# Recognising Dangerous Drop Out Incidents as opposed to accidents to improve the efficiency of triggers reducing customer churn. Application to RFM customer segments of a Fast Moving Customer Goods retail chain.


Michel CALCIU
Associate Professor
IAE Business School, University Lille1
104 Av. Du Peuple Belge
59000 Lille
tel: +(33)03 20 12 34 09
e-mail: michel.calciu@univ-lille1.fr

Dominique Crié
Professor
IAE Business School, University Lille1
104 Av. Du Peuple Belge
59000 Lille
tel: (+33) 03 20 12 34 95
e-mail: dominique.crie@iae.univ-lille1.fr

Andrea Micheaux
Associate Professor
IAE Business School, University Lille1
104 Av. Du Peuple Belge
59000 Lille
tel: +33(0)3 20 12 34 61
e-mail:andrea.micheaux@iae.univ-lille1.fr



*Citation:*
CALCIU M, CRIE D., MICHEAUX A, “Recognising Dangerous Drop Out Incidents as opposed to accidents to improve the efficiency of triggers reducing customer churn. Application to RFM customer segments of a Fast Moving Customer Goods retail chain”, *Proceedings International Marketing Trends Conference 2015*

Recognising Dangerous Drop Out Incidents as opposed to accidents to improve the efficiency of triggers reducing customer churn. Application to RFM customer segments of a Fast Moving Customer Goods retail chain.


**Abstract:**

Drop Out Incidents (DOIs) can be identified by fixing threshold levels for the probability of being active on a global basis or by segment. In other words a customer whose probability of being active is below a given threshold is experiencing a DOI. These DOIs can then be used to calculate additional indicators such as Duration between DOIs (DDOI) to signal potential customer churn.

Contrary to previous studies which analyze the efficiency of stochastic models against a “dead or alive” outcome, we take the view that customers will tend to become occasional rather than lapse altogether. Rather than focusing on a single Drop Out Incident, we recognize that the first incident can be an “accident” or sudden death. We observe distributions of inter-DOI durations which reverse mirror the inter-purchase-span distributions.


**Introduction**

Stochastic models such as Pareto NBD and more recently BGNDB or MBGNBD that capture continuous time dynamic customer behavior that includes customer churn have been praised in marketing literature for their ability to calculate the customer’s probability of being active (alive). This probability of being active (pactive) integrates customer heterogeneity and takes into account each individual’s buying recency, frequency and anciency. As such it potentially constitutes an individually calibrated and rather powerful and flexible tool for predicting customer churn. Drop Out Incidents (DOIs) can be identified by fixing threshold levels for the probability of being active on a global basis or by segment. In other words a customer whose probability of being active is below a given threshold is experiencing a DOI. These DOIs can then be used to calculate additional indicators such as Duration between DOIs (DDOI) to signal potential customer churn.

Some authors have shown that stochastic models are efficient in predicting overall transaction volume in a given future period but that they are not all that useful for targeting customers in danger of lapsing. Conversely, a managerial hiatus such as a recency threshold has been shown to be more useful in identifying customers at risk than the stochastic model, albeit less efficient in forecasting transaction volume (Wübben and Wangenheim 2008). We show that by taking into account the underlying pattern of DDOIs, a stochastic model can be used both for forecasting transaction volume and for targeting follow-up resources to individual customers. Contrary to previous studies which analyze the efficiency of stochastic models against a “dead or alive” outcome, we take the view that customers will tend to become occasional rather than lapse altogether. Rather than focusing on a single Drop Out Incident, we recognize that the first incident can be an “accident” or sudden death. We observe distributions of inter-DOI durations which reverse mirror the inter-purchase-span distributions.

An application is given for a large data file comprising transaction data from 2011-01-01 to 2014-06-30. All 44 675 962 transactions have been extracted pertaining to 199 352 customers randomly selected from the top four RFM segments of a Fast Moving Customer Goods (FMCG) retailer. The top four RFM segments were defined by management heuristics as segments which contributed the highest overall sales value during the four months prior to the start of the study. The dataset therefore contained a large representative sample of the retailer’s highest value active customers, i.e. those customers the retailer was most anxious to retain.

**Stochastic consumer behaviour models and their use to detect Drop Out Incidents**

Probability or stochastic models consider observed behaviour as the result of an underlying stochastic process controlled by unobserved latent characteristics that vary across individuals. According to Gupta et al. (2006), they try to find a simple representation that describes and predicts observed behaviour instead of trying to explain differences between explained and observed behaviour as a function of covariables (as in regression models). They assume that a given behaviour varies across the population according to a probability law.

Historically, it is the NBD model (Negative Binomial Distribution - Ehrenberg 1959) that has been used on panel data in order to forecast the number of consumer goods purchased in a given time period while considering heterogeneity among individual customers. In this model the individual number of purchases follows a Poisson distribution with time, and the inter-individually heterogeneous purchase frequency follows a gamma distribution. This model does not take into account the customers' "inactivity" or "mortality". The first model that was able to estimate the probability for a customer to be «alive», the Pareto/NBD model, was developed by Schmittlein, Morrison and Colombo (1987) and extended by Schmittlein and Peterson (1994). It added to the NBD model a stochastic representation of the customer survival by postulating that at an individual level the survival probability diminishes exponentially (follows an exponential probability law) with time since last purchase and that the inter-individual heterogeneous mortality rate follows a gamma distribution. This model is particularly useful in non-contractual customer relationships where the firm cannot know when exactly a customer becomes inactive. The model can compute the probability for a customer to be active or "alive" by relying on his number of purchases (or the purchase frequency) and on time since his/her last purchase (or recency).

