
Eugenia Y. Huang
National Chengchi University
huang.eugenia@gmail.com

Chia-jung Tsui
National Chengchi University
chiajung@nccu.edu.tw

Wellington K. Kuan
National Chengchi University
kuan-wk@nccu.edu.tw

Hung-Shin Chen
National Chengchi University
terria0317@gmail.com

Ming-chin Hung
Soochow University
nhungg@scu.edu.tw

Abstract

In a challenging environment where the electronic marketplace is transparent and competitors are only a click away, web retailers are more vulnerable to customer attrition. Despite that customer retention was widely regarded central to business survival and growth and some studies attempted to find factors that influence customer retention, few measured customer retention quantitatively. Businesses have long been eager to know, at a given time how many customers can be considered as retained. Conversely, for how long an existing customer did not return to purchase can we regard him or her as non-retained? In this paper, we answer these questions using real purchasing data from a web retailer. We propose that the 80th percentile of maximum interpurchase times serve as the threshold to determine if a customer is retained. Our approach not only enables researchers to undertake longitudinal studies of customer repatronage behavior, but also helps practitioners monitor customer retention.

1. Introduction

Customer retention is particularly important in the transparent electronic business market [10]. In the environment of electronic marketplace, customers’ switching costs are low compared with the traditional economy, and web retailers are more vulnerable to customer attrition. Wirtz and Lihotzky [17] pointed out that the lower switching costs result mainly from the market transparency in the electronic business where the price comparison service is popular, from the lack of physical distance among customers, suppliers and competitors, and from the absence of personal retailer-customer relationships.

Customer retention has never been easy, and it is especially challenging in the electronic business market. Yet, businesses have to be prepared to meet the challenge, because customer retention gives rise to numerous benefits [1]. For example, Reichheld and Scheffer [13] believed that the repeat purchases due to customer retention result in sustainable profitability. Other benefits include cross-selling opportunities to the retained customers [6, 18]. Hennig-Thurau and Klee [7] concluded that customer retention opens up considerable cost-reduction potentials.

As a result, businesses competing in the digital economy are turning their attention and resources towards increasing retention of their customers. In this paper, we define retained customers of a web retailer as those who will keep doing business with the retailer. However, how can one decide if a customer is retained or not? For example, if a customer purchased something from a web retailer but did not repurchase this year, can we consider him or she non-retained? For how long an existing customer did not return to purchase can we regard him or her as non-retained? When we claim that a customer is non-retained if he or she had not repurchased from the web retailer for a whole year, how sure are we in saying this? The problem of such claim is that we can never be sure, as there could be some customers who would return to purchase even after many years. Furthermore, if we use a year or twelve months as a threshold to assess retention status, can the threshold of twelve months be applied to customers acquired at different times? In other words, can we apply the same threshold of certain period to all existing customers acquired at different times and decide their retention status with the same level of assurance?

We answer these questions in this paper by offering an analytical approach, based upon real purchasing data, to measure customer retention. In the following, the literature concerning customer retention is briefly reviewed and discussed. Then we illustrate our approach with an empirical study of a web retailer’s transaction data which span 6 years. And finally we conclude by discussing the validity...
and benefits of our approach for customer retention research and practice.

2. Literature review

Aspinall et al. [3] proposed the multifaceted nature of customer retention which involves behavioral and attitudinal aspects. They suggested some possibilities for each of the two aspects. For example, behavioral variables of customer retention include number of active customers, frequency of buying, recency of buying, size of expenditure, contract, etc. On the other hand, attitudinal aspect may contain psychological commitment/loyalty, trust, empathy, propensity to pay more, customer satisfaction, etc. Through an extensive survey of employees at managerial level, they found that behavioral rather than attitudinal aspect of customer retention was more frequently used. Another research by Hennig-Thurau and Klee [7] argued that customer retention does clearly not contain any attitudinal factors, and that customer retention is merely of behavior facet. As a result, in the following paper we focus customer retention on the behavior facet alone.

