Panteion University of Social and Political Sciences

Department of Economic and Regional Development

Post Graduate Program: "Applied Economics and Regional Development"

Direction: "Applied Economics and Administration"

"BUSINESS ANALYTICS AND MARKETING CAMPAIGNS"

Author
Ekatereni Panaousis

Supervisor
Pr. Grigorios Siourounis

January, 2017

ABSTRACT

Objectives of the Study

The measurability of user traffic in various websites, game platforms or applications allows somebody to understand important characteristics of users' visits in conjunction with specific reinforcing actions that have been made. This present study examines how a marketing action affected the number of daily users in a trivia game platform. More specifically, the objective is to show if the marketing campaign implementation in mobile apps had a positive effect on daily users. In addition, the effect of this marketing action on the retention rate of users is being investigated, in an attempt to qualitatively enhance the findings. To sum up, this thesis suggests a conceptual framework based on previous research, in order for marketing campaign effectiveness in mobile apps to be assessed.

Academic Background and Methodology

Taking into account the total of the literature review has been used for the purpose of this thesis, it could be said that, as for the theoretical part, it mainly draws from the field of marketing research, with a bit of psychology. The focal hypothesis tested is based on previous research in the fields just mentioned. For the case study conducted in this thesis, Times Series Analysis is used and more specifically, ARMA methodology is applied. Based on the basic principles of econometrics, five models have been built and estimated, in an attempt to test if the hypothesis mentioned above is accepted or not.

Findings and Conclusions

The case study contacted in this thesis identifies that a campaign effect actually exists in case the marketing campaign is implemented in mobile apps. While the mean daily users in the trivia game platform after the campaign ended are calculated to be four times more than before its launch, an effect on user retention is captured, as well. The first-mentioned effect is identified when the first of the dummy variables, which created for the purpose of this thesis, is added in the main model. Furthermore, when the main model includes the second dummy variable and the second and third one simultaneously, the effect on user retention becomes evident, as well.

Key Words

Business analytics, engagement, marketing campaign effectiveness, mobile apps, retention rate of users

ΠΕΡΙΛΗΨΗ

Στόχοι της Μελέτης

Η δυνατότητα καταμέτρησης της κινητικότητας των χρηστών σε διάφορες ιστοσελίδες, πλατφόρμες παιχνιδιών ή εφαρμογών επιτρέπει σε κάποιον να κατανοήσει σημαντικά χαρακτηριστικά των επισκέψεων των χρηστών, σε συνδυασμό με συγκεκριμένες ενισχυτικές ενέργειες που έχουν λάβει χώρα. Η παρούσα εργασία εξετάζει πώς μία ενέργεια μάρκετινγκ επηρέασε τον αριθμό των χρηστών ενός παιχνιδιού ερωτήσεων σε καθημερινή βάση. Πιο συγκεκριμένα, ο σκοπός είναι να γίνει φανερό αν η εκτέλεση της καμπάνιας μάρκετινγκ σε εφαρμογές για κινητά είχε θετική επίδραση στους καθημερινούς αυτούς χρήστες. Επιπρόσθετα, ερευνάται η επίδραση αυτής της ενέργειας μάρκετινγκ στη διατήρηση των χρηστών, σε μια προσπάθεια να ενισχυθούν ποιοτικά τα ευρήματα της παρούσας εργασίας. Συνοψίζοντας, η εργασία αυτή προτείνει ένα εννοιολογικό πλαίσιο, βασισμένο σε προηγούμενη έρευνα, προκειμένου η αποτελεσματικότητα της καμπάνιας μάρκετινγκ σε εφαρμογές για κινητά να εκτιμηθεί.

Ακαδημαϊκό Υπόβαθρο και Μεθοδολογία

Λαμβάνοντας υπόψη το σύνολο της βιβλιογραφίας που χρησιμοποιήθηκε για τον σκοπό της παρούσας εργασίας, θα μπορούσε κανείς να πει ότι, όσον αφορά στο θεωρητικό κομμάτι, πηγάζει κυρίως από το πεδίο της έρευνας μάρκετινγκ και σε μικρό βαθμό από αυτό της ψυχολογίας. Η κεντρική υπόθεση που εξετάζεται βασίζεται σε προηγούμενη έρευνα στα δύο αυτά πεδία που μόλις αναφέρθηκαν. Αναφορικά με τη μελέτη περίπτωσης της παρούσας εργασίας, χρησιμοποιείται η Ανάλυση Χρονοσειρών και πιο συγκεκριμένα, εφαρμόζεται η μεθοδολογία ARMA. Με βάση τις βασικές αρχές οικονομετρίας, πέντε μοντέλα έχουν χτιστεί και εκτιμηθεί, ώστε να εξεταστεί η αποδοχή ή όχι της υπόθεσης που αναφέρεται παραπάνω.

Ευρήματα και Αποτελέσματα

Η μελέτη περίπτωσης στην παρούσα εργασία εξακριβώνει ότι πραγματικά η καμπάνια επιδρά στους καθημερινούς χρήστες ενός παιχνιδιού ερωτήσεων, όταν αυτή υλοποιείται μέσω εφαρμογών για κινητά. Ο μέσος αριθμός των ημερήσιων χρηστών μετά το τέλος της εκστρατείας υπολογίζεται να είναι τέσσερις φορές μεγαλύτερος από αυτόν πριν την έναρξή της, ενώ παράλληλα παρατηρείται και μια επίδραση στη διατήρηση των χρηστών μετά το τέλος αυτής. Η πρώτη επίδραση που αναφέρεται παραπάνω εντοπίζεται όταν η πρώτη εκ των ψευδομεταβλητών της παρούσας εργασίας εισάγεται στο βασικό μοντέλο. Επιπλέον, όταν στο κύριο μοντέλο συμπεριλαμβάνεται η δεύτερη και οι δεύτερη και τρίτη ψευδομεταβλητές ταυτόχρονα, η επίδραση στη διατήρηση των χρηστών γίνεται επίσης εμφανής.