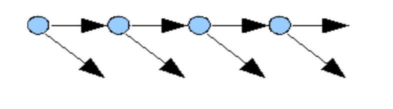
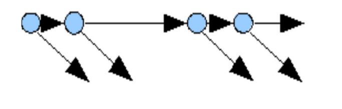
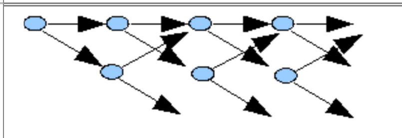
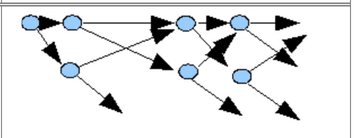
By fixing a threshold level for this probability, Drop Out Incidents (DOIs) can be detected. This probability acts as a threshold value that qualifies a person as a customer or a prospect (Reinartz and Kumar 2000, 2003). Yoo, Hanssens & Kim (2012) use the probability of being active to indicate whether an individual has purchased while in "acquisition" or "retention" state.

The probability function associated to the Pareto/NBD model is quite complex; it needs numerous evaluations of the gaussian hypergeometric function. Therefore its implementation is difficult not only because it is non-familiar to most marketing researchers, but also because it demands a great deal of computational resources. There are rather few articles that have applied it (Reinartz and Kumar 2003; Castéran, Meyer-Waarden and Benavent 2007). Fader, Hardie and Lee (2005) have therefore suggested an alternative model: the Beta-Geometric NBD. The main difference between these two models relies in the way attrition is modelled. While in the Pareto/NBD model, a customer can leave at any time and its survival probability diminishes with time since last purchase following an exponential distribution, in the BG/NBD model, departure (attrition) can only occur immediately after a transaction. Attrition probability follows individually a geometric distribution and the inter-individually heterogeneous attrition rate follows a beta distribution. In the BG/NBD model, customers who have not made a repeat purchase are considered as being active. In order to correct this aspect, Batislam et al. (2007) have modified the BG/NBD model by including the special case of attrition at zero moment; that is attrition just after a first purchase. In their model that can be called MBG/NBD (Modified BG/NBD) all other assumptions remain unchanged.

## A Collection of Customer Value calculation models as statistical packages

We have grouped together several stochastic models that represent dynamic customer behaviour and which have been classified by Jain and Singh (2002) as Customer Lifetime Value (CLV) calculation models. They are integrated by us as R packages and organized using three classification criteria: dynamic customer behavior contexts (contractual, non-contractual), chronology aspects of the transaction occasions (discrete and continuous time) and customer heterogeneity (deterministic and stochastic models). By crossing the first two dimensions several modeling contexts can be identified as shown in table 1.
For deterministic models, two packages for discrete transaction occasions have been developed: one for the contractual context the so-called retention model (ltvret) and one for the non-contractual migration model (ltvmigr).

*Table 1 - Adequation between activities and/or industries and the typology of customer relationships, examples*

| | | | | |
|---|---|---|---|---|
| Type of customer relationship | Contractual | Industry | Magazine subscriptions, Insurance policies, etc. | Credit cards, Mobile phone usage |
| | | Dynamic |  |  |
| | Non-contractual | Industry | Catalogue sales, Events attendance, Charity fund drives, | Retail purchases, Doctor visits, Hotel stays |
| | | Dynamic |  |  |
| | | | Discrete | Continuous |
| | | | Transaction occasions | |

*Source: adapted from Fader & Hardie (2009) p.63*

Stochastic models, unlike deterministic models, take into account customer heterogeneity. They are grouped together in one package (ltvmstochastic) that covers both discrete and continuous buying transaction occasion contexts. The models and their authors are given in table 2

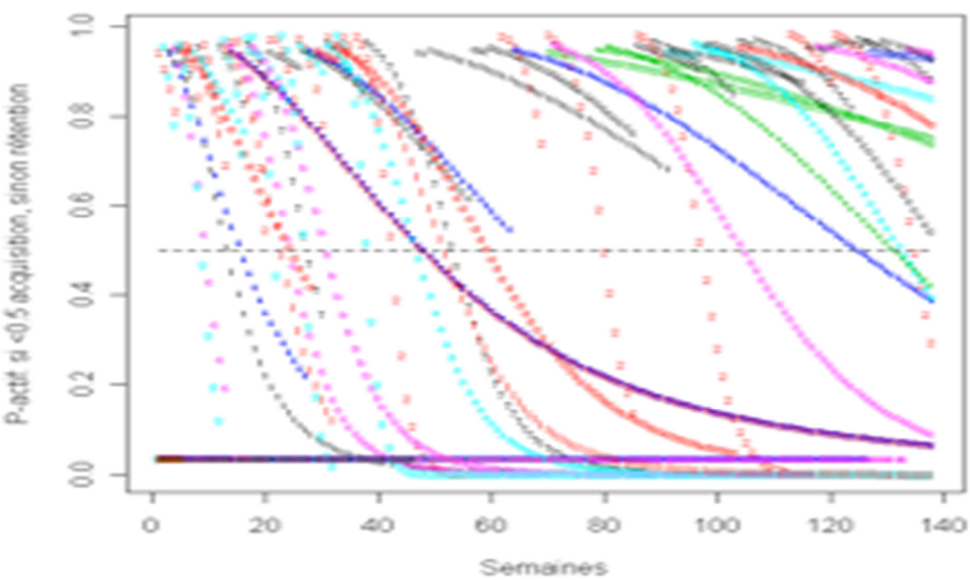
*Table 2 - Stochastic Models for various dynamic customer behavior contexts*

| | | | |
|---|---|---|---|
| Type of customer relation-ship | Contractual | *Shifted Betageometric* (**sBGD**-Fader & Hardie) | *Exponential Gamma* (**EGD** or Pareto Distribution of second kind) |
| | Non-contractual | Betageometric Betabinomial (**BG/BBD** - Fader, Hardie & Berger) | **Pareto/NBD** (Schmittlein, Morrison & Colombo); *Betageometric/NBD* (Fader,Hardie & Lee); *Modified Betageometric/NBD* (Batislam et al.)\| |
| | | Discrete | Continuous |
| | | Transaction occasions | |

