Finally, can we use competing risk models, commonly used in biostatistics to predict likely causes of death, to model *why* customers churn as opposed to merely *who* is at risk?

- Multiple campaign management: To our knowledge, this area has received virtually no attention in academic research. Research is needed that helps managers plan how many retention campaigns to run and at what timing, e.g., via dynamic programing optimization. Another dimension that needs research is how to coordinate proactive and reactive campaigns. Finally, how can we coordinate discrete retention campaigns with ongoing, continuous efforts to manage retention at the customer level?

- Strategic integration: How does one optimally allocate acquisition and retention funds over time? How does one balance the sometimes-conflicting goals of high acquisition rates and high retention rates? A wide-open area is the coordination of retention activities with other elements of the marketing mix. For example, how should firms allocate funds between discrete retention campaigns and brand-building activities such as advertising and product development? How do we ensure that retention campaigns and marketing strategy work in sync?

In summary, while we have learned a lot about customer retention, management attention and academic research efforts need to be broadened beyond identifying customers who are most likely to churn. More work can be done in this crucial area, but there are several other areas very fertile for future work. It is clear that the customer retention problem is not going away. Researchers need to guide managers on how to deal with this issue successfully. Managers need to think about retention more broadly than simply a matter of 0:1 subscription renewals. We hope this paper serves as a useful starting point for both researchers and managers.

References

1. Ali ÖG, Aritürk U (2014) Dynamic churn prediction framework with more effective use of rare event data: the case of private banking. Expert Syst Appl 41(17):7889–7903
2. Ascarza E (2017) Retention futility: targeting high risk customers might be ineffective. Forthcoming at the J Mark Res
3. Ascarza E, Hardie BGS (2013) A joint model of usage and churn in contractual settings. Mark Sci 32(4):570–590
4. Ascarza E, Iyengar R, Schleicher M (2016) The perils of proactive churn prevention using plan recommendations: evidence from a field experiment. J Mark Res 53(1):46–60
5. Ascarza E, Ebbes P, Netzer O, Danielson M (2017) Beyond the target customer: social effects of CRM campaigns. J Mark Res 54(3):347–363

6. Ascarza E, Fader PS, Hardie BGS (2017) Marketing models for the customer-centric firm. In: Wierenga B, van der Lans R (eds) Handbook of marketing decision models. Springer

7. Ascarza E, Netzer O, Hardie BGS (2017) Some customers would rather leave without saying goodbye. Forthcoming at Mark Sci

8. Athey S, Imbens GW (2016) Recursive partitioning for heterogeneous causal effects. Proc Natl Acad Sci 113(27):7353–7360

9. Bell DR, Ho TH, Tang CS (1998) Determining where to shop: fixed and variable costs of shopping. J Mark Res 35(3):352–369

10. Benedek G, Lublóy Á, Vastag Y (2014) The importance of social embeddedness: churn models at mobile providers. Decis Sci 45(1):175–201

11. Berger PD, Nasr NI (1998) Customer lifetime value: marketing models and applications. J Interact Mark 12(1):17–30

12. Berson A, Smith S, Thearling K (2000) Building data mining applications for CRM. McGraw-Hill, New York

13. Blattberg RC, Deighton J (1996) Manage marketing by the customer equity test. Harv Bus Rev 74(3):136–144

14. Blattberg RC, Getz G, Thomas JS (2001) Customer equity: building and managing relationships as valuable assets. Harvard Business Press, Boston

15. Blattberg RC, Kim B-D, Neslin SA (2008) Database marketing: analyzing and managing customers. Springer, New York

16. Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. J Mach Learn Res 3(Jan):993–1022

17. Braun M, Schweidel DA (2001) Modeling customer lifetimes with multiple causes of churn. Mark Sci 30(5):881–902

18. Bult JR, Wansbeek T (1995) Optimal selection for direct mail. Mark Sci 14(4):378–394

19. Castanedo F, Valverde G, Zaratiegui J, Vazquez A (2014) Using deep learning to predict customer churn in a mobile telecommunication network. Available at www.wiseathena.com/pdf/wa_dl.pdf. Accessed 4 Nov 2016

20. Chen K, Hu Y-H, Hsieh Y-C (2015) Predicting customer churn from valuable B2B customers in the logistics industry: a case study. IseB 13(3):475–494

21. Coussement K, Van den Poel D (2009) Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers. Expert Syst Appl 36(3):6127–6134