Λέξεις Κλειδιά

Επιχειρηματική αναλυτική, συμπλοκή, αποτελεσματικότητα καμπάνιας μάρκετινγκ, εφαρμογές για κινητά, βαθμός διατήρησης χρηστών

ACKNOWLEGMENTS

~ My first gratitude goes to my supervisor, Pr. Grigorios Siourounis, to which I am highly indebted and whose expertise, generous guidance and support made it possible for me to work on a topic that was to a great interest to me, for providing me with information I could not possibly have discovered on my own and for being a source of motivation. It was a pleasure working with him.~

~ I would like to express my gratitude to Pr. Clive Richardson and Pr. Stavros Degiannakis for finding out time to reply to my e-mails, giving their precious and kind advice regarding this study. ~

~ Words are powerless to express my gratitude to my family, Antonios, Zoi and Pavlos, for their unconditional support and love, during the writing period of this thesis.

This is for you! ~

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

It is a fact that significant progress has been achieved in the field of Business Analytics over the past few years. This improvement has been crucial for those people in a company their main responsibility of which can be summarized in the phrase "Decision Making". While, nowadays, there is a number of approaches to making decisions, i.e. tradition ("We've always done it this way!"), intuition ("gut feeling") and rules of thumb ("As the restaurant owner, I schedule twice the number of waiters and cooks on holidays!"), businesses need to obtain competitive advantage against their competitors by grasping the chance that BA generously gives: converting data into knowledge (Camm et al., 2015). But what BA is exactly?

A formal definition of this term is the following: BA is the process of scientifically transforming data into insight for making better decisions. It involves simple tools as reports and graphs, and more sophisticated ones as optimization, data mining and simulation (Camm et al., 2015). In addition, there are several types of applications of analytics by application area such as Financial Analytics, Human Resource Analytics, Marketing Analytics, Supply Chain Analytics, Analytics for Government and Nonprofits, Health Care Analytics and –finally– Web Analytics (Camm et al., 2015).

In today's world, in order for businesses to be competitive, their marketing executives need to withstand unprecedented challenges to create new and loyal customers. Until a few years ago, the traditional ways of applying marketing principles were driven solely by the ideas of big companies and their competitors (Burby & Atchison, 2007). Now, this function is only a matter of a single click. Those clicks, afterwards, are the main input for companies which apply Business Analytics, in an attempt to evaluate the effectiveness of a site or a marketing campaign implemented in apps, making it clear what is working and what is not (Burby & Atchison, 2007).

In the case of a marketing campaign evaluation, however, it is quite difficult for someone to determine how this campaign affects consumption in the case of products or traffic in the case of online or mobile apps, games and websites, as far as it is often holistic, in a way that involves both digital and traditional promotional activities (Fagerstrom & Ghinea, 2010). Analytics, though, allows business executives to dig in and understand everything they can about the desired behavior based on the data they have collected.

In this study, we will examine how a marketing campaign in mobile apps affects the number of users of a game in a daily base. In particular, this thesis is conducting a case study to examine the number of daily users in a trivia game platform before, during and after the campaign implementation in mobile apps. In addition, apart from examining the "Campaign Effect" and in an attempt to obtain more qualitative findings for the effectiveness of the campaign, the retention rate of users as a result of the action above is being investigated, while we also refer to the "Cost per User".

1.1. Research Question

A major challenge with analytics has been the evaluation of marketing campaign effectiveness. This thesis proposes a framework for identifying the success rate of a marketing campaign implemented in mobile apps. To be more specific, the goal is to show if the campaign was successful in the following way:

Did the marketing campaign affect positively the number of daily users? And if so, what was the contribution of the campaign to the retention rate of users?

To test all the above, a case study of proprietary data, which were tracked for a similar purpose and for a brand new unknown start-up company, is conducted.

Robehmed (2013) has gathered many perspectives about what a startup is exactly. “A startup is a company working to solve a problem where the solution is not obvious and success in not guaranteed” and “a business or undertaking that has recently begun operation” are only two of those perspectives, while the author concludes that, all in all, a startup is characterized by its ability to grow.

The case study examines the campaign effect on the number of daily users in a trivia game platform and the effect on the retention of users as well.

1.2. Structure of Thesis

The present thesis is structured in the following way: it starts with an introduction about Business Analytics and its importance for marketing executives and after this, the research question is briefly discussed.

The second chapter is a literature review of the field of marketing research and a bit of psychology. In particular, the most common general approaches toward some useful – for the purpose of this study- concepts are defined, in order for the hypothesis that is being examined to be extracted.

The third chapter introduces the conceptual framework this thesis proposes and the research method used for the analysis. To be more specific, it firstly presents a framework about those constructs that should be taken into account in order for the effectiveness of a marketing campaign implemented in mobile apps to be assessed. Afterwards the theory about Time Series Analysis and specifically about ARMA(k,l) models is presented.

The fourth chapter presents the results of the case study conducted in this thesis. In particular, it presents the outputs of the models estimated and the various tests applied, while they are briefly discussed, as well.

The fifth and final chapter of this thesis presents a brief discussion about the findings of the analysis, the hypothesis acceptance or rejection and the limitations should be taken into account.

2. LITERATURE REVIEW

This chapter presents a brief literature review of the field of marketing research and psychology, with a focus on advertising effectiveness. It has been made an attempt the most common approaches toward the concepts that seem to determine the advertising results to be presented in this structural way, in order to make clear the way we have been led to the research questions.

