CUSTOMER DEFECTION: INSIGHTS FROM A MAJOR BRAZILIAN NEWSPAPER PUBLISHER

DESERÇÃO DO CLIENTE: PERCEPÇÕES DE UM EDITOR DOS PRINCIPAIS JORNAIS BRASILEIROS

Karine de Almeida Karam
Pontifícia Universidade Católica – RJ

Jorge Ferreira da Silva
Pontifícia Universidade Católica – RJ

Flávia de Holanda Schmidt
Pontifícia Universidade Católica – RJ

Jorge Manoel Teixeira Carneiro
Pontifícia Universidade Católica – RJ

RESUMO
A importância que a retenção de clientes e dos seus antecedentes tem para a lucratividade das empresas na atualidade foi a premissa sobre a qual foi realizado este estudo. O objetivo que o norteou foi: (i) identificar variáveis (de relacionamento, geográficas, socio-demográficas e relacionadas à lealdade) que possam ser relacionadas ao risco de cancelamento de clientes e (ii) determinar o perfil dos clientes com maior risco de desligamento. Foram coletados dados secundários sobre aproximadamente 36.000 clientes de um dos principais jornais brasileiros e utilizados modelos de regressão logística para analisar os dados. Com base na revisão de literatura realizada sobre retenção de consumidores e seus antecedentes, catorze variáveis foram preliminarmente identificadas como associadas ao cancelamento de clientes. O modelo final de estimação é composto de 11 variáveis e a tabela de classificação mostrou uma taxa de acerto geral de 75,3%. A utilização de um modelo como este pode permitir às firmas alocar de forma muito mais eficiente seus esforços destinados à retenção de consumidores, evitando assim que a adoção de ações uniformes sobre toda a base de clientes, além de definir com maior clareza para os gestores quais são os tipos de consumidores que devem ser priorizados como alvo. Algumas implicações teóricas decorrentes do estudo são ainda discutidas.

PALAVRAS-CHAVE
ABSTRACT

The recognition that customer retention and its antecedents thereof may be important drivers of profitability has underscored this present study. Our objective was (i) to identify variables (relationship-related, geographic, socio-demographic and loyalty-related) that might be associated with customer defection (the opposite of retention) and (ii) to determine the profile of (most likely to be) defectors. We collected secondary data about 36,000 customers of a major Brazilian commercial newspaper and applied logistic regression to analyze the data. Based on a literature review on customer retention and its antecedents, we preliminarily identified 14 variables that might be associated with customer churn. The final estimation model was composed of 11 variables and exhibited an overall hit rate (percentage of cases correctly classified as either churners or non-churners) of 75.3%. By applying such a model a firm may allocate in a much more efficiently manner its efforts towards customer retention instead of spreading such efforts uniformly across its customer base and, moreover, it might better define which types of customers to target in the first place. Some theoretical implications are also discussed.

KEYWORDS

INTRODUCTION

The ability to attract and retain customers has been recognized as a key driver of firm profitability especially in marketing contexts where a frequent repurchase pattern or a continuous customer provider relationship is observed (ATHANASSOPOULUS 2000; BHATTACHARYA 1998; COLGATE; DANAHER, 2000; RUSTER; ZAHORIK, 1993). Researchers and practitioners acknowledge that small changes in loyalty and retention (e.g., 5%) can yield disproportionately large changes in profitability (e.g., 25%-100% cf. REICHHELD; MARKEY; HOPTON, 2000; REICHHELD; TEAL, 1996).

Given such impact on profits, a change from a mere transactional to a relational paradigm has been advocated (SINGH; SIRDESHUMUKH, 2000). Relationship programs seek to increase customer lifetime value since, as has been argued by Vavra (1996), the longer a customer stays with a firm the more profitable he tends to become. He not only buys more but also becomes dependent of the firms’ products and services, which makes him less prone to defect to competitors’ even if they charge lower prices.

In Brazil almost 80% of the sales of quality newspapers (those directed to higher income customers) are derived from subscriptions. So, preventing churn – i.e., retaining their present subscribers – seems to be a more sensible approach to keeping sales up instead of relying on transactions at newsstands or merely searching for ever more subscribers.