22. Crammer K, Kulesza A, Dredze M (2009) Adaptive regularization of weight vectors. Adv Neural Inf Proces Syst 22:414–422

23. Datta H, Foubert B, Van Heerde HJ (2015) The challenge of retaining customers acquired with free trials. J Mark Res 52(2): 217–234

24. de Fortuny J, Enric DM, Provost F (2013) Predictive modeling with big data: is bigger really better? Big Data 1(4):215–226

25. Delanote S, Leus R, Nobibon FT (2013) Optimization of the annual planning of targeted offers in direct marketing. J Oper Res Soc 64(12):1770–1779

26. Dodson JA, Tybout AM, Sternthal B (1978) Impact of deals and deal retraction on brand switching. J Mark Res 15(1):72–81

27. Fader P (2012) Customer centricity: focus on the right customers for strategic advantage. Wharton Digital Press, Philadelphia

28. Fader PS, Hardie BGS (2009) Probability models for customer-base analysis. J Interact Mark 23(1):61–69

29. Fader PS, Hardie BGS (2010) Customer-base valuation in a contractual setting: the perils of ignoring heterogeneity. Mark Sci 29(1):85–93

30. Fader PS, Hardie BGS (2012) Reconciling and clarifying CLV formulas. Available at http://brucehardie.com/notes/024. Accessed 4 Nov 2016

31. Fader PS, Hardie BGS, Lee KL (2005) 'Counting your customers' the easy way: an alternative to the Pareto/NBD model. Mark Sci 24(2):275–284

32. Fader PS, Hardie BGS, Shang J (2010) Customer-base analysis in a discrete-time noncontractual setting. Mark Sci 29(6):1086–1108

33. Forbes Insights (2014) Companies struggling to win customers for life, says new study by Forbes Insights and Sitecore. Available at http://www.forbes.com/sites/forbespr/2014/09/10/companies-struggling-to-win-customers-for-life-says-new-study-by-forbes-insights-and-sitecore/#7ac948205759. Accessed 2 Dec 2016

34. Gedenk K, Neslin SA (1999) The role of retail promotion in determining future brand loyalty: its effect on purchase event feedback. J Retail 75(4):433–459

35. Giudicati G, Riccaboni M, Romiti A (2013) Experience, socialization and customer retention: lessons from the dance floor. Mark Lett 24(4):409–422

36. Guelman L, Guillén M, Pérez-Marín AM (2015) Uplift random forests. Cybern Syst 46(3–4):230–248

37. Gupta S, Lehmann DR, Stuart JA (2004) Valuing customers. J Mark Res 41(1):7–18

38. Haenlein M (2013) Social interactions in customer churn decisions: the impact of relationship directionality. Int J Res Mark 30(3):236–248

39. Haenlein M, Libai B (2013) Targeting revenue leaders for a new product. J Mark 77(3):65–80

40. Handley L (2013) Customer retention: brave new world of consumer dynamics. Mark Week Online Ed 21. Available at https://www.marketingweek.com/2013/03/20/customer-retention-brave-new-world-of-consumer-dynamics. Accessed 4 Nov 2017

41. Hill S, Provost F, Volinsky C (2006) Network-based marketing: identifying likely adopters via consumer networks. Stat Sci 21(2): 256–276

42. Hogan JE, Lemon KN, Libai B (2003) What is the true value of a lost customer? J Serv Res 5(3):196–208

43. Huang B, Kechadi MT, Buckley B (2012) Customer churn prediction in telecommunications. Expert Syst Appl 39(1):1414–1425

44. Kosinski M, Stillwell D, Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci 110(15):5802–5805

45. Kumar V, Bhagwat Y, Zhang X (2015) Regaining 'lost' customers: the predictive power of first-lifetime behavior, the reason for defection, and the nature of the win-back offer. J Mark 79(4): 34–55

46. Labrecque JS, Wood W, Neal DT, Harrington N (2016) Habit slips: when consumers unintentionally resist new products. J Acad Mark Sci 45(1):1–15

47. Lemmens A, Croux C (2006) Bagging and boosting classification trees to predict churn. J Mark Res 43(2):276–286

48. Lemmens A, Gupta S (2017) Managing churn to maximize profits," Available at SSRN: https://ssrn.com/abstract=2964906 or https://doi.org/10.2139/ssrn.2964906