2.1. Basic Concepts

2.1.1. Attitude

Attitude is a concept that many marketing researchers have studied over the years. According to Mitchell and Olson (1981), there are two main reasons for this enduring interest. First, attitudes are often considered relatively stable and as a consumer behavior indicator, in a way that they should provide useful predictions of consumer behavior toward a product or service. Stahl et al. (2012), based on the theory of reasoned action (Engel et al., 1995) and hierarchy-of-effects models of consumer behavior (e.g., Lavidge and Steiner, 1961), pointed out, as well, that consumer attitudes are a precursor to consumer actions. Second, a satisfying number of theoretical models of the attitude construct has been provided by the field of social psychology, which in turn has stimulated much of the attitude research in marketing.

Fishbein and Ajzen (1975) claimed that

"*A person's attitude is a function of his salient beliefs at a given point in time*".

Mitchell and Olson (1981) indicated that those beliefs regarding the attributes of a product or service have a major mediating effect on brand attitudes, whereas attitudes mediate to a significant extent behavioral intentions.

They pointed out, in addition, that the product attribute belief as an index is not the one and only mediator of attitude formation. Rather, they indicated that individuals can base completely on visual information with no clear reference about the brand of the product or service, even in case of visual stimuli that has apparently no relevance with the product or service brand.

2.1.2. Routes of Attitude Changes or Persuasion

Consumers' attitude changes (or persuasion from the point of view of marketing executives) toward a product or service brand consist of two main distinct routes.

According to Petty, Cacioppo and Schumann (1983), the first route is the central one, which attributes consumers' attitude change to a person's diligent information processing. This information is central for the consumer to the true merits of a particular attitudinal position. For example, the way a person evaluates several alternatives, according to the manner he/she combines relevant beliefs about the present issue is a factor that characterizes the central route. A quite important take home message for marketing executives is that this kind of attitude changes are postulated to be long-lasting and a good predictor of behavior (Petty et al., 1983).

Afterwards, the second route is the peripheral one. Attitude changes that occur via the peripheral route do so because the attitude issue or object relates with positive or negative cues. For instance, a person may accept an opinion simply because it was presented by an expert. In contrast to the case of the preceding paragraph, attitude changes induced via peripheral route are postulated to last only a limited period of time and be inappropriate means of behavior prediction (Petty et al., 1983).

Massaro (1988) noticed the following, regarding a Petty and Cacioppo's (1986) research about attitude change:

"The most innovative contribution of this (Petty and Cacioppo's) approach is the distinction between two major routes of attitude change. The central route involves the careful and thoughtful (perhaps conscious, controlled, and effortful) assessment of the message, whereas the peripheral route involves a fairly direct change in attitude

without careful thought and consideration (perhaps unconscious, automatic, and effortless)".

It should be noticed, though, that if somebody wants to understand how attitude changes occur, he/she must consider that the same person can be either enthusiastic with information collection/processing or deliberately avoid any process of information assessment, depending on the situation he/she might be (Petty et al., 1983).

2.1.3. Experiences during Web Navigation

When a consumer/user is connected with a site (or uses a mobile application in the case of the present study), he/she has some experiences. These experiences are defined as the consumer's/user's beliefs about what this specific web context offers him/her (Calder et al., 2009). A web context can provide utilitarian experiences and intrinsically enjoyable experiences, as well. Depending on the consumer's/user's personal needs and his/her central and peripheral route of attitude, a web context can be engaging in many different ways.

Kim, Lin and Sung (2013) noticed that companies are able to take advantage of this emerging platform for marketing communication, named mobile applications. All they need to do, in order to engage with consumers more effectively, is to provide them unique experiences associated with their brand, product or service, when they actually use those apps, in an attempt to take care of all manner of daily tasks.

At the same time, however, marketers need to be sparing with the amount of ads and campaigns implemented in apps, so that ad avoidance becomes less intense. The point is that it is not only about engagement with a means, as far as consumers/users tend to be affected to a great extent by information they have learnt from their prior personal experiences (Fazio et al., 1978; Smith and Swinyard, 1982). Those experiences are characterized as negative ones when, for example, dissatisfaction, lack of utility and incentive are born (Cho and Cheon, 2004). And that seems to be the case when advertising in the Internet and apps is perceived to be intrusive, as it interrupts consumers' goals or convinces them that the amount of ads is excessive.

2.1.4. Involvement

There have been various definitions in an attempt the term "Involvement" to be clarified. A major characteristic of the existing literature is that, in many cases, the term "Involvement" is equated with the term "Engagement".

Mittal and Lee (1989) tried to broadly interpret the concept of involvement, based on the pre-existing literature. According to their definition,

"Involvement is the perceived value of a goal-object that manifests as interest in that goal-object".

Depending on what this goal-object is exactly, the authors above claimed that there can be the two following forms of involvement. Firstly, in case of a product, the form is called product involvement and it is an interest that a consumer expresses, when he/she believes that the product class meets important values and goals. Secondly, in case of a purchase decision, the form is called purchase involvement or brand-decision involvement and it is the interest of a consumer when he/she is about to make a brand selection (Mittal and Lee, 1989).

In addition, the authors noticed that purchase involvement can be characterized as low, when consumers have to occasionally select a brand (Mittal and Lee, 1989). In contrast, a high purchase involvement is occurring in case of a deeply deliberated brand choice decision process (Mittal and Lee, 1989).

Petty, Cacioppo and Schumann (1983) performed an experiment, in which they exposed some undergraduates to a magazine ad, under conditions of either high or low product involvement. The results showed that the manipulation of the hortatory information quality (strong or weak argument for the product) had a greater impact on attitudes under high than low involvement, while the manipulation of product endorser (important sport celebrities or average citizens) had a greater impact under low than high involvement.