All packages contain polymorphic functions using customer behavior data like purchase recency and/or frequency in order to calculate log-likelihood, purchase probabilities, probabilities of being active (noted Pactive, only for non-contractual contexts), expected transactions (ET, noted eYt), discounted expected transactions (DET) and CLV for a given lifetime or for the long term.
More recently the authors of some of these models, especially those who capture non-contractual purchasing behavior, have published an R package called BTYD as it implements "Buy 'Til You Die" situations, i.e. people buying until they die (become inactive as customers). The main models presented in the package are the Pareto/NBD, BG/NBD and BG/BBD models, which describe scenario of the firm not being able to observe the exact time at which a customer drops out.

**Choosing the critical threshold for the probability of being active**
At the individual level, the probabilities of each customer being active at a given time follow trajectories like the ones that can be observed in Figure 2.

*Figure 2 - Using a 0,5 probability of being active threshold to signal DOIs*




For frequently purchasing customers, the probability of being active diminishes at a higher rate during periods when they do not buy, in comparison with other less frequently purchasing customers. This trajectory depends on the number of purchases up to the moment of the calculation and on the time since the last purchase. The moments when the curves get interrupted correspond to new purchases. Using the resulting probability of being active which is unique for each client, we can distinguish whether customers are “alive” or not. This distinction requires a threshold level for the probability of being active which allows the tracing of a frontier between the two states. Traditionally (Reinartz and Kumar 2000, 2003), this threshold is fixed at an intuitive level of 0.5. A threshold of 0.5, especially in the case of frequently purchased consumer goods, does not necessarily ensure optimality of the number of correctly classified individuals (Calciu and Mihart, 2010). Using the algorithm suggested by Wübben and Wangenheim (2008) the threshold value can be improved. In our study a threshold of 0.93 appears to be optimal.

**Data preparation and validation**

The 199 352 customer file initially contained Pactive values calculated by a consulting firm for each of up to 34 successive fortnights. Pactive had been calculated by the consulting firm using a stochastic model on a sliding 4 month period of transaction data from 2010-09-01 to 2012-05-15. The Pactive values had been calculated on the first and on the 16th day of each month using the four months preceding transactions up to and including the previous day. The first Pactive value provided was 2011-01-01 and the last was 2012-05-16. We carried out data validation steps, calculated intermediate recency – frequency - anciency variables and determined which BGNBD model had been used to calculate Pactive.

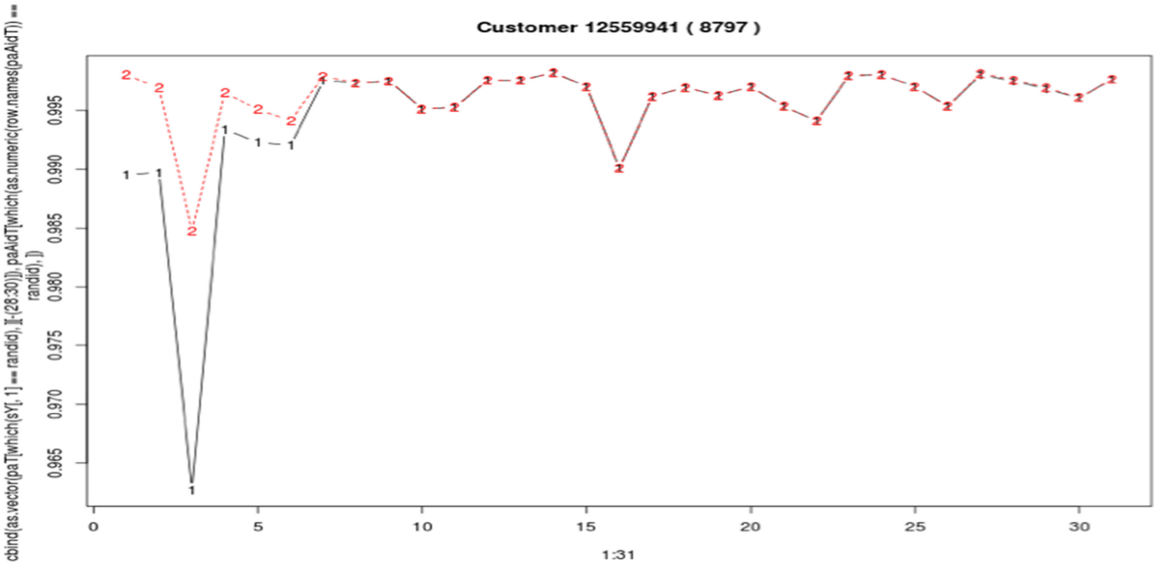
The formula of the BGNBD Probability of being active which had initially been used was

$$\frac{1}{1+\delta_{x>0}\frac{a}{x+b-1}\left(\frac{\alpha+T}{\alpha+t_x}\right)^{r+x}}$$

where (x, $t_x$ ,T) represent purchase history with x = number of purchases, $t_x$ = recency, T= anciency; (α,r,a,b) are the estimated beta distribution parameters defining the heterogeneous buying and dying probabilities. δ(x>0) means applies only to cases where x > 0 otherwise Pactive = 1

The comparison between the authors’ calculations and those by the consultants showed that the result for Pactive is identical except for the first eight forecasts due to the 2010 transaction data not being available to the authors (endnote 1)[i].