Another research stream discussed determinants of customer retention. Many studies [7, 15] pointed out that customer satisfaction has been treated as the essential premise for the retention of customers. Some [4, 5] addressed the importance of the link between customer satisfaction and retention. Specifically, in numerous models of customer retention, satisfaction has been explored as a key determinant in customers’ decisions to continue or terminate a business relationship. For example, Reichheld and Teal [12] found that satisfaction measures account for up to 40 percent of the variance in models of customer retention. Kotler summed this up when he stated: “The key to customer retention is customer satisfaction” [9].

Still other studies attempted to find useful strategies for customer retention. For example, by empirical research consisting of website observation and executive interviews, Wirtz and Lihtzky [17] discovered seven relevant retention strategies broadly used by B2C electronic businesses. They are: trust building, community, convenience, free services, individualization, contractual agreements, and technical integration. Yelkur [19] showed that high quality customer service and support is the key to improving customer retention. Yim et al. [20] reported that an effective CRM strategy is crucial to both customer satisfaction and customer retention. Ang and Buttle [2] investigated the associations between customer retention and a number of management processes. They found that excellence at customer retention is positively and significantly associated with the presence of documented complaints-handling processes. None of the other management processes is significantly associated with customer retention.

Customer acquisition is a related concept. Like customer retention, customer acquisition plays an important role in the B2C electronic business. In other words, they both contribute to the ongoing development of a company, as a company’s revenue sources are from both new customers (customer acquisition) and returning customers (customer retention). In literature, most studies emphasized customer retention more than customer acquisition. For example, Aspinall et al. [3] found that 54 percent of companies reported that customer retention was more important than customer acquisition. Lindgreen et al. [11] viewed the two from cost perspective and computed that it can be up to ten times more expensive to win a new customer than to retain an existing customer.

Customer acquisition is usually an action. When a company successfully acquires a customer, it means that the customer is doing business with the company now and was not before. In the paper, we define customer acquisition as the action of the first purchase by a new customer. Compared to customer acquisition, customer retention is not always an action. Rather, it is generally a state that the customer keeps doing business with the company as before [16]. In our analytical approach, we focus on defining the state which constitutes customer retention.

In summary, existing researches about customer retention in electronic business underlined its importance and provided some insights into possible strategies, but none made an attempt to measure customer retention quantitatively and none provided empirical results using real purchasing data. Aiming to fill the gap in the literature, this study conducts an empirical study to measure customer retention.

3. Research method

3.1. Data

This study utilizes data from a web retailer, which offers a comprehensive line of beauty products. The data include real individual-level purchasing data of more than forty five thousand customers over a six-year period, which is from 2005 to 2010. Established in 1999, the company represents one of a few web retailers which survive the 2001 internet bubble burst and bode well for the future.
3.2. Data analysis

First, customers are grouped according to the year they were acquired, i.e., the year they made their first purchase from the web retailer. Then, we observe the composition of new and returning customers over the years. Logically and practically, the number of returning customers can only decrease over the years since their first purchase.

At any point of time, marketers are interested in the ability to assess the level of customer retention, and to predict which customer is retained. To fully utilize the data at hand, we assume this specific point of time to be the end of 2010, which coincides with the end of data duration.

There are two reasons that the customers are grouped based on the year when they made their first purchase. First, customers acquired in the same year, generally share some similar external or internal factors. External factors could include the overall economic condition and the competition intensity in the market during that year. As for the internal factors, they include the company’s yearly promotion campaigns, management practices, business strategies, etc. We believe comparisons among these groups have a high potential in offering insights since most extraneous variables are effectively contained within each group. Second and more importantly, customers acquired in the same year present the same observation duration. For example, for the customers acquired in 2005, data observation is possible from 2005 to 2010, while for the customers acquired in 2010, data observation is only possible in 2010. Obviously, longer observation duration provides more information. Thus, we group the customers based on the year of their first purchase, as the first step in measuring customer retention.

After grouping the customers, we analyze their repeat purchase behavior using real purchasing data on an individual level. Compared to aggregate company-wide level of analysis, the individual-level approach enables observations on finer granularity and makes the analysis or prediction of individual customer possible [7]. With individual-level approach, we are able to determine, in the subsequent analysis, if a specific customer is retained.