Last but not least, Mittal and Lee (1989) proved the significance of product involvement as an antecedent of brand-decision involvement.

2.1.5. Engagement

Apart from the fact that the term "Engagement" is very often equated with the term "Involvement", another main characteristic of the existing literature is the number of definitions and interpretations of the former term.

In the case of organizational behavior literature, for example, Saks (2006) stated that this interest in the concept is attributed to reports for a significant (or at least to some extent) disengaged percentage of the workforce from their workplace, which actually costs U.S. businesses $300 billion per year in lost productivity.

Regarding the subject of this present study, Calder, Malthouse and Schaedel (2009) claimed that the most of the circulated definitions are consequences of engagement rather than engagement itself. In particular, they argued that the desire to visit a specific website (or a trivia game platform in the case of this study), download its pages, highly recommend it to somebody or be disappointed in case of non-availability, are all main characteristics of an "engaged" user behavior (Calder et al., 2009). According to their opinion, engagement needs to happen first in order for usage, affect and advertising outcomes to follow, a process closely related to the different experiences a consumer/user has during navigation. The authors used Figure 1 to make this state clear.

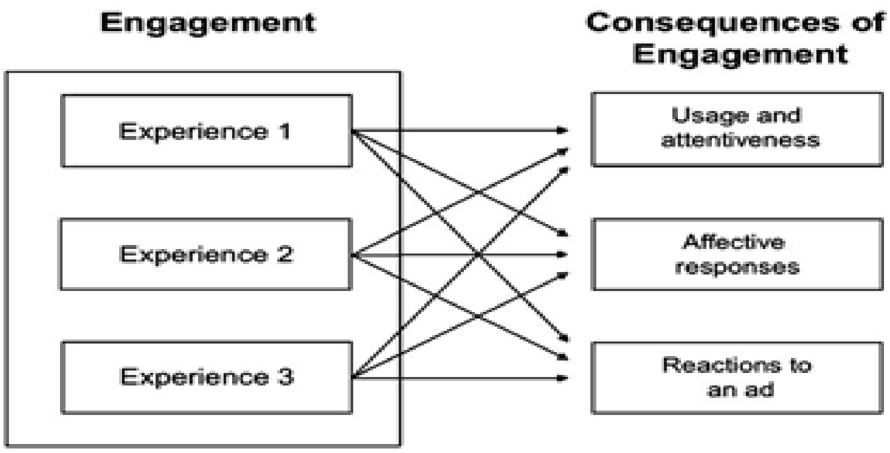


Figure 1. Engagement and its Consequences

They, also, tested the following hypothesis:

"Engagement with the surrounding online media vehicle context increases advertising effectiveness".

The conclusion of this study was that online media are as capable as traditional media to create engaged web users, on which advertising has its own impact (Calder et al., 2009).

In addition, they showed that engagement consists of the two following factors: personal engagement and social-interactive engagement. The first one is manifested in experiences that traditional media offer, whereas the second one mainly refers to websites. Both of these two forms of engagement contribute to advertising effectiveness as well.

In conclusion, over the past few years, different media context effects on advertising have been examined bringing to the fore a consensus of opinion among researchers: the higher the engagement the more effective the advertising (Bronner and Neijens, 2006; Wang, 2006).

In the case of mobile phones, Gupta (2013) claimed that people simply do not like ads on their screens, inducing marketing executives to create or use apps with certain attributes. Either consumers/users attitude toward a product, a service or a brand changes via the central route or the peripheral one, mobile applications that will be used for a marketing campaign implementation need to do at least one of the following: i) add convenience (i.e. by offering utilitarian experiences), ii) offer unique value (by offering either utilitarian experiences or intrinsically enjoyable ones, according to the consumer's/user's profile), iii) provide social value, iv) offer incentives, v) entertain (Gupta, 2013). The characteristics above can promise the creation of long-term engaged customers.

Fagerstrom and Ghinea (2010) have written an article disclosing the remarkable results in case marketing executives realize the importance of those five characteristics above. According to this article, the SCA Libresse launched a campaign and specifically, a two-month online design competition, in attempt to encounter its strong competitors. For the purpose of this campaign a package crafted was created, as the main goal of this attempt was to invite the target segment (girls between 14 and 25 years old) to design a pair of underpants on the Libresse Web site. This attempt was based on a prior research which indicated that girls in that age are interested enough in fashion design.

The authors (2010) claimed that, through this action, SCA Libresse offered unique and social value and incentive, as –apart from the cash prize- the winner's underpants would be promoted for sale in a fashion chain of stores for teenagers, making the wish of being a designer for once come true. In addition, the design competition can be characterized as entertaining, as the participants could use templates, complete figures and freehand drawing, turning on their creativity. The option to vote for their favorite piece is another reason for which the competition can be considered this way.

According to the authors (2010), the results of this attempt were overwhelming. The aim of the company was to increase the web site traffic by 25%, while the actual increase touched 75%! In addition, brand awareness and positive attitude toward the brand were achieved and a sales increase, as well. This is the reason why the SCA Libresse case is a living example of how to create engaged consumers by using means and ways with characteristics as those we have mentioned above.

3. CONSEPTUAL FRAMEWORK AND RESEARCH METHOD

This chapter presents the conceptual framework the study generally proposes, in order for marketing campaign effectiveness in mobile apps to be assessed, and the quantitative research method used for this purpose, as well. At first, this framework, which is based on previous research, is introduced. After this, Time Series Analysis is briefly discussed.