*Figure 1 comparison between Pactive calculated from the BGNBD model by the authors (line 1) and by the consulting firm (line 2).*




After presentation of the initial data integration and validation steps in a decision support document using RESTfull web services (Calciu and Micheaux, 2014), the FMCG retailer provided the additional transaction data up until 2014-06-30. The full dataset available for the study thus comprised firstly transaction data from 2011-01-01 to 2014-06-30, secondly customer data (RFM segment) valid on 2011-01-01, thirdly, Pactive values calculated by the authors for 84 fortnights from 2011-01-01 to 2014-06-16 (endnote 2) [ii].

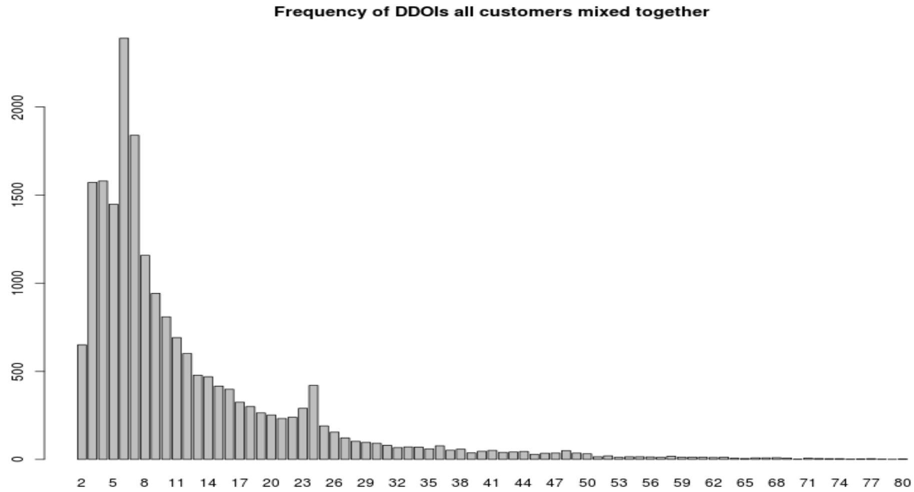
A Drop-Out Incident (DOI) Pactive threshold had been defined by the consulting firm as 0,93, based on some analysis of lapse behavior. A DOI in the current fortnight *f* was defined as *pactive (f) <= 0,93* AND *pactive (f-1) > 0,93*. i.e. a DOI occurs when the Pactive value for a customer in the current fortnight drops under the threshold from a value which had been above the threshold in the previous fortnight. Thus the first possible DOI occurrence is the second fortnight 2011-01-16 since the first Pactive value provided was 2011-01-01. It also follows that customers who begin the study period under the threshold are not counted as having DOIs until they have been over the Pactive threshold at least once. We drew a random sample of 10 000 customers from the 199 352 customer dataset. Of these, 6915 had at least one fortnight with pactive <=0.93. 1827 had one DOI, 1262 had two DOI's and 3820 had three or more DOI's (endnote 3)[iii]. The authors then calculated Duration (in fortnights) between successive DOIs (DDOI).

**Analysis of duration between drop-out incidents**

The research question was to determine whether DDOIs thus predicted by the stochastic model followed any underlying trend. Out of the 10 000 customers in the overall population sample, DDOI's could only be calculated for the 5082 customers with two or more DOIs. Among these, the modal average DDOI per customer was between 7 and 8 fortnights. There was no difference in the distribution of average DDOI per customer between the segments.

The overall distribution of DDOIs among all customers and all DDOIs mixed together follows a curve skewed to the left (figure 2).

*Figure 2: Distribution of DDOIs calculated from the DOIs of 5082 customers with at least 2 DOIs among the sample drawn from the overall population*

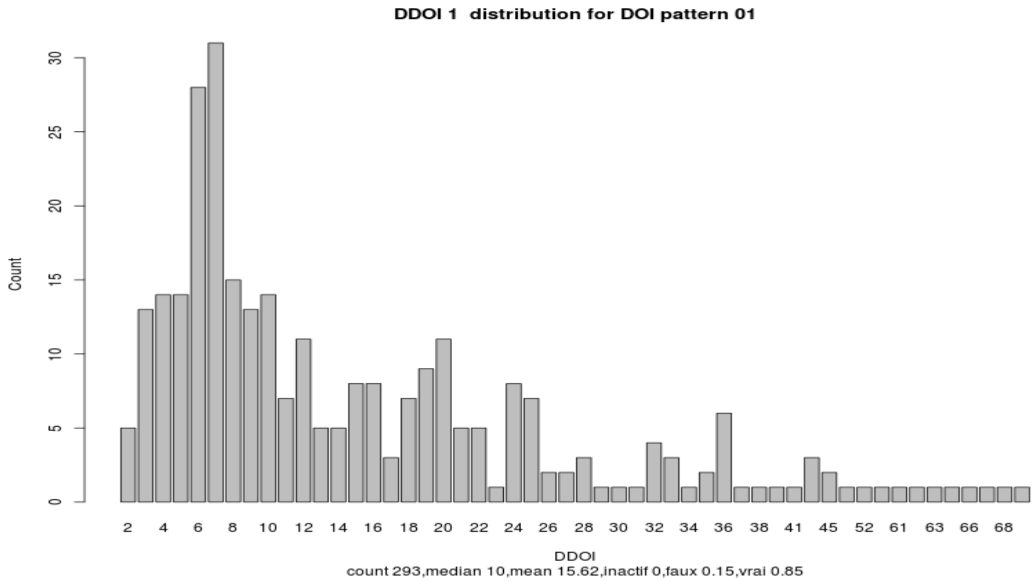



The DDOI distribution reads the opposite way to the DIHA (inter purchase interval) curve. In the DIHA distribution the less active customers would gather to the right of the curve. In the DDOI distribution the less active customers would be found to have a shorter duration between successive Drop-Out Incidents i.e. they would gather to the left of the curve. The modal DDOI was 6 fortnights.