49. Leone R, Rao VR, Keller KL, Luo AM, McAlister L, Srivastava R (2006) Linking brand equity to customer equity. J Serv Res 9(2): 125–138

50. Lewis M (2005) Incorporating strategic consumer behavior into customer valuation. J Mark 69(4):230–238

51. Lim M, Hastie T (2015) Learning interactions via hierarchical group-lasso regularization. J Comput Graph Stat 24(3):627–654

52. Little JDC (1970) Models and managers: the concept of a decision calculus. Manag Sci 16(8, Application Series):B466–B485

53. Martens D, Provost F, Clark J, Fortuny EJ d (2016) Mining massive fine-grained behavior data to improve predictive analytics. MIS Q 40(4):869–888

54. Matz SC, Netzer O (2017) Using big data as a window into consumers' psychology. Curr Opin Behav Sci 18:7–12

55. McCarthy DM, Fader PS, Hardie BGS (2017) Valuing subscription-based businesses using publicly disclosed customer data. J Mark 81(1):17–35

56. Montoya R, Netzer O, Jedidi K (2010) Dynamic allocation of pharmaceutical detailing and sampling for long-term profitability. Mark Sci 29(5):909–924

57. Musalem A, Joshi YV (2009) Research note-how much should you invest in each customer relationship? A competitive strategic approach. Mark Sci 28(3):555–565

58. Neslin SA (2013) Dynamic customer optimization models. In: Coussement K, De Bock KW, Neslin SA (eds) Advanced database marketing. Gower Publishing Limited, Surrey

59. Neslin SA, Gupta S, Kamakura W, Lu J, Mason CH (2006) Defection detection: improving the predictive accuracy of customer churn models. J Mark Res 43(2):204–211

60. Nitzan I, Libai B (2011) Social effects on customer retention. J Mark 75(6):24–38

61. Nitzan I, Libai B (2013) If you go, I will follow … social effects on the decision to terminate a service. GfK Mark Intell Rev 5(2):40–45

62. Nobibon F, Leus TR, Spieksma FCR (2011) Optimization models for targeted offers in direct marketing: exact and heuristic algorithms. Eur J Oper Res 210(3):670–683

63. Oliver RL, Rust RT, Varki S (1997) Customer delight: foundations, findings, and managerial insight. J Retail 73(3):311–336

64. Ovchinnikov A, Boulu-Reshef B, Pfeifer PE (2014) Balancing acquisition and retention spending for firms with limited capacity. Manag Sci 60(8):2002–2019

65. Pennebaker JW, Francis ME, Booth RJ (2001) Linguistic inquiry and word count: LIWC 2001. Lawrence Erlbaum Associates, Mahway

66. Peppers D, Rogers M (2004) Managing customer relationships: a strategic framework. John Wiley & Sons, Hoboken

67. Perlich C, Dalessandro B, Stitelman O, Raeder T, Provost F (2014) Machine learning for targeted display advertising: transfer learning in action. Mach Learn 95(1):103–127

68. Perro J (2016) Mobile apps: what's a good retention rate? Localytics, March 28. (Available at http://Info.Localytics.Com/Blog/Mobile-Apps-Whats-A-Good-Retention-Rate. Accessed 4 Nov 2016

69. Provost F, Fawcett T (2013) Data science for business: what you need to know about data mining and data analytic thinking. O'Reilly Media

70. Provost F, Dalessandro B, Hook R, Zhang X, Murray A (2009) Audience selection for on-line brand advertising: privacy-friendly social network targeting. In: Proceedings of the Fifteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 707–716

71. Putter H, Fiocco M, Geskus RB (2007) Tutorial in biostatistics: competing risks and multi-state models. Stat Med 26(11):2389–2430

72. Rand W, Rust RT (2011) Agent-based modeling in marketing: guidelines for rigor. Int J Res Mark 28(3):181–193

73. Reinartz W, Thomas JS, Kumar V (2005) Balancing acquisition and resources to maximize customer profitability. J Mark 69(1):63–79

74. Rust RT, Oliver RL (2000) Should we delight the customer? J Acad Mark Sci 28(1):86–94

75. Rust RT, Zeithaml VA, Lemon KN (2001) Driving customer equity: how customer lifetime value is reshaping corporate strategy. Simon and Schuster, New York