3.1. Framework for the Study and Hypothesis Tested

It is a fact that, nowadays, mobile phones –and especially smartphones- are extensions of their owners' hands. This phenomenon allows marketers to exploit the opportunity that the extensive use of mobile phones, and specifically the use of mobile apps offers: create new and loyal customers. By using various applications, which either help people to cope with activities during daily routine or contribute to their entertainment, companies are able to advertise their products, services and brands, while there is a great opportunity for consumers/users to get engaged with this specific object they are exposed to.

The remarkable point here is that consumers do not perceive this tactic as advertising, as they appreciate the various kinds of benefits they enjoy by using those apps (Gupta, 2013). In order for those benefits to be preserved, however, marketers need to take heed of the challenges and issues that consumers/users face when using a specific app, as far as a negative experience is able to cause a reduction of the advertising effectiveness.

A recent study about user-reported issues of iOS apps, uncovered twelve (12) different types of user complaints (Khalid et al., 2015). Functional errors, feature requests and app crashes are presented as the most frequent complaints, while those about privacy, ethical issues and hidden app costs are the most negatively-perceived ones by users, causing a really low rating of the app (Khalid et al., 2015). Those findings provide marketing executives insight into which specific mobile applications they should choose or what they should take into account in case they create or choose one for the campaign launch, in order for the engagement process not to be impeded.

At this point, taking into account the following: i) companies can use mobile apps in order to create engaged consumers/users, as apps enhance consumers' life in many different ways (Gupta, 2013), and ii) users' experiences and user engagement must preexist so that effectiveness of marketing campaigns in mobile apps turns up (Calder et al., 2009), we propose our focal hypothesis:

The use of mobile applications for marketing campaign implementation increases daily users on average in the trivia game platform.

Apart from trying to provide insight into the effectiveness of the marketing campaign implemented in this specific way, in terms of an increased number of daily users, it would be very interesting to show what part of them could be considered as loyal or, in other words, retained.

Based on the preexisting marketing literature, Gustafsson, Johnson and Roos (2005) have chosen the following ones as the three of the most important drivers of retention: i) overall customer satisfaction, in terms of a general assessment of performance to date, ii) affective commitment, which actually portrays feelings as trust and mutual support and calculative commitment, a less emotional construct than those just mentioned, as it actually relates to economic incentives or a lack of options and, finally, iii) specific factors or events that caused a change in customers' point of view regarding to the perceived performance, characterized as "triggers".

Based on prior research (Fornell, 1992; Fornell et al., 1996), the authors made the assumption that customer loyalty can be affected to a great extent and in a positive way by the former driver, which, according to Boulding et al. (1993), results in promotional actions as positive word-of-mouth is. In addition, they assumed that affective and calculative commitment positively affect customer retention, as well, while triggers result in a weaker relationship between satisfaction and retention.

The findings of their research seem to be surprising as Gustafsson, Johnson and Roos (2005) have been partially proved right. To be more specific, customer satisfaction and calculative commitment proved to affect retention positively, in concordance with their prior hypotheses. On the contrary, the effect of affective commitment on retention cannot be captured, when the first is included with customer satisfaction. Regarding the effect of triggers on the customer retention or the relationship between the latter and satisfaction, the findings are not consistent to the authors' primary hypothesis, while they have noticed that it remains controversial as prior studies support the opposite (Bolton, 1998; Seiders et al., 2005).

Bowden (2009), also, referred to the terms of satisfaction and affective and calculative commitment. In particular, she depicted the process of customer engagement through a figure (Fig. 2), in which a combination of rational and emotional bonds are formed. In essence, she tried to make clear the way an individual becomes loyal to a service brand, for which engagement plays a major role.

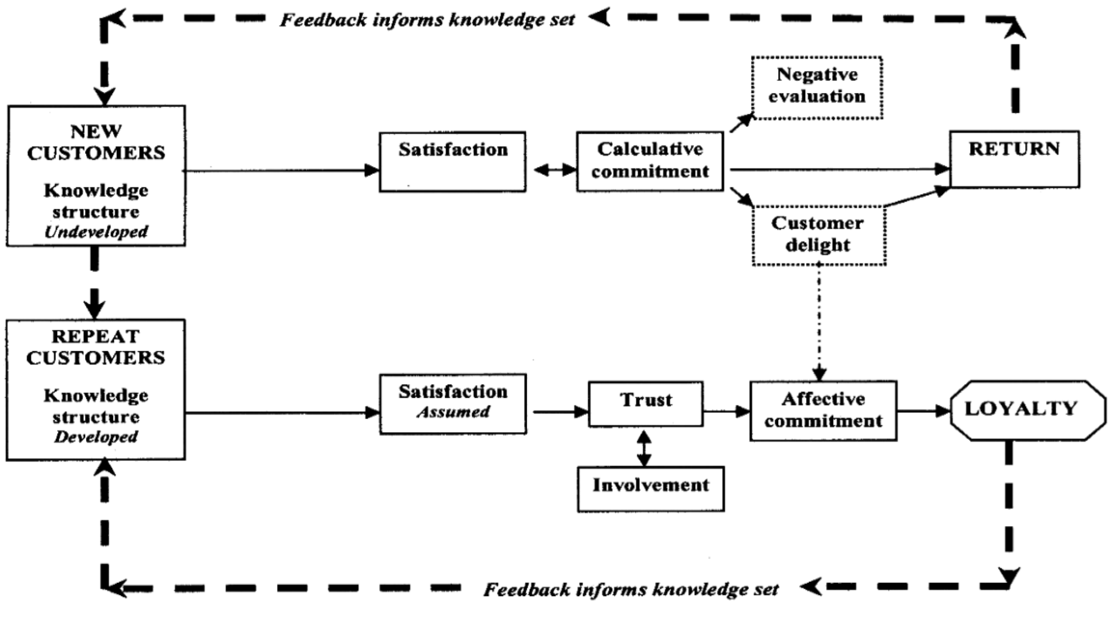


Figure 2. A Conceptual Framework for the Process of Engagement

All in all, the author (2009) was based on previous research in order to make the following four propositions, noting, however, that empirical testing of them and the conceptual model (Fig. 2) should follow:

"Proposition 1: Calculative commitment will have a greater impact than affective commitment in explaining new customers' intention to return and to make positive

recommendations to others.