In this study, our interest focuses on repeat purchase behavior and how it relates to customer retention. We first locate the customers who had repeat purchase, i.e., who purchased at least twice in the observation duration, excluding the customers who made only one purchase. Then, we calculate statistics of interpurchase times (in months) for each individual customer who had repeat purchase. The statistics calculated are average and maximum interpurchase times on an individual level. For example, if a customer purchased four times in total during the observation duration, we get three interpurchase times (in months). Then, we calculate the maximum and the average of the three interpurchase times. The procedure is performed for every customer who had repeat purchase, producing two statistics, maximum and average interpurchase time, for each customer.

Next, for each customer group we calculate aggregate statistics of the above two individual-level statistics for all the customers with repeat purchase behavior in that group. Specifically, for average interpurchase times we calculate aggregate average and the standard deviation. For maximum interpurchase times, in addition to aggregate average and the standard deviation, we calculate aggregate maximum and the 80th percentile. Like interpurchase time, both individual and aggregate statistics are in months.

We posit that the aggregate 80th percentile of maximum interpurchase times is meaningful in the managerial sense. Literally, the 80th percentile is the value (in months) below which 80 percent of the observed maximum interpurchase times on the individual level may be found. Statistically, the 80th percentile can be interpreted as the duration within which 80 percent of the customers with repeat purchase behavior returned to purchase. Furthermore, we can infer that with 80 percent assurance, if a customer would return to purchase, the interpurchase time since the last purchase will not exceed the 80th percentile.

The 80th percentile for each customer group is useful in measuring customer retention. The 80th percentile of maximum interpurchase times is calculated based on the data up to the end of 2010. Say, at the end of 2010, the marketer would like to assess the customer retention, and he/she is essentially asking, “Based on the historical data, can I tell at this point which customer is retained?” To answer this question, in practice, we observe if an existing customer purchased within the 80th percentile months right before the end of 2010. If the customer did not return to purchase during that duration, with 80 percent of assurance we say that the customer will not be back; otherwise, the customer is retained. The former represents the state of customer attrition, while the latter represents the state of customer retention. As a result, we can measure customer retention for each group with the same level of assurance.

We acknowledge the fact that one cannot foretell if an existing customer would return to purchase. A customer not retained according to our approach
could still return to purchase at an unknown future point of time. On the other hand, a customer regarded as retained is not guaranteed to return. However, we propose to view customer retention as a state at a certain point of time, which satisfies a certain level of assurance. With 80 percent assurance, the measured customer retention at the end of observation duration in our approach is both meaningful and useful.

4. Results

Figure 1 displays the composition in percentage of the new and returning customers every year between 2005 and 2010. Please note that we group the customers according to the year they were acquired, i.e. the year of their first purchase at the web retailer. Since we did not include real purchasing data from customers who were acquired before 2005, customers in 2005 consist of only the new customers acquired in that year. As for the following years between 2006 and 2010, customers consist of both new and returning customers. If we take a further look at Figure 1, the new customers constitutes between fifty and seventy percent of the total customers for years 2006 to 2010. In other words, the percentage of the new customers (customer acquisition) is consistently equal to or higher than that of the returning customers (customer retention) every year between 2006 and 2010 for the web retailer. Besides, as we can imagine, the number of the returning customers of each group is decreasing over the years after their first purchase.

Next, we analyze customers’ repeat purchase behavior using real purchasing data on the individual level. For each group, we first sort out the customers who had ever returned to purchase since their first purchase. Table 1 shows the customers with repeat purchase behavior. For the customers acquired in the last two years (the last two groups), their observation durations are less than two years, which we consider are not long enough. As a result, the reported numbers for the last two years are for reference only and we mark them gray in Table 1. For the customers acquired in other years, the percentages for customers with repeat purchase behavior are generally high, between 40% and 60%. When inspecting the percentage of each group, we assume that for the customers who are acquired earlier and thus with a longer observation duration they are more likely to return to purchase at least once. However, we find that for the customers acquired earlier their percentage for customers with repeat purchase behavior is, on the contrary, lower. For example, the percentage is only 40% for the customers acquired in 2005. On the other hand, for the customers acquired in 2008 with the observation duration less than three years, the percentage is, nevertheless, close to 60%. As a result, we can infer that the customers acquired recently by the web retailer are of better quality in that the percentage for customers with repeat purchase behavior is even higher than that with longer observation duration.