This study has two main objectives: (i) to identify which variables are associated with churning and (ii) identify the profile of chuners.
The study is circumscribed to the customers of a single major Brazilian newspaper publisher, based in Rio de Janeiro. Some characteristics of the newspaper industry suggest it as an interesting case study for a deeper investigation of the phenomena of consumer defection: churn is an important issue to this industry (LEMMENS; CROUX, 2006) – as newspaper subscription is a continuously provided service –; the value of a customer is directly tied to the length of time of the relationship and subscriptions are usually fostered by short-term promotions (LEWIS, 2006) – so one can compare retention levels of promotionally-acquired consumers vs. others, since the former might have considerably lower lifetime values than the latter.

The study is organized as follows. After this introduction, we present a literature review on customer retention and its antecedents. Then, we discuss the methods employed for data collection, treatment and analysis. Findings are afterwards reported and discussed. Final remarks conclude the paper.

LITERATURE REVIEW

We organized our literature review around three main themes: profit impact of customer retention; association between relationship marketing and customer retention; and association among satisfaction, loyalty and retention. We thoroughly searched for material in three top-tier marketing journals – Journal of Marketing, Journal of Marketing Research and Journal of Consumer Research – and also articles in some other journals on a cross-reference basis.

PROFIT IMPACT OF CUSTOMER RETENTION

It has been argued that keeping present customers tends to be more profitable than attracting new ones (BERRY; 1995; NASH, 1993; PARVATIYAR; SHETH, 2000). Moreover, some researchers have contended a firm’s value and future prospects may be estimated from its customer base (GUPTA; LEHMANN, 2003; RUST; ZEITHAML; LEMON, 2000).

Some studies have suggested that the longer a customer stays with a firm, the greater the profits that accrue (REICHHELD, 1996). A study by Reichheld and Sasser Jr. (1990) indicated that profits could be raised by at least 25% and as much as 85% if a firm were able to reduce customer defection by just 5%. In Oliver’s (1999) study of 14 industries, a 5% increase in the customer retention rate was shown to lead to a very relevant increase – from 25% up to as much as 95% – in net present value (NPV). Considering the customer base as an asset of the firm has become a central point in customer management applications (LEWIS, 2006), enhancing the use of the customer life-value concept (CLV). CLV might be calculated using predictions of customer behavior over the long run (VENKATESAN; KUMAR; BOHLING, 2007).

In some industries, retention may be negatively associated with the frequency of use of promotions and acquisition discounts (BLATTBERG; NESLIN, 1990; NESLIN, 2002). Lewis’s (2006) study of the USA newspaper industry also showed that discount level is negatively related with the propensity to renew. In the wireless telecommunications industry, defection has been shown to happen more often during the first year, which could be related to the fact that several operators offer free cellular phone packages combined with one-year subscription. In fact, empirical findings suggest that a customer is more likely to churn when his telecom plan is cheaper ((LEMMENS; CROUX, 2006).

Association between relationship marketing and customer retention

Relationship marketing can be defined as attempts by firms to establish, maintain, and enhance relationships with customers and other parties at a profit so that the objectives of all the parties involved are met by mutual exchange and
fulfillment of promises (GRONRÖOS, 1991). It is a recognition of the long-term value of a customer and thus of the importance of the relationship.

Relationship marketing would involve a continuous and purposeful process of interactions – instead of just a number of independently-managed transactions – that help build stronger bonds between customers and the firm (GRÖNROOS, 2000; KOTLER; ARMSTRONG, 2005).

By managing the relationship with a customer, a firm can prevent defection and thus have the customer buy more, cost less to be served (given both parties’ accumulated experience from repeated transactions) and even spontaneously indicate new customers. Furthermore, if companies can more accurately predict which customers are most likely to depart, they can target incentives to those customers in order to induce them to stay. Firms would also save money and organizational resources by refraining from investing efforts in providing incentives to customers that would probably stay nonetheless (NESLIN et al., 2006). Therefore, it is important that firms develop the ability to predict churn with accuracy.

Other researchers confirm the importance of predicting churn to enable the elaboration of targeted retention strategies to limit customer losses (BOLTON; KANNAN; BRAMLETT, 2000; GANESH; ARNOLD; REYNOLDS, 2000; SHAFFER; ZHANG, 2002). Lemmens and Croux (2006) suggests, for instance, that specific incentives may be offered to the most risky customer segments, in order to try to keep them in the customer base of the firm.