Proposition 2: For new customers, the experience of delight accelerates the development of commitment and loyalty.

Proposition 3: The higher the level of involvement with the service brand, the greater the degree of brand trust leading to increased levels of customer commitment.

Proposition 4: Affective commitment will have a greater impact than calculative commitment in explaining repeat purchase customers' intention to return and recommend.".

Based on the discussion above, it is apparent that there are various qualitative variables that we should take into account, in order to make an assumption regarding the effect of the marketing campaign implementation in apps on user retention. As the available data are limited from a quality point of view, it is more preferable not to propose a hypothesis for the retention rate of users and, instead, the relevant results of the case study only to be presented.

3.2. Research Method

The goal of this study is to identify the nature of a possible phenomenon that the sequence of observations – data may represent. In particular, we examine how the marketing campaign implementation in mobile apps affects the number of daily users in a trivia game platform. As time is a closely related factor to the objective of this study, Time Series Analysis is used. Although it includes many different methods for modeling and forecasting procedures, the one chosen in this case is ARMA methodology, in an attempt to cope with problems as considerable error of observations and unclear data patterns are. In addition, three dummy variables that have been created for the purpose of the study are included in the models estimated, while at the same time three separate measures, mean daily users before the campaign started, mean daily users after the campaign and –finally– the per user campaign cost are also calculated for further insight. In conclusion, the statistical package used for the analysis is Eviews 7.0.

3.3. Theory the Analysis Was Based On

3.3.1. ARMAX Models

ARMAX (Autoregressive Moving Average with Exogenous Variables) models, like those estimated in this study, are a combination of AR (Autoregressive) and MA (Moving Average) models, in case exogenous variables exist. Firstly, an AR(k) model has the form as below:

$$y_t = \boldsymbol{X}_t\beta + e_t$$

$$e_t = c_1 e_{t-1} + c_2 e_{t-2} + \ldots + c_k e_{t-k} + \varepsilon_t$$

$$\varepsilon_t \sim N\left(0,\sigma^2\right).$$

Furthermore, an MA(l) model has the following form:

$$y_t = \boldsymbol{X}_t \beta + e_t$$

$$e_t = \varepsilon_t + w_1 \varepsilon_{t-1} + w_2 \varepsilon_{t-2} + \ldots + w_l \varepsilon_{t-l}$$

$$\varepsilon_t \sim N\left(0, \sigma^2\right).$$

As we have mentioned before, the combination of those two models above results in the ARMAX(k,l) model, which is written as below:

$$y_t = \boldsymbol{X}_t \beta + e_t$$

$$e_t = c_1 e_{t-1} + c_2 e_{t-2} + \ldots + c_k e_{t-k} + \varepsilon_t + w_1 \varepsilon_{t-1} + w_2 \varepsilon_{t-2} + \ldots + w_l \varepsilon_{t-l}$$

$$\varepsilon_t \sim N\left(0, \sigma^2\right).$$

3.3.2. ARMAX Models with ARCH Innovations (Degiannakis and Xekalaki, 2004)

Regarding the ARCH process, residuals are not considered as independent and normally distributed, with a zero mean and a constant variance. On the contrary, they are considered as the product of an i.i.d. (independent identically distributed) process, which follows the standard normal distribution and a positively measurable function of information, that is available at particular times in the past. This process is written briefly as below:

$$\varepsilon_t = z_t \sigma_t$$

$$z_t \overset{i.i.d.}{\sim} N(0,1)$$

$$\sigma_t^2 = g\left(I_{t-1}\right),$$

where g(x): a function which attributes positive values and

I_t *: the information available at t specific moment.*

Making the assumption that σ_t is a function of information, which is available at particular times in the past, we are able to estimate this function, in terms of a Time Series Analysis.

Regarding the mean and the variance of residuals, we have the following two equations:

$$E\left(\varepsilon_t\right) = E\left(z_t\right) E\left(\sigma_t\right) = 0 E\left(\sigma_t\right) = 0$$

and

$$V\left(\varepsilon_t\right) = E\left(\varepsilon_t^2\right) - \left(E\left(\varepsilon_t\right)\right)^2 = E\left(z_t^2 \sigma_t^2\right) - 0^2 = E\left(z_t^2\right) E\left(\sigma_t^2\right) = V\left(z_t\right) E\left(\sigma_t^2\right) = E\left(\sigma_t^2\right) \equiv \sigma^2.$$

Furthermore, given all the information that is available, the mean and variance of residuals can be written, respectively, as follows:

$$E(\varepsilon_t \mid I_{t-1}) = E(z_t \mid I_{t-1})E(\sigma_t \mid I_{t-1}) = E(z_t)E(\sigma_t \mid I_{t-1}) = 0$$

and

$$V(\varepsilon_t \mid I_{t-1}) = E(\varepsilon_t^2 \mid I_{t-1}) - (E(\varepsilon_t \mid I_{t-1}))^2 = E(z_t^2)E(\sigma_t^2 \mid I_{t-1}) = E(\sigma_t^2 \mid I_{t-1}) = \sigma_t^2.$$

Consequently, there have been built a model for time series, which calculates both the total variance of a time series, σ^2, and the variance of every single moment, σ_t^2.