Table 1. Customers with repeat purchase behavior

<table>
<thead>
<tr>
<th>Year</th>
<th>Customers with repeat purchase</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>40.25%</td>
</tr>
<tr>
<td>2006</td>
<td>48.72%</td>
</tr>
<tr>
<td>2007</td>
<td>56.07%</td>
</tr>
<tr>
<td>2008</td>
<td>58.35%</td>
</tr>
<tr>
<td>2009</td>
<td>51.89%</td>
</tr>
<tr>
<td>2010</td>
<td>31.43%</td>
</tr>
</tbody>
</table>

After we sort out the customers who had ever returned to purchase since their first purchase, we calculate average interpurchase time and maximum interpurchase time for each of them. Then, we calculate aggregate statistics of the two statistics on the individual level for each group. Table 2 shows aggregate statistics of average interpurchase times for customers acquired in different years. We mark the reported numbers for the last two years gray in Table 2 because of inadequate observation durations. We do the same in the following analysis. In Table 2, the aggregate averages of average interpurchase times are between 5.62 and 6.45 months. If the values are rounded, they are all six months. The six-month period reflects the property of the web retailer’s products. Beauty products are habitual consumer products. For habitual consumer products, their interpurchase times tend to be regular. The aggregate averages of average interpurchase times suggest that
the average replenishment cycle for beauty products is about six months.

Table 2. Aggregate statistics of average interpurchase times

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>6.45</td>
<td>7.00</td>
</tr>
<tr>
<td>2006</td>
<td>6.22</td>
<td>5.82</td>
</tr>
<tr>
<td>2007</td>
<td>5.73</td>
<td>5.75</td>
</tr>
<tr>
<td>2008</td>
<td>5.62</td>
<td>4.93</td>
</tr>
<tr>
<td>2009</td>
<td>4.41</td>
<td>3.57</td>
</tr>
<tr>
<td>2010</td>
<td>2.51</td>
<td>1.83</td>
</tr>
</tbody>
</table>

Unit: Month

Table 3 shows aggregate statistics of maximum interpurchase times for customers acquired in different years. In Table 3, the aggregate 80th percentile of maximum interpurchase times increase smoothly as the observation duration extends, from 14 months in 2008 to 17 months in 2005.

Table 3. Aggregate statistics of maximum interpurchase times

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean</th>
<th>SD</th>
<th>Maximum</th>
<th>the 80th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>10.36</td>
<td>9.62</td>
<td>66</td>
<td>17</td>
</tr>
<tr>
<td>2006</td>
<td>10.55</td>
<td>8.36</td>
<td>53</td>
<td>16</td>
</tr>
<tr>
<td>2007</td>
<td>9.90</td>
<td>7.81</td>
<td>44</td>
<td>15</td>
</tr>
<tr>
<td>2008</td>
<td>8.91</td>
<td>6.23</td>
<td>35</td>
<td>14</td>
</tr>
<tr>
<td>2009</td>
<td>6.16</td>
<td>4.12</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>2010</td>
<td>2.92</td>
<td>2.04</td>
<td>11</td>
<td>5</td>
</tr>
</tbody>
</table>

Unit: Month

Before we move on to apply the aggregate 80th percentile of maximum interpurchase times to measure customer retention for the web retailer, we further examine the distribution of maximum interpurchase times. Figure 2 is the distribution of maximum interpurchase times for customers acquired in 2005. In general, the longer the maximum interpurchase time gets, the less the customers can be found there. There are 12.7 percent of customers whose maximum interpurchase time is just one month. This indicates that 12.7 percent of customers did not return to purchase beyond one month during the 6-year observation duration from 2005 to 2010. These customers with one-month maximum interpurchase time may purchase more than twice, but their interpurchase times were all equal to or less than one month.

The percentage of customers drops to 8.7 percent when the maximum interpurchase time is two months. The percentage of customers continues to drop when the maximum interpurchase time gets longer. According to Table 3, the 80th percentile of maximum interpurchase times is 17 months for customers acquired in 2005. This means that there are 80 percent of customers whose maximum interpurchase time is equal to or less than 17 months. In other words, 80 percent of customers returned to purchase within 17 months. Therefore, we use 17 months as the threshold to measure customer retention for the customers acquired that year. If a customer who made his or her first purchase in 2005 purchased again during the 17-month duration before the end of 2010 (from August 2009 to December 2010), we conclude that he or she is a retained customer.