In the context of relationship marketing, the emphasis has shifted toward measuring the value of the customer base; understanding the impact of marketing expenditures on customer value; and actively employing marketing actions to maximize customer value and, thus, firm value (WEBSTER, 1992).

Satisfaction, Loyalty and Retention

As their expectations are exceeded, met or frustrated, customers may experience several different levels of satisfaction or of disappointment after each service encounter (LOVELOCK; WRIGHT, 2001). Yi (1990) conceptualizes satisfaction as the attitudinal judgment made after the purchase or after a series of interactions between the customer and the product. Satisfaction is a result of the perception by the customer that the consumption experience has fulfilled some need, desire or objective in a pleasant fashion or at a favorable performance level as compared to some known reference or expected pattern.

Reichheld (1996) defines loyalty in terms of the permanence of the client in the company for a long time. For Gerpott, Rams and Schindler (2001), loyalty would be a favorable attitude of a customer towards a firm and its offers.

Oliver (1999) understands loyalty as a sequence of phases. First, the customer becomes cognitively loyal, then affectively loyal, and moves on to become conatively loyal and ultimately actively loyal. Cognitive loyalty is directed toward the brand – given its attribute performance levels – and is considered a shallow state, such that in this phase satisfaction is not yet processed. When satisfaction starts mediating the repurchase intentions of the customer, but there is still room for switching brands, loyalty is said to be affective. In conative loyalty, the intention of behavior is influenced by repeated episodes of positive affect toward the brand. However, intention of behavior may not always turn into realized action. Active loyalty, on the other hand, represents the motivated intention that turns into readiness to act, despite possible obstacles.

Caruana (2004) makes a distinction between loyalty and retention. While loyalty is a psychological state of the customer, retention is a strategy implemented by a firm. Although distinct, the two concepts are related and are also associated with
satisfaction. Szymanski and Henard’s (2001) meta-analysis showed that satisfaction has a positive impact on self-reported customer loyalty; Jones, Motherbaugh and Beatty (2000) found a positive association between satisfaction and customer retention; and Bansal and Taylor (1999) found a positive association between retention and profitability.

The existing relationships among some constructs and loyalty have been addressed by many researches but there is still no clear consensus among the findings. While loyalty and repurchase intention has been found to be positively associated with satisfaction (BLOEMER; KASPER, 1995; LAM et al., 2004), trust (AGUSTIN, SINGH, 2005; GARBARINO; JOHNSON, 1999), and value (SIRDESHMUKH; SINGH; SABOL, 2002), the link between satisfaction and actual customer loyalty may be complex (JONES; SASSER, 1995; REICHHELD, 1996). Some studies suggested that the relationship between the two constructs might be moderated by some factors such as time of relationship, product usage, variety seeking behavior, switching costs, consumer knowledge, and socio-demographics (BOLTON, 1998; BOWMAN; NARAYANDAS, 2001; CAPRARO; BRONIARCZYCK; SRIVASTAVA, 2003; HOMBURG; GIERING, 2001; JONES; MOTHERSBAUGH; BEATTY, 2001; MITTAL; KAMAKURA, 2001).

CONCEPTUAL MODEL

Our interest in this study is to uncover which variables would better predict whether a customer tends or not to defect and, moreover, what is the profile of would-probably-be defectors.

The appropriate selection of explanatory variables has been addressed as an important issue in the predictive accuracy of consumer defection models (NESLIN et al., 2006). The authors stated that academics and practitioners tend to rely on theory, factor analysis and cluster analysis – in this order – to select variables for churn modeling.

Conceptual and empirical studies have suggested that several types of variables would affect the probability of defection:

- Demographic variables, such as income and household size have been argued to lead to statistically significant improvements in model fit estimation (GUPTA; CHINTAGUNTA, 1994);
- The existence and level of relationship between customers and the firm, such as participation in loyalty programs that provide economic rewards, have been suggested to lengthen customer relationships and to enhance customer share, while time of relationship would have a positive effect on customer retention (VERHOEF, 2003);
- Gilly and Gelb (1982) empirically supported the hypotheses that the way complaints are answered does affect satisfaction/dissatisfaction, thus reinforcing the importance of controlling for complaints when conceiving of retention strategies;
- Company interaction data, such as calls to the customer service center, together with behavioral and geographic variables, have also been argued to affect defection (NESLIN et al., 2006);
- Lemmens and Choux (2006), in a study of the wireless telecommunications industry, suggested that relationships tend to be much stronger for younger customers than for older ones, indicating that some demographic variables may have an impact on the probability of defection; the authors also found evidence that the base cost of the telecom plan was negatively related with retention, indicating that promotional campaigns may also affect defection.