Most of the time, $g(x)$ is a linear or nonlinear function of the previous values of the following: i) squared residuals, $\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \ldots$, ii) conditional variance, $\sigma_{t-1}^2, \sigma_{t-2}^2, \ldots$, and iii) exogenous variables, $\upsilon_{t-1}, \upsilon_{t-2}, \ldots$, which are all included in the (I_{t-1}) term.

To sum up, an ARMAX model with an ARCH process of innovation can be generally written as follows:

$$y_t = \boldsymbol{X}_t \beta + e_t$$

$$e_t = c_1 e_{t-1} + c_2 e_{t-2} + \ldots + c_k e_{t-k} + \varepsilon_t + w_1 \varepsilon_{t-1} + w_2 \varepsilon_{t-2} + \ldots + w_l \varepsilon_{t-l}$$

$$\varepsilon_t = z_t \sigma_t$$

$$z_t \overset{i.i.d.}{\sim} N(0,1)$$

$$\sigma_t^2 = g(\sigma_{t-1}^2, \sigma_{t-2}^2, \ldots; \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \ldots; \upsilon_{t-1}, \upsilon_{t-2}, \ldots).$$

3.3.3. Assumptions need to be met

In order for the estimation outputs to be reliable, the following assumptions have to be met. The first is the one for a normal distribution of residuals, the second is the one for uncorrelated residuals over time and the third is the one for uncorrelated squared residuals over time.

Statistically, we can test for normality, autocorrelation and heteroskedasticity of residuals, applying the following three tests, respectively.

3.3.3.1. Jarque Bera Normality Test (Jarque and Bera, 1987)

This test can be used to find if the residuals are normally distributed or not. To be more specific, the test examines the null hypothesis, according to which residuals are normally distributed, against the alternative one, according to which they are not. In order for this test to be applied, the Jarque Bera statistical quantity is calculated. This quantity follows the Chi-Square distribution ($\mathbf{X}^2$), with two degrees of freedom. Apart from the examination of these two hypotheses, the normality of residuals can be tested by observing their histogram, which should look like bell-shaped.

To sum up, the Jarque Bera Normality Test can be written briefly, as follows:

$$H_0: e_t \sim \boldsymbol{N}$$
$$H_1: e_t \nsim \boldsymbol{N}$$
$$JB = \frac{N}{6} * \left(s^2 + \frac{(k-3)^2}{4} \right),$$

where N: the number of the independent observations on a random variable, s:skewness and k:kurtosis

$$JB \sim X_2^2$$

If $P_{value} < \alpha$, H_0 rejected for (1-α) confidence interval.

3.3.3.2. Serial Correlation LM Test (Breusch, 1978; Godfrey, 1978)

This test can be used to find serial correlation of any order and it does not make the assumption that the independent variables of the model estimated are not previous values of the dependent one.

In order for the null hypothesis to be tested, according to which residuals are not characterized by serial correlation (autocorrelation) of *i* order, LM test relies on a model that consists of the following variables: i) the residuals of the primary model estimated, which play the role of the dependent variable and ii) the independent variables of the primary model estimated and *i* first previous values of the residuals, which play the role of the independent variables.

Its written form is the following one:

$$\hat{e}_t = \boldsymbol{X}_t \beta^* + \gamma_1 \hat{e}_{t-1} + \ldots + \gamma_i \hat{e}_{t-i} + \varepsilon_t.$$

The Breusch and Godfrey's LM testing function follows Chi-Square ($\mathbf{X}^2$) distribution, with *i* degrees of freedom and it is calculated as the number of observations (T) multiplied with the coefficient of determination (R^2).

To sum up, the Serial Correlation LM Test can be written briefly as below:

$$H_0: \rho_1 = \rho_2 = \ldots = \rho_i = 0$$
$$H_1: \text{at least one } \rho_i \neq 0, \text{ where } \rho_i = Corr(\varepsilon_t, \varepsilon_{t-i})$$
$$\hat{e}_t = X_t \beta^* + \gamma_1 \hat{e}_{t-1} + \ldots + \gamma_i \hat{e}_{t-i} + \varepsilon_t$$
$$TR^2 \sim X_i^2$$

If $P_{value} < \alpha$, H_0 rejected for (1-α) confidence interval.

3.3.3.3. ARCH LM Test (Engle, 1982)

The ARCH (AutoRegressive Conditional Heteroskedasticity) test can be used to find if serial correlation among squared residuals of the initial model estimated exists. More specifically, it examines if the residuals of the model are characterized by a specific type of heteroskedasticity, named conditional.

For this reason, in order for the null hypothesis to be tested, according to which residuals are not characterized by the ARCH form of heteroskedasticity of *i* order, LM Test relies on a model, that consists of the following variables: i) the squared residuals of the initial model estimated, which play the role of the dependent variable and ii) the *i* first previous values of the squared residuals of the initial model estimated, which play the role of the independent variables. Mathematically, it has the following form:

$$\hat{e}_t{}^2 = \gamma_0 + \gamma_1 \hat{e}_{t-1}^2 + \ldots + \gamma_i \, \hat{e}_{t-i}^2 + \varepsilon_t \, .$$

The Engle's LM testing function follows the Chi-Square ($\mathbf{X}^2$) distribution, with *i* degrees of freedom and it is calculated as the number of observations (T) multiplied with the coefficient of determination (R^2). To sum up, the ARCH LM Test can be written briefly as below:

$$H_0: \sigma^2 = c$$

$$H_1: \sigma^2 \neq c$$

$$\hat{e}_t{}^2 = \gamma_0 + \gamma_1 \hat{e}_{t-1}^2 + \ldots + \gamma_i \hat{e}_{t-i}^2 + \varepsilon_t$$

$$TR^2 \sim X_i^2$$

If $P_{value} < \alpha$, H_0 rejected for (1-α) confidence interval.