According to the aggregate 80th percentile of maximum interpurchase times for customers acquired in different years, we can measure customer retention state for each individual customer in the specific group. Table 4 is the measurement results of customer retention in the end of 2010 for the web retailer. According to Table 4, about 29% of customers who made their first purchase in 2008 are retained. As for customers who made their first purchase in 2005, only 3% are regarded as retained. This suggests that attrition rate increases quite rapidly as the time span since a customer’s first purchase advances. For the customers who first made their purchase in 2005, the retention rate drops to only 3% after about six years by the end of 2010.
the aggregate 80th percentile of maximum interpurchase times. When comparing the 80th percentiles between 2005 and 2008, the 80th percentile is 17 months for year 2005 and 14 months for year 2008. The former is a more tolerant criterion, while the latter is more stringent. For customers acquired in 2005, they need to have repeat purchase behavior during the 17 months before the end of 2010 in order to be regarded as retained. Conversely, for customers acquired in 2008, the 14-month period before the end of 2010 is shorter and thus stricter. This answers one of the questions we raised before: Can we apply the same threshold of time duration to all existing customers acquired at different times and decide their retention status with the same level of assurance? The answer is no. Our approach applies the 80th percentile of maximum interpurchase times using real purchasing data to measure customer retention at the end of observation duration. The values of the aggregate 80th percentile of maximum interpurchase times for customers acquired in different years are different. Therefore, our approach of using the aggregate 80th percentile as the threshold would mean that the threshold is varying according to different years. However, it is reasonable that the criterion is more tolerant for customers acquired earlier but more stringent for those acquired more recently. The thresholds could be different for customers acquired in different years, but all the results attain 80 percent assurance.

In addition to the aggregate 80th percentile of maximum interpurchase times, we notice that other aggregate statistics in Table 2 and 3 increase with the length of observation duration. For example, the aggregate average of average interpurchase times increases from 5.62 months in 2008 to 6.45 months in 2005. In addition, the aggregate average of maximum interpurchase times increases from 8.91 months in 2008 to 10.36 months in 2005. We believe that the phenomenon is due to the following two reasons. First, as time passes by, some frequent buyers tend to drift away, making the interpurchase times longer. Second, as observation duration is getting longer, extreme cases are more likely to show up. For example, for customers acquired in 2005, the most extreme maximum interpurchase time on the individual level is 66 months. However, the corresponding time for customers acquired in 2008 is just 35 months.

5. Discussion

5.1. Benefits of the approach

According to the measurement results of customer retention for the web retailer, we found that as the time span since a customer’s first purchase advances, customer attrition rate increases rapidly. It is well recognized in the literature that customer retention is particularly challenging in the transparent electronic business market. However, without a systematic approach to measure customer retention quantitatively, a web retailer can only rely on guts felling and is not able to clearly perceive how challenging customer retention is. Our approach offers to measure customer retention using real purchasing data so that a web retailer is able to have a clear view on its actual customer retention condition.

Besides comparing customer retention results based on different years of customers’ first purchase and with different lengths of observation duration, we can also compare retention results based on different years of customers’ first purchase and with the same length of observation duration. For example, the current data is from 2005 to 2010. In the future, when we obtain the data in 2011 from the web retailer, we can update the measurement results of customer retention in Table 4. Currently, in Table 4, the customer retention result is 19.41% for customers acquired in 2007 and the result is based on the observation duration of four years (from 2007 to 2010). When we update Table 4 with 2011 data, the new customer retention result for customers acquired in 2008 will also be based on the observation duration of four years (from 2008 to 2011). Comparing the two results means that we are comparing the retention results between customers acquired in 2007 and 2008 with the same length of observation duration. If one is better than the other, we can conclude objectively that under the same length of observation duration and the same level of assurance the one with the better retention result suggests that customers acquired in that year are of better quality. The web retailer then needs to ponder what external and internal factors cause the difference.
5.2 Validity of the approach

For the customers acquired in 2005, an independent-samples t-test is conducted to compare the number of orders from retained customers and from non-retained customers. The result shows that there is a significant difference in the number of orders from retained customers ($M=13.40$, $SD=9.06$) and from non-retained customers ($M=2.16$, $SD=2.46$); $t(461) = -26.57$, $p = 0.00$. The large discrepancy in the number of orders for retained and non-retained states is consistent across years (Table 5). These results suggest that our approach to determine if a customer is retained effectively portrays a customer’s potential contribution in terms of number of orders. Specifically, the results suggest that when a customer is regarded as retained based on our approach, he or she placed more orders. This establishes the validity of our approach in that retained customers are also well recognized for their repeat purchases.