Even though we were limited in this study to using variables that were part of a database of the firm, we nonetheless resorted to theory in order to choose from the several available variables. In the predictive model developed for this research, we selected 14 explanatory variables, which as a set cover several areas of theoretical influence on
the propensity to defect. Specifically, we chose nine relationship-related variables, one geographic variable, three socio-demographic variables and one loyalty related variable. These variables and their operationalizations are presented in Table 1. Only 11 variables were retained after the estimation of the model. The dependent variable in our model is the status of the customer, which can take on two distinct levels: active or inactive (i.e., defector).

### TABLE 1
Operationalization of the variables in the predictive model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Operationalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationship - related variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. subscription period</td>
<td>period of subscription contracted by the customer</td>
<td>year, semester, quarter, month</td>
</tr>
<tr>
<td>2. sales channel</td>
<td>form of contact employed to acquire the customer</td>
<td>agent, old customer (i.e., migrated from a previous database), internet, indicated by a friend, active telemarketing, receptive telemarketing, other, nil</td>
</tr>
<tr>
<td>3. payment method</td>
<td>form of payment chosen by the customer</td>
<td>payment at the bank, credit or debit card, automatic bank transaction, programmed bank installment</td>
</tr>
<tr>
<td>4. time of relationship</td>
<td>for active clients, for how long he/she has been a client; for inactive clients, how long he/she stayed as a client</td>
<td>0-3 months, 3 - 6 months, 6 - 12 months, 1-2 years, 2-3 years, more than 3 years</td>
</tr>
<tr>
<td>5. quantity of aggregated products</td>
<td>quantity of aggregated products</td>
<td>up to 3, 4-7, 8-10, more than 10</td>
</tr>
<tr>
<td>6. lifetime value</td>
<td>category of client according to the following formula: (time of relationship + expected future time of relationship) * (monthly margin)</td>
<td>platinum, diamond, silver, bronze, tin, nil</td>
</tr>
</tbody>
</table>
7. aggregated product purchase * whether the customer has or not purchased any additional product (e.g., book, guide, CD, collectible) yes, no

8. advertisement purchase* whether the customer has or not bought any ad yes, no

9. complaint whether the customer has or not filed any complaint thru customer service or the internet yes, no

Geographic variable

10. region neighborhood or city where the customer receives his/her newspaper, aggregated according to distribution costs several geographical regions in Rio - and in Brazil

Socio - demographic variables

12. sex* sex man, woman, nil

13. age bracket age bracket less than 20, 20-25, 26-30, 31-35, 36-40, 41-50, 51-60, 61-80, over 80, nil

14. socio-economic level aggregated score composed of several social and demographic variables provided by a public research institute solidity, security, equilibrium, stability, survival, privation, subsistence, nil

Loyalty - related variable

15. participation in frequent-user programs whether the customer has or not participated in some relationship initiative, such as, discounts, bonuses, exclusive promotions

* dropped after estimation of the final model
nil = field not filled in the database
DATA AND METHODS

The unit of analysis is the individual customer, for which we collected several characterization variables (relationship-related, geographic, socio-demographic and loyalty-related variables, as well as client status – i.e., active vs. inactive). Note that we deliberately chose to circumscribe our scope to customer persons, leaving out client firms which are also a group of subscribers. We believe that this delimitation helps to homogenize the sample and avoids including a possibly confounding influence. The population of interest was composed of all (person) subscribers and ex-subscribers to the newspaper. Although the publishers produce and sell two newspapers, this study covered only subscribers to one of them – a quality paper that dominates 36% of volume market share in the city of Rio de Janeiro, 81% of which is on a subscription base as of 2005. Having acquired such a large percentage of the industry customers, it would seem reasonable for the publishers to concentrate on retaining their current customers instead of just fighting to attract new ones.