To determine whether successive DDOI's followed the same pattern, we examined cohorts of customers with the same initial DOI profile. This allowed us to control for the bias due to the cut off study period of maximum 84 fortnights. Among the 10 000 sample, 293 customers had their first DOI in the second fortnight, i.e. they had a DOI profile of 01, where *pactive* was >= 0,93 in the first fortnight and *pactive* <0,93 in the second fortnight.

Among these, 293 had two or more DOI's allowing the calculation of at least one DDOI. For these customers, the modal value of the first DDOI value was 7 fortnights (figure 3). The distribution remained skewed to the left with the median DDOI of 10 fortnights and mean DDOI of 15,6 fortnights.

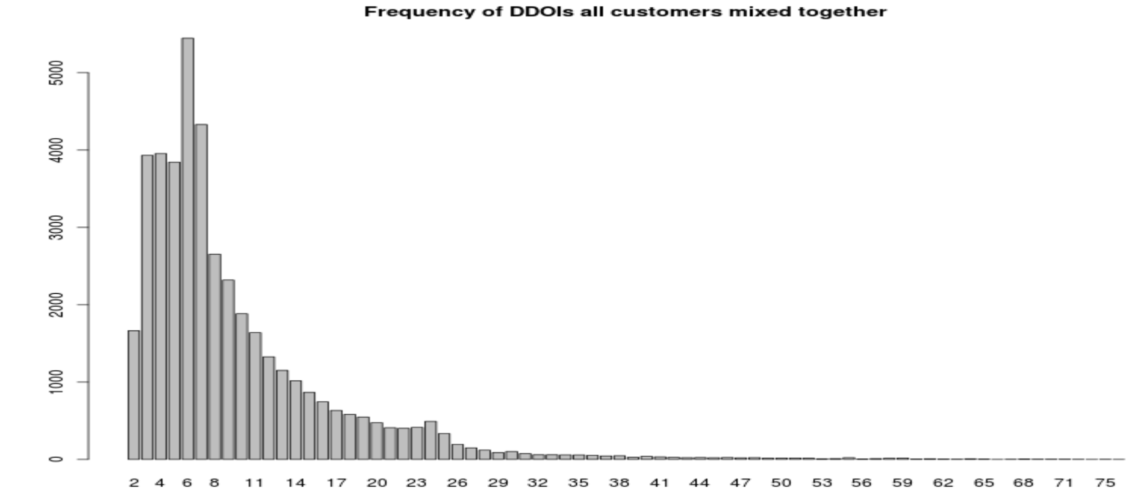
*Figure 3 Distribution of the first occurrence of DDOIs among 293 customers in the overall sample with a DOI profile beginning 01*




In order to examine further successive DDOI's we extracted from the overall population of 199 352 customers, 12 719 "Mortal" customers with at least one DOI while at least one Pactive value was missing. This means that "mortal" customers all spent more than one consecutive fortnight under the threshold and they all lapsed at one point during the study period. A lapse was defined by the retailer as no transactions during a four month period. Since Pactive was calculated over a four months sliding period, if customers lapsed they also had no Pactive value. "Mortal" customers therefore included those who lapsed definitively ("died") and those who "resuscitated", either to stay alive or to lapse again.

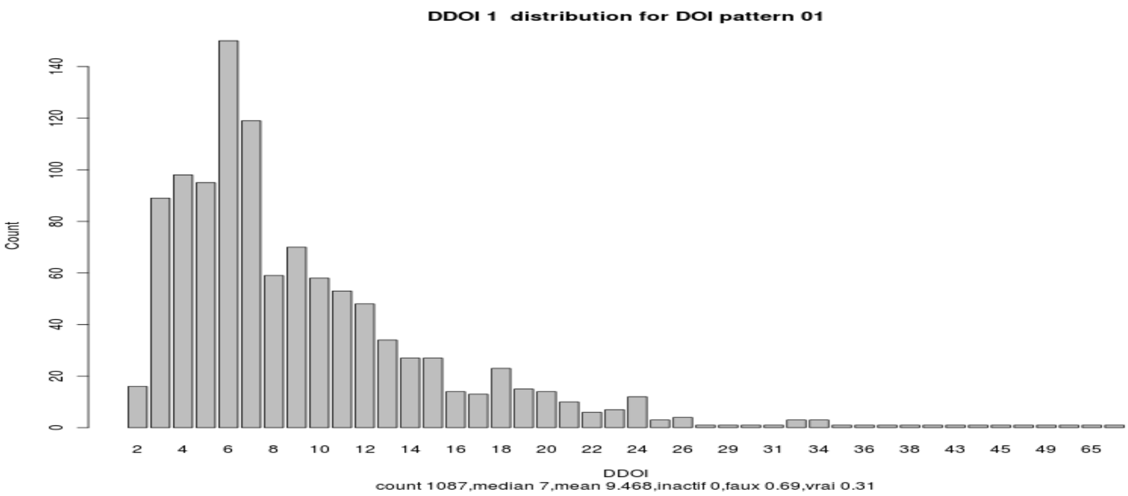
The "mortal" population all had at least one DOI. 3155 had one DOI, 2161 had two DOIs and 7403 had three DOIs or more. The Duration between DOIs (DDOI) could only be calculated for the 9564 customers with two DOI's or more. The modal average DDOI per customer was between 6 and 7 fortnights, this was shorter than for the overall population. There was no difference in the distributions of average DDOI per customer between the RFM segments. The DDOI modal value remained 6 fortnights (Figure 4). As for the overall population (Figure 2) the histogram of DDOIs for all "mortal" customers mixed together is rather leptokurtic or more concentrated around the modal DDOI.