76. Rust RT, Kim J, Dong Y, Kim TJ, Lee S (2015) Drivers of customer equity. Handbook of research on customer equity in marketing. Edward Elgar, pp 17–43

77. Schmittlein DC, Morrison DG, Colombo R (1987) Counting your customers: who are they and what will they do next? Manag Sci 33(1):1–24

78. Schrift RY, Parker JR (2014) Staying the course: the option of doing nothing and its impact on postchoice persistence. Psychol Sci 25(3):772–780

79. Schwartz EM, Bradlow ET Fader PS (2014) Model selection using database characteristics: Developing a classification tree for longitudinal incidence data. Mark Sci 33(2):188–205

80. Schweidel DA, Knox G (2013) Incorporating direct marketing activity into latent attrition models. Mark Sci 32(3):471–487

83. Schweidel DA, Bradlow ET, Fader PS (2011) Portfolio dynamics for customers of a multiservice provider. Manag Sci 57(3):471–486

82. Seitz P (2015) Apple music facing subscriber retention problems. Investor's business daily. Available at https://www.investors.com/apple-music-hitting-some-sour-notes. Accessed 4 Nov 2017

83. Springer T, Kim C, Azzarello D, Melton J (2014) Breaking the back of customer churn. Available at http://www.bain.com/publications/articles/breaking-the-back-of-customer-churn.aspx. Accessed 4 Nov 2016

84. Stahl F, Heitmann M, Lehmann DR, Neslin SA (2012) The impact of brand equity on customer acquisition, retention, and profit margin. J Mark 76(4):44–63

85. Statista (2016) Average monthly churn rate for wireless carriers in the United States from 1st quarter 2013 to 1st quarter 2016. Available at http://www.statista.com/statistics/283511/average-monthly-churn-rate-top-wireless-carriers-us/. Accessed 4 Nov 2016

86. Stauss B, Friege C (1999) Regaining service customers: costs and benefits of regain management. J Serv Res 1(4):347–361

87. Tamaddoni A, Stakhovych S, Ewing M (2016) Comparing churn prediction techniques and assessing their performance: a contingent perspective. J Serv Res 19(2):123–141

88. Taylor J, Tibshirani RJ (2015) Statistical learning and selective inference. Proc Natl Acad Sci 112(25):7629–7634

89. Thomas JS, Blattberg RC, Fox EJ (2004) Recapturing lost customers. J Mark Res 41(1):31–45

90. Tibshirani RJ (1996) Regression shrinkage and selection via the lasso. J R Stat Soc B Methodol 267–288

91. Vafeiadis T, Diamantaras KI, Sarigiannidis G, Chatzisavvas KC (2015) A comparison of machine learning techniques for customer churn prediction. Simul Model Pract Theory 55:1–9

92. van Baal S, Dach C (2005) Free riding and customer retention across retailers' channels. J Interact Mark 19(2):75–85

93. Venkatesan R, Kumar V (2004) A customer lifetime value framework for customer selection and resource allocation strategy. J Mark 68(4):106–125

94. Verbeke W, Martens D, Mues C, Baesens B (2011) Building comprehensible customer churn prediction models with advanced rule induction techniques. Expert Syst Appl 38(3):2354–2364

95. Verbeke W, Martens D, Baesens B (2014) Social network analysis for customer churn prediction. Appl Soft Comput 14:431–446

96. Verhoef PC (2003) Understanding the effect of customer relationship management efforts on customer retention and customer share development. J Mark 67(4):30–45

97. Verhoef PC, Donkers B (2005) The effect of acquisition channels on customer loyalty and cross-buying. J Interact Mark 19(2):31–43

98. Voss GB, Voss ZG (2008) Competitive density and the customer acquisition-retention trade-off. J Mark 72(6):3–18

99. Wang JC, Hastie T (2014) Boosted varying-coefficient regression models for product demand prediction. J Comput Graph Stat 23(2):361–382

100. Webb (2016) Apple to Revamp Streaming Music Service After Mixed Reviews, Departures. Available at http://www.bloomberg.com/news/articles/2016-05-04/apple-to-revamp-streaming-music-service-after-mixed-reviews-departures. Accessed 4 Nov 2016

101. Wood W, Neal DT (2009) The habitual consumer. J Consum Psychol 19:579–592

102. Zou H, Zhang HH (2009) On the adaptive elastic-net with a diverging number of parameters. Ann Stat 37(4):1733–1751