<table>
<thead>
<tr>
<th>Year</th>
<th>Two conditions</th>
<th>$M$</th>
<th>$SD$</th>
<th>$t$</th>
<th>df</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>Retained customers</td>
<td>13.40</td>
<td>9.06</td>
<td>26.57</td>
<td>461</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Defected customers</td>
<td>2.16</td>
<td>2.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>Retained customers</td>
<td>11.38</td>
<td>6.99</td>
<td>-30.72</td>
<td>589</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Defected customers</td>
<td>2.36</td>
<td>2.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>Retained customers</td>
<td>9.02</td>
<td>5.48</td>
<td>-47.22</td>
<td>1665</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Defected customers</td>
<td>2.30</td>
<td>2.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>Retained customers</td>
<td>6.64</td>
<td>4.06</td>
<td>-50.24</td>
<td>2136</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Defected customers</td>
<td>1.88</td>
<td>1.47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>Retained customers</td>
<td>4.64</td>
<td>2.64</td>
<td>-51.07</td>
<td>2118</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Defected customers</td>
<td>1.46</td>
<td>.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>Retained customers</td>
<td>1.78</td>
<td>1.24</td>
<td>-23.44</td>
<td>4794</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Defected customers</td>
<td>1.21</td>
<td>.53</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.3 Implications for relational marketing

Customer retention is generally accepted as a key objective in relationship marketing [8]. Sheth [14] defines relationship marketing as the retention of profitable customers through continuing one-to-one collaborative and partnering activities. The ability to determine if a customer is retained is crucial in relationship marketing. Our approach differentiates customers who are prone to repurchase from those who can be regarded as non-retained. The two types of customers should be approached quite differently in relational marketing. Besides, customers regarded as retained according to our approach placed significant more orders than those regarded as non-retained. As a result, the retained customers according to our approach are also those profitable customers of special interest in the relationship marketing.

5.4 Limitations and future studies

Despite the advantages mentioned above, our approach to measure customer retention has some limitations. First, the analysis was based on the data from a web retailer which offered products in a specific product category. Future research can use data from other web retailers or product categories to verify the applicability of the approach. Second, the 80th percentile of maximum interpurchase times was used as a threshold in measuring customer retention to achieve 80 percent measurement assurance. In general, higher threshold provides more assurance. A different choice of threshold can be considered by the web retailer, especially when the aspects of industry competition and product category are factored in. Third, we demonstrated that on average the retained customers identified by our approach placed more orders than those not-retained. This observation is consistent for all customer cohorts, i.e. the customers acquired in different years. However, we are not sure if the retained customers do contribute more to the web retailer in terms of total purchase amount. Potentially profitable customers are recognized not only by the number of orders but also by the total monetary contribution. For the retained customers to be more profitable, the evidence of monetary contribution is desirable.

We aim to address the three limitations in our future studies.

6. Conclusion

In this study, we have offered rich descriptive statistics of online repeat purchase behavior. We propose to use the aggregate 80th percentile of maximum interpurchase times to measure customer retention, in other words, to determine if a customer is retained or not, at the end of the observation duration. Customer retention is a challenging metric for doing business in the transparent electronic business market, whether in measuring or improving it. The measurement of customer retention is central to how well a web retailer can evaluate its performance. By linking customer retention measurement to actions in the market, businesses
have much better chance to excel. Our approach to measure customer retention has shown the benefits with proven validity.

7. Acknowledgement

This research is partially funded by National Science Council in Taiwan under project number NSC 100-2410-H-004-026-MY2 and a grant from National Chengchi University, Taiwan, R.O.C.

8. References