*Figure 4 Distribution of DDOIs calculated from the DOIs of the 9 564 "Mortal" customers with at least 2 DOIs .*



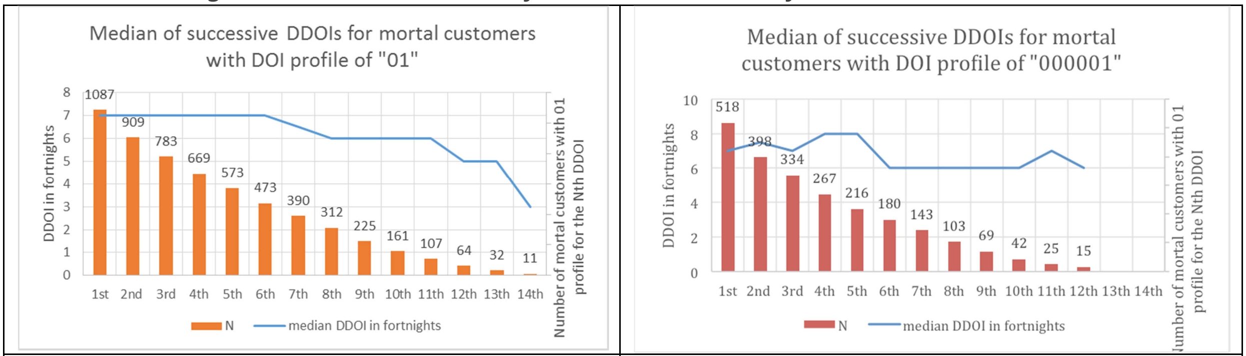

To determine whether successive DDOI's followed the same pattern, we examined the "mortal" population among cohorts of customers with the same initial DOI profile. This allowed us to control for the bias due to the cut off study period of maximum 84 fortnights. We selected "mortal" customers with their first DOI at the beginning of the study period, these 1195 customers had a DOI profile beginning 01. Among these, 1087 had 2 or more DOIs allowing the calculation of at least one DDOI. For these customers, we compared distributions of successive DDOIs.

*Figure 5 Distribution of the first DDOI among 1087 "Mortal" customers with a DOI profile beginning 01*




As for the overall sample population, the modal value of the first DDOI was 6 fortnights, the median value was 7 fortnights. The distribution was leptokurtic with a mean DDOI of 8.6 fortnights. Analysis of the successive fortnights showed the median to remain constant at 7 fortnights for the first six successive DDOI's. There remained 473 customers with "01" DOI profiles and 7 DOIs required to compute a sixth DDOI (Figure 6a). Beyond the sixth DDOI the median decreased due to the cut off period (end note 4)[iv]. The distribution remained leptokurtic throughout the successive DDOIs.

*Figure 6 - Median values of successive DDOIs of “mortal” customers*




| a) customers with “01” DOI profiles | b) customers with “000001” DOI profiles |
|---|---|

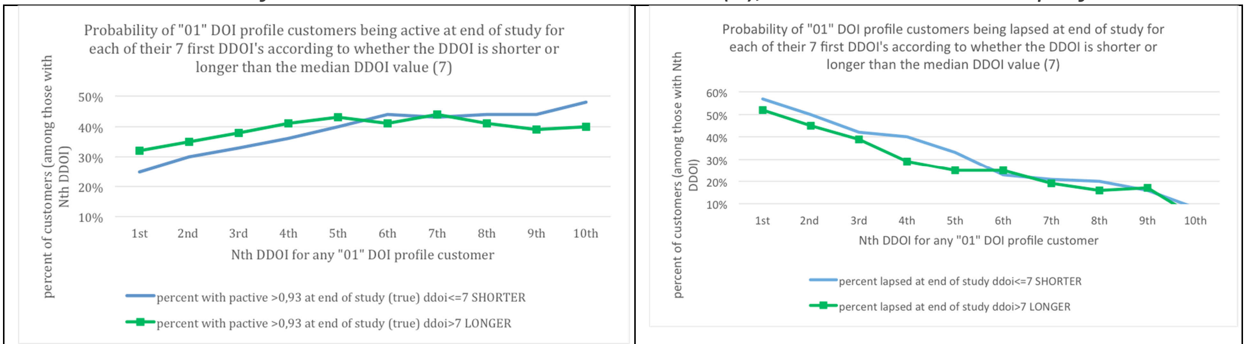
We reproduced the analysis for mortal customers with DOI profiles of 000001 (figure 6b), i.e. their first DOI was two months later than the “01” profile cohort shown in figure 6a. The median value remained constant within the initial observation period unaffected by the cutoff (7 for the first three successive DOIs and 8 for the 5th and 6th DOI). Again the distribution remained leptokurtic throughout the successive DDOIs.

Finally, we examined DDOI values with respect to the Pactive status of the customer at the end of the study; Pactive over the threshold; “true”, under the threshold “false” or no Pactive value i.e. the customer had lapsed altogether.

Customers in the cohort with DDOI values under the median value (left side of the distribution i.e. with relatively more frequent DOIs in a given period) are less likely to be active at the end of the study and more likely to have lapsed altogether, while customers with longer DDOI values i.e. right side of the distribution with relatively more sporadic DOIs are more likely to still be active at the end of the study period and less likely to have lapsed altogether.

We considered the first seven successive DOIs for the “01” DOI profile cohort and compared the Pactive values at the end of the study for customers with DDOIs below or above the median value of 7 fortnights. Figures 7a and 7b show the efficiency gain in predicting customers who are still active or who are completely lapsed (no purchase during the previous four months) at the end of the study from each successive DDOI value using the hiatus of the median DDOI.

*Figure 7 Efficiency gain in predicting active status and lapsed status at end of study for successive DDOIs of above versus below median value (7), Customers with DOI profile “01”*




| a) active status | b) lapsed status |
|---|---|

For the first 5 successive DDOIs (the first 6 successive DOIs), customers with DDOIs which are shorter than the median 7 fortnights are, as expected, less likely to be active at the end of the study and more likely to have lapsed at the end of the study than customers with DDOIs which are longer than the median 7 fortnights. The finding is reversed from the 6th successive DDOI in Figure 7a and the difference is no longer significant from the 6th DDOI in Figure 7b. This result could be due to the truncated study period; three to four DDOIs of 24 fortnights or less (90% of all mortals' DDOI's) can comfortably be observed in the 84 fortnight period. However since the median and mean values were constant for the first seven DDOIs calculated for the "01" DOI profile customers, the inverted (non–significant for lapse) efficiency from the 6th to 8th DDOI could be due to regular repetitive behavior such as second residence. A longer study period is necessary to confirm.

**Discussion and Conclusions**

*Managerial implications:* Marketing managers often use a recency hiatus for targeting follow-up messages to customers at risk of attrition. However, contrary to the Pactive value, recency does not take into account the individual pattern of transaction occasions previous to the most recent one. Direct and digital marketing actions which are personalized according to individual customers' caracteristics and behaviour are widely used by firms and have been shown to be more efficient than non-personalized messages (see Arora et al, 2008, for a review). Since Pactive more closely represents individual behavior than recency, it should be more appropriate and more efficient than recency for use by management in the personalization of the content and the timing of follow-up messages. However, there are several difficulties which can deter marketing managers from using a Drop Out Incident Pactive threshold for marketing operations. Our research addresses these difficulties.

Firstly, Pactive is given by the stochastic model as the probability of the customer being "alive" at the date of calculation. Notwithstanding analysis of lapse behavior to set the Pactive threshold, managers may find it difficult to brief their advertising agency to create a follow up message for targeting a customer with "a probability of less than 50 percent of being alive", or, in the case of the present study, "a probability of less than 93 percent of being alive". The duration between Drop Out Incidents (DDOI) is expressed in fortnights after the previous Drop Out Incident, a more intuitive and familiar concept than the Pactive threshold, which is closer to recency, all the while reflecting the individual pattern of transaction occasions.

Secondly, and more importantly, managers intending to use stochastic models for targeting follow-up marketing actions are faced with the difficult task of setting the Pactive threshold such that they follow-up the greatest amount of customers most at risk of attrition all the while avoiding the erroneous targeting of customers who have temporarily changed their purchase pattern without actually being in danger of lapsing. These are drop-out "accidents" rather than truly dangerous Drop Out Incidents.

Erroneous targeting potentially has negative consequences for the customer and for the retailer. One the one hand there is the risk of creating perceived pressure among customers receiving unfounded follow-up actions which they perceive as irrelevant (Micheaux 2011). On the other hand, the retailer risks losing revenue in offering discounts destined to incite leavers to reactivate but which are windfalls if the "resuscitated" customer was never really "dead". We propose that if managers base the follow-up hiatus on the duration between successive drop-off incidents instead of on the Pactive threshold, the danger of unnecessary marketing pressure and of unprofitable promotional effort is reduced.

By using the DDOI, the first DOI is not considered alone but in conjunction with the next successive DOI. The marketing action would only be considered on the occurrence of the second DOI. Customers with only one DOI can fall into one of two categories; those with no further transactions after the DOI; "sudden deaths" and those who "resuscitate" and remain above the Pactive threshold. In the FMCG retail sector a common explanation of a sudden death, is a customer who moves out of the catchment area of the stores. A reactivation message would not be relevant in this case. Other marketing programs would apply, such as those targeted to people moving into a new area. In the case of the permanently resuscitated

customer, the observed DOI can be considered as an isolated accident which would not require any action. Indeed, any such action could be considered by the customer as intrusive. Using the DDOI rather than the DOI as a trigger for action thus avoids following up isolated DOIs and sudden deaths, releasing marketing resources for the more dangerous successive DOIs.
The question for optimal resource allocation is therefore, how long is a “dangerous” DDOI? The median value DDOI provides a useful hiatus for management in detecting customers to follow up. Our research suggests that customers with two successive DOIs within a period of less than the median DDOI (in this study seven fortnights) are less likely to remain active and more likely to be in danger of ultimately lapsing altogether than those with successive DOIs separated by a longer period. Managers can apply this result to refine the target set of customers for reactivation while differentiating the marketing treatment according to the proximity of the customers’ DDOI to the median value. For example, in our case study, only customers with two successive DOIs in a period of less than seven fortnights would be offered a follow-up promotional incentive. Customers with DDOIs above but close to the median value could be put “under surveillance”. They may receive a lower value incentive and a satisfaction questionnaire. DDOI’s over “rare” values (in our study, 19 fortnights) could be ignored.
By using the method proposed in this paper, managers can concentrate resources on active customers at risk rather than wasting marketing budget on sudden deaths or accidental DOI. In addition to providing more efficient transaction volume forecasts than management heuristics, stochastic models offer the advantages of being simple to apply, requiring few variables and are intuitive when DDOIs are considered in conjunction with the Pactive threshold. Additional more in-depth analysis bringing in additional variables can complement this method among populations identified at risk. This will enable follow up actions to be more closely tailored for greater relevance and efficiency.