The firm’s database contained information on all presently active subscribers who subscribed from 1976 on. For those who cancelled the subscription, the database contains information only about subscribers who defected in July 1995 or later. We defined a window for analysis which covered subscribers who defected between June 2004 and May 2005 as well as active subscribers as of May 2005. We deemed this window to be large enough to run the statistical analysis and reasonably representative of the (present and future) population of interest. From a total of 230,858 active subscribers and 105,524 defected subscribers in the window of analysis, we randomly extracted two sub-samples – an estimation sub-sample with 35,549 cases and a hold-out sub-sample with 4,796 cases. In each sub-sample the proportion of active vs. inactive subscribers was about the same as that in the whole original sample (the little differences come from the removal of cases due to data errors as described hereinafter). The initial estimation sample had 36,000 subscribers, but cases that exhibited missing data on the following variables were removed from the sub-samples: payment method and subscription period. Also subscribers registered as below 20 or over 80 years of age were removed since one would suspect that there might have been a mistake in data input. The final estimation sample contained 35,549 cases – 24,492 active and 11,057 inactive subscribers. Following population distribution, both estimation and hold-out sample contain 69% of active subscribers and 31% of inactive ones.

We chose logistic regression in order to predict group membership (i.e., ongoing vs. defected customer) and to identify which variables seem to be associated with the likelihood of defection. In cases such as this – where there are only two levels for the dependent variable and assumptions of multivariate normality and of equal variance-covariance matrices across groups are not met –, logistic regression is to be preferred to discriminant analysis (HAIR et al., 2005). Moreover, Neslin et al. (2006) have argued that logistic regression, as compared to decision trees, would be a more competitive methodological approach in terms of predictive accuracy.

DATA ANALYSIS

SPSS™ 13.0 was used to run the analysis. Stepwise forward procedure was chosen since it minimizes the number of variables while maximizing the precision of the logit model.

The baseline model – against which one has to compare the improvement provided by the final predictive model – considers one single constant, i.e., it would be assumed that all subscribers would have canceled. The hit rate of the baseline model was just 31% (Table 2).
Subscription period was the variable with the highest Wald statistic and, as such, was chosen to be the first one to enter the model. The next variables inserted were respectively time of relationship and complaint. These three first variables contributed 92.1% explanatory power for the model. The stepwise analysis with the Wald statistic took 11 steps until the final model was reached. The coefficients were all statistically significant at each sequential step.

So, out of the 14 original variables in our conceptual model, 11 were kept. The three unused variables were: sex, aggregated product purchase, and advertisement purchase.

As each additional predictor was inserted, there could be observed a decrease in the likelihood statistic (–2log) and an increase in pseudo R² measures. In the last step the value of Nagelkerke’s pseudo R² indicates 29% improvement in the explanatory power of the final model (R² = 40.0%) from the base line model (R² = 31%) (see Table 3). The Hosmer and Lemeshow measure of overall fit (see Table 4) indicates no statistically significant difference between observed and predicted classifications of cases in all estimated models with two or more predictors. The overall hit rate (cases correctly classified) of the final model is 75.3% and hit rates for each individual group are also high: 77.4% for defectors and 74.3% for ongoing customers. It is worth noticing that the final model exhibits a hit rate 2.4 times as high as the baseline model (75.3% vs. 31.1%). However, it should also be noticed that after the ninth step hit rates stabilize and there is very little improvement in pseudo R². This suggests dropping the two predictors inserted last (respectively, participation in frequent-user programs and quantity of aggregated products); nonetheless these two variables were kept, not to increase the precision of the model, but rather to help establish the profile of defecting customers.

### TABLE 2

<table>
<thead>
<tr>
<th>Predicted Status</th>
<th>Correctly classified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Status</td>
<td>Inactive</td>
</tr>
<tr>
<td>Inactive</td>
<td>11,057</td>
</tr>
<tr>
<td>Active</td>
<td>24,492</td>
</tr>
<tr>
<td>Overall hit rate</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 3

<table>
<thead>
<tr>
<th>Step</th>
<th>-2 Log likelihood</th>
<th>Cox &amp; Snell R Square</th>
<th>Nagelkerke R Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35227.478</td>
<td>.220</td>
<td>.310</td>
</tr>
<tr>
<td>2</td>
<td>33727.999</td>
<td>.253</td>
<td>.355</td>
</tr>
<tr>
<td>3</td>
<td>33430.169</td>
<td>.259</td>
<td>.364</td>
</tr>
<tr>
<td>4</td>
<td>33157.816</td>
<td>.264</td>
<td>.372</td>
</tr>
<tr>
<td>5</td>
<td>32940.573</td>
<td>.269</td>
<td>.378</td>
</tr>
<tr>
<td>6</td>
<td>32711.454</td>
<td>.274</td>
<td>.385</td>
</tr>
<tr>
<td>7</td>
<td>32545.559</td>
<td>.277</td>
<td>.390</td>
</tr>
<tr>
<td>8</td>
<td>32397.244</td>
<td>.280</td>
<td>.394</td>
</tr>
<tr>
<td>9</td>
<td>32271.724</td>
<td>.283</td>
<td>.398</td>
</tr>
<tr>
<td>10</td>
<td>32221.754</td>
<td>.284</td>
<td>.399</td>
</tr>
<tr>
<td>11</td>
<td>32195.392</td>
<td>.284</td>
<td>.400</td>
</tr>
</tbody>
</table>

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.
b. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.
c. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.
d. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.
CUSTOMER DEFECTION: INSIGHTS FROM A MAJOR BRAZILIAN NEWSPAPER PUBLISHER

After the initial estimation, the model was validated by means of the hold-out sub-sample. The overall hit rate with the validation sub-sample was 75.5% and within each sub-group it was 76.1% and 75.3%. These rates were very similar to the ones obtained from the estimation sample and provide confidence in the model.

Some limitations of the method must be acknowledged. Since our data came from an already existing database, we were limited to using these variables, but not others that might have been suggested from theoretical reflection and which might have added predictive power. Moreover, the definition of variables to compose the final predictive model and also the pecking order (i.e., the order in which they were inserted into the model) might have been affected by multicollinearity issues (when two or more exogenous observation series are highly correlated and thus have about the same predictive power with respect to the endogenous variable, they are called multicolinear). Nevertheless, as we were interested in building a predictive model, rather than an explanatory model, multicollinearity would in fact not be a serious problem (although it might be a problem for explanatory purposes, since it affects the presence/absence and the weights of explanatory variables).

FINDINGS

The following 11 variables explained 75.3% of the likelihood of defection:

- seven relationship-related variables: subscription period, sales channel, payment method, time of relationship, quantity of aggregated products, lifetime value, complaint;
- one geographic variable: region;
- two socio-demographic variables: age bracket, socio-economic level;
- one loyalty-related variable: participation in frequent-user programs.

Three variables provided negligible predictive power (after the insertion of the previous 11) and were not included in the final model: aggregated product purchase, advertisement purchase, and advertisement purchase.

Together those 11 predictor variables provide a likelihood score of defection as shown in Table 5. All in all, most of the subscribers exhibit a low probability of defection: 62% of them have only a 20% chance of defecting.

<table>
<thead>
<tr>
<th>Step</th>
<th>Chi-square</th>
<th>df</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.000</td>
<td>4</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>139.350</td>
<td>7</td>
<td>.000</td>
</tr>
<tr>
<td>3</td>
<td>135.151</td>
<td>7</td>
<td>.000</td>
</tr>
<tr>
<td>4</td>
<td>187.903</td>
<td>7</td>
<td>.000</td>
</tr>
<tr>
<td>5</td>
<td>129.408</td>
<td>8</td>
<td>.000</td>
</tr>
<tr>
<td>6</td>
<td>134.869</td>
<td>8</td>
<td>.000</td>
</tr>
<tr>
<td>7</td>
<td>134.234</td>
<td>8</td>
<td>.000</td>
</tr>
<tr>
<td>8</td>
<td>120.247</td>
<td>8</td>
<td>.000</td>
</tr>
<tr>
<td>9</td>
<td>144.911</td>
<td>8</td>
<td>.000</td>
</tr>
<tr>
<td>10</td>
<td>142.347</td>
<td>8</td>
<td>.000</td>
</tr>
<tr>
<td>11</td>
<td>135.804</td>
<td>8</td>
<td>.000</td>
</tr>
</tbody>
</table>