3.4. Test of Statistical Significance of the Parameters (Green, 2002)

Another test that is applied in the present study is the one that examines if the parameters of the models estimated are statistically significant for (1-α) confidence interval. The statistical quantity calculated for this test is the *t* one and follows the t – student distribution, with n – 2 degrees of freedom. In conclusion, the key points of this specific tests are the following:

$$H_0: \beta_i = 0$$

$$H_1: \beta_i \neq 0$$

$$t_{stat} \sim t_{n-2},$$

where n: the number of the independent observations on a random variable

$$t_{stat} = \hat{\beta}_i \, / \, \hat{s}_t$$

If $P_{value} < \alpha$, H_0 rejected for (1-α) confidence interval.

3.5. Data

The data were tracked by a Big Four US mobile advertising company, which ran the marketing campaign for a brand new trivia game in Apple Store, in order to evaluate the campaign effectiveness. Those proprietary data pertain to three distinct time periods, that is to say three different time series, which are the number of daily users, before, during and after the marketing campaign. This campaign was effective from Sep 23, 2015 until Nov 1, 2015, whereas the number of daily users was tracked from Sep 1, 2015 until Feb 29, 2016.

4. RESULTS OF THE CASE STUDY

As mentioned before, Time Series Analysis is used in this thesis and more specifically, the ARMA methodology is applied. In summary, the analysis is presented step by step, in order to show how we have concluded in the variables used and the models estimated.

In particular, the analysis starts with an AR(1) model, which is evolved into an ARMA(1,3) one, in which the dependent variable is the log of daily users. Afterwards, D_1 and D_2 variables are added, one at a time, in an attempt to examine the campaign effect on daily users and user retention, as well. In addition, D_2 and D_3 variables are included simultaneously in the main model, in order to have an extra insight into the percentage of the retained users.

The first dummy variable, D_1, takes on the value "0" or "1" to indicate the absence or presence of the marketing campaign implementation, respectively. The second dummy variable, D_2, takes on the value "1" in case the campaign implementation has been completed and "0" otherwise. Finally, the third dummy variable, D_3, takes on the value "1" in case the campaign has not started yet and the value "0" otherwise.

At the same time, the necessary assumptions for reliable results to be obtained are tested.

4.1. Mean Daily Users Before and After the Campaign Implementation

In an attempt to provide a primary insight into how the marketing campaign implementation has affected the number of daily users in the trivia game platform, we have calculated the two following measures: i) the mean daily users before the campaign started and ii) the mean daily users after the campaign ended.

As we know, the mean of a sample of n values is the sum of these sampled values divided by the size of the sample. Its writing form is the following:

$$\bar{x} = (\sum_{i=1}^{n} x_i \,) \,/\, n \,,$$

where $\bar{x}$: the mean of the sampled values

x_i: the value i

and

n: the number of items in the sample.

Using this type of the mean, the mean daily users before the campaign started is calculated to be equal to 25.5 users per day, while the mean daily users after the campaign ended is calculated to be equal to 109.13 users per day.

If we compare the values of the two measures above, we understand that the campaign has obviously positively affected the daily users, as the number of them after its implementation is calculated to be almost four times larger than the one before it.

Let us see, however, if the results of the empirical analysis are consistent to the finding above.

4.2. Per User Campaign Cost

The total cost of the marketing campaign implementation in mobile apps was equal to twenty thousand US dollars (USD 20 K). In order to be able to calculate the campaign cost per user, we will need to calculate first the sum of daily users for the period Sep 23, 2015 until Nov 1, 2015, that is to say for the period the campaign was effective. Consequently, the campaign cost per user is calculated as a fraction, where the numerator is the total campaign cost and the denominator the number of users, for the dates of interest.

Finally, as the number of daily users for this period is equal to 24,963, the campaign cost per user for every single day is equal to 0.80$. As we cannot identify if the users of a specific date were all new ones or they had visited the trivia game platform before, all we can notice is that the company actually paid 0.80$ for every user of every single day, for the period we referred to.

4.3. AR(1) Model Estimation with Maximum Likelihood Method of Estimation

We, first, started the analysis running a simple AR(1) model, in which the log of daily users is the dependent variable (Y_t) and its first previous value the independent one (Y_{t-1}). Table 1 presents the output of the AR(1) model estimation, using Maximum Likelihood Method (ML).

Table 1. Output of AR(1) Model Estimation with Maximum Likelihood Method

Dependent Variable: Log Daily Users
Method: ML
Sample (adjusted): 2 182
Included observations: 181 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.998869	0.435995	11.46541	0.0000
AR(1)	0.947858	0.021005	45.12578	0.0000

R-squared	0.919200	Mean dependent var	4.803435
Adjusted R-squared	0.918748	S.D. dependent var	1.053340
S.E. of regression	0.300251	Akaike info criterion	0.442595
Sum squared resid	16.13702	Schwarz criterion	0.477938
Log likelihood	-38.05487	Hannan-Quinn criter.	0.456924
F-statistic	2036.336	Durbin-Watson stat	2.325521
Prob(F-statistic)	0.000000		
Inverted AR Roots	.95		

According to the estimation output above, the model can be written as follows:

$$Y_t = 4.998 + e_t,$$
$$\text{where } e_t = 0.947e_{t-1} + \varepsilon_t .$$

As the extremely high value of the R^2 indicates (0.919), the first lag of the dependent variable interprets the values of the latter to a great extent. The value of the coefficient of the AR(1) term (0.947 > 0.00) means that if Y_{t-1} variable increases by one unit then Y_t increases by 0.947, as well. As we can see, the first previous value of the dependent variable is statistically significant for any level of statistical significance (Prob. = 0.00).

However, while estimating the model with ML Method, the three