*Contribution and further research:* The literature has shown that stochastic models are efficient in predicting overall transaction volume in a given future period but that they are not all that useful for targeting customers in danger of lapsing. Conversely, a managerial hiatus such as a recency threshold has been shown to be more useful in identifying customers at risk than the stochastic model, albeit less efficient in forecasting transaction volume (Wübben and Wangenheim 2008). We show that by taking into account the underlying pattern of DDOIs, a stochastic model can be used both for forecasting transaction volume and for targeting follow-up resources to individual customers. Contrary to previous studies which analyze the efficiency of stochastic models against a “dead or alive” outcome, we take the view that customers will tend to become occasional rather than lapse altogether. Rather than focusing on a single Drop Out Incident, we recognize that the first incident can be an “accident” or sudden death. We observe a fractal pattern of distributions of inter-DOI durations which reverse mirror the inter-purchase-span distributions.
As regards the size of the data files that had to be dealt with this study can be regarded also as a "big data" exercise. At this stage of the research process, the study remains in an exploratory phase, further work will be undertaken to consolidate and formalize these encouraging findings.

## References


Arora N, X. Dreze, A. Ghose, J. Hess, R. Iyengar, Z. Zhang, et al (2008). Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Letters,* 19(July) 305-321.

Ehrenberg, A. S. C. (1959), The Pattern of Consumer Purchases, Applied Statistics, 8(1), p. 26-41.

Batislam, E.P., M. Denizel, A. Filiztekin (2007), Empirical validation and comparison of models for customer base analysis, *International Journal of Research in Marketing,* 24, 201-209

Castéran H., L. Meyer-Waarden, C. Benavant (2007), Une évaluation empirique des modèles NBD pour le calcul de la Valeur Actualisée Client dans le domaine de la grande distribution, *Actes du XXIIIème Congrès International de l'AFM* (31st May & 1st June), Aix-les-Bains, France.

Calciu M., A. Micheaux (2014), Words fly, Documents rest. RESTfull Decision Supportive Documents for Marketing Managers. Towards a paradigmatic change in the Interaction between Managers and Marketing Scientists, *43rd EMAC Conference*, June 3rd – 6th, Valencia, Spain.

Calciu M., Mihart I. (2010) "Modelling short and long term customer acquisition and retention effects in consumer goods sales using stochastic and econometric methods", *Journal of Marketing Trends*, 1,2 (April),74-86

Fader, P.S., B.G.S. Hardie, K. Lee. (2005), Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model. *Marketing Science*, 24(2), 275-84

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

Fader P.S., B.G.S. Hardie (2009) Probability Models for Customer-Base Analysis, *Journal of Interactive Marketing*, 23, 61–69

Fader, P.S., B. Hardie, P. D. Berger (2004) Customer-Base Analysis with Discrete-Time Transaction Data. Available at SSRN: http://ssrn.com/abstract=596801 or http://dx.doi.org/10.2139/ssrn.596801 accessed on 25/11/2014

Gupta, S., D.M. Hanssens, B. Hardie, W. Kahn, V. Kumar, N. Lin, N. Ravishanker & S. Siram (2006), Modeling Customer Lifetime Value, *Journal of Service Research,* 9, 2, 139-155

Jain D., S.S. Singh (2002). Customer Lifetime Value Research in Marketing. A Review and Future Directions. *Journal of Interactive Marketing*, 16(2), 34-46

Micheaux A., (2011). Managing E-mail Advertising Frequency from the Consumer Perspective. *Journal of Advertising*, 40 (4) 45-65

Reinartz W., V. Kumar (2000), On the Profitability of Long-Life Customers in a Non-contractual Setting: An Empirical Investigation and Implications for Marketing. *Journal of Marketing*, 64, 17-35

Reinartz, W., V. Kumar (2003), The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *Journal of Marketing*, 67, 17-35

Schmittlein, D., D. Morrison, R. Colombo (1987), "Counting Your Customers: Who They Are and What Will They Do Next?" *Management Science*, 33, 1-24

Schmittlein, D., R.A. Peterson (1994), "Customer Base Analysis: An Industrial Purchase Process Application," *Marketing Science*, 13 (1), 41-67

Wübben M., F. V. Wangenheim (2008), Instant Customer Base Analysis: Managerial Heuristics Often "Get It Right", *Journal of Marketing,* 72, (May), 82–93

Yoo S., D.M. Hanssens, H. Kim (2012), "Marketing and the Evolution of Customer Equity of Frequently Purchased Brands" October, Working Paper,http://www.anderson.ucla.edu/faculty/dominique.hanssens/content/CECPG_October9%202012.pdf, accessed on 08/09/2014

---

[i] 1. The retail company had changed their marketing information system on 2011-01-01. The 2010 transaction data had been archived and was not available to the authors at the time of the study.

[ii] 2. Due to the missing initial history, the consultant firm's pactive calculations were used for the first 26 fortnights, the remaining fortnights from 27 to 84 were the authors' calculations, which were of course still based on the consultant firm's initial estimation and the parameters of BGNBD which they had obtained.

[iii] 3. Of the 10 000 sample, only 6 had a pactive value <=0.93 at the start of the study period: 2011-01-01. As discussed, these were not counted as having a Drop Off Incident. These 6 customers were coded DOI = 0 as opposed to DOI missing for customers with no fortnights with pactive <=0.93.

[iv] 4. Overall, 70% of the DDOIs of "mortal" customers (all customers mixed together) lasted 10 fortnights or less, 90% of all DDOIs lasted 19 fortnights or less and 95,5% of all DDOIs lasted 24 fortnights or less. Thus three to four DDOIs could comfortably be observed in the cut off period of 84 fortnights.