TABLE 4

Hosmer and Lemeshow Test
TABLE 5
Defection likelihood for customers in the study’s database

<table>
<thead>
<tr>
<th>Likelihood range</th>
<th>active</th>
<th>defected</th>
<th>overall</th>
<th>% defection</th>
</tr>
</thead>
<tbody>
<tr>
<td>between 0.00 and 0.20</td>
<td>15,249</td>
<td>1,601</td>
<td>16,850</td>
<td>9%</td>
</tr>
<tr>
<td>between 0.20 and 0.40</td>
<td>4,763</td>
<td>1,656</td>
<td>6,419</td>
<td>18%</td>
</tr>
<tr>
<td>between 0.40 and 0.60</td>
<td>2,828</td>
<td>2,635</td>
<td>5,463</td>
<td>32%</td>
</tr>
<tr>
<td>between 0.60 and 0.80</td>
<td>1,512</td>
<td>3,848</td>
<td>5,360</td>
<td>54%</td>
</tr>
<tr>
<td>between 0.80 and 1.00</td>
<td>140</td>
<td>1,317</td>
<td>1,457</td>
<td>61%</td>
</tr>
</tbody>
</table>

The profile of most-likely defectors could also be derived from the model:

- chose a yearly subscription;
- were induced to subscribe by means of active telemarketing efforts;
- chose as payment method either direct payment at the bank or automatic debt to the bank account;
- newly acquired customer;
- had not purchased any aggregated product;
- LTV, a segmentation score that accounts for customer life time value, considering, for each customer, the permanence as a newspaper subscriber and the profit margin of the subscription;
- never filed a complaint;
- dwellers of either the low-income suburbs in the Rio de Janeiro city or of a state other than Rio de Janeiro state (note: the capital of the state of Rio de Janeiro is the city of Rio de Janeiro);
- young (below 30);
- from a lower socio-demographic stratum (the lowest four strata out of seven strata in the respective classification scale);
- participation in frequent-user programs, indicating whether the customer has ever participated in the newspaper relationship programs.

Some findings in this study are in line with other results from the literature. They provide support to the predictive power of logit models (NESLIN et al., 2006). The use of demographic, behavioral and customer interaction variables is also in line with the work of Neslin et al. (2006).

Furthermore, the fact that defectors tended to be younger than active subscribers confirms findings by Lemmens and Choux (2006) and Gupta and Chintagunta (1994).

The authors believe the insights delivered may be useful for defining new retention strategies. Results demonstrate that subscribers and non-subscribers had almost the same proportion of participation in loyalty programs, thus indicating that these programs are not being effective to retain customers. This result contrasts with Vefhoef’s (2003) suggestion that participation in loyalty programs tends to lengthen customer relationships. Or, one could interpret the finding as probable evidence of the lack of ability of the particular firm under study here to implement such programs.

FINAL DISCUSSION AND CONCLUSION

We conducted a single-firm study in order to investigate which variables could better explain the likelihood of defection of subscribers of a major Brazilian newspaper headquartered in Rio de Janeiro.
CUSTOMER DEFECTION: INSIGHTS FROM A MAJOR BRAZILIAN NEWSPAPER PUBLISHER

Janeiro. In the particular case of this newspaper, measures ought to be taken given that they had suffered a defection of around 30% subscribers in a one-year period.

The choice of a single case study makes sense both from a managerial and a theoretical standpoint. For managers of the newspaper, they were able to gain insight into what could explain defection, so that they could take measures to circumvent the loss. From a theoretical perspective, we suspected that this particular case might offer conceptual insights since managers told us that they were astonished with the high churning rate – especially because they believed they had taken several marketing measures that they expected would keep defection low. So, both for academics and for practitioners it would be important to understand why some generally accepted marketing actions did not seem to pay off in this particular case.

A logistic regression model was built from an estimation sample of 35,549 (current plus former) subscribers and 11 variables were found to be associated with the probability of defection. The overall hit rate of the model was 75.3% and it could explain as much as 40% of variation in the dependent variable (status of the customer, i.e., active vs. inactive).

From the specific variables associated with defection and the profile of defecting customers, some suggestions can be advanced for the newspaper publisher. First we will address current subscribers. Then we will discuss prospects.

As for current subscribers one can consider some actions to raise the probability of retaining them:

- packaging lower-cost offers (e.g., only weekend subscription instead of a seven-days-a-week package) especially for those young and more economically needy and offering such options in case they communicate their intention to cancel;
- inserting booklets with news and information about poor suburbs and other states in order to target at customers living in these places (although this would come at a cost for the publishers and should thus be carefully scrutinized);
- initiating follow-up actions with newly acquired customers in order to help them make full use of what the newspaper has to offer, solve possible doubts and, as a result, increase their retention rate.

Besides, one should also discount the future receivables by a rate that reflects the risk (probability) of desertion. If the publishers were to actively initiate retention actions, they should concentrate efforts on those customers who show a higher probability of defection (thus expecting to retain some of the would-be defectors) instead of spreading efforts evenly and ending up addressing customers that would not defect anyway. In fact, if customers were targeted at random by retention actions, around 69% of them would be those that would have probably remained with the company anyway. However, applying the model in order to identify most probable defectors would enable the firm to reach around 75% (vs. 31%) of probable defectors.

As for potential prospects one can consider:

- preferably addressing older (over 30) people who live in middle- and higher-class neighborhoods in Rio;
- offering them (not necessarily at the time of the subscription) some aggregated product as long as some advantage is included for current customers that would not be available for new ones;
- suggesting lower-cost packages (as the ones discussed before) for less favored people;
- actively assessing the degree of satisfaction as a way of letting customers spell out what they think instead of just walking away without complaining and, besides, giving the firm some valuable information on how to address and retain, or improve the relationship with, other (present and future) subscribers.

All in all, it would be interesting to investigate what actions could be taken that might make customers loyal – not necessarily actively loyal, but
at least affectively loyal so that they go on subscribing to the paper.

One should be careful when interpreting other predictive variables of the model. For example, the fact the probability of defection was found to be associate with choosing a particular payment method – either direct payment at the bank or automatic debt to the bank account – meets no theoretical support. It might be a spurious result, which can be used for predictive, but not for normative, purposes. Moreover, although customers attracted by active telemarketing efforts do seem to have a higher probability of defection should not lead one to immediately conclude that such efforts should be minimized or terminated. Rather, as long as the net present value of such actions (expected additional revenues minus respective serving costs minus acquisition costs) is expected to be positive – despite higher churning rates – they should be continued.

Some of the variables already available in database ought to be redefined or refined. Complaints, for instance, have been recorded as a dichotomous variable: “yes” for clients that have complained at least once and “no” for those who have never filed a compliant. However, determining whether the solution provided by the firm was or not satisfactory to the client would probably provide more meaningful insights about the link between transaction- or relationship-related variables and retention. Moreover, the database also fails to record perceptual measures of satisfaction – such indicators could be input by salespeople and call-center personnel who have contact with the customers.

Lifestyle evolution and increase in the popularity and reach of technology, as the internet must have also changed reading habits in the most recent years, so that the habit of reading of a traditional newspaper might become somehow difficult to keep. So, tracking whether clients have switched to reading newspapers online in lieu of their printed versions should also be investigated.

Table 6 presents these as well as some other variables that were not available in the company database, but which might be added in the future.

| TABLE 6 |
| Suggestions of variables to be added to the newspaper publishers’ database |

- reason for defection
- client default history
- qualitative information about complaints
- discount offer in the moment of renewal
- reason for contacting call-center
- educational level
- occupation
- marital status
- household income
- reading habits
- usage of the Internet as an information tool
Some managerial implications for industries that rely on subscription models to keep their clients, like cable TV, telecommunications (mobile or not), credit cards, magazines, insurance companies, might be suggested. Researchers following this stream have to define which variables should be considered for data modeling in the chosen industry. However, since variables that might lead to retention may vary from industry to industry the results of this present study may not be immediately generalizable, although some of them would surely seem to make theoretical sense in other business environments also.
REFERENCES


HOMBURG, C.; GIERING, A. Personal Characteristics as Moderators of the


NEAL, William D. Satisfaction is nice, but value drives loyalty. Marketing Research. [S. I.], v. 11, n. 1, p. 20–23, 1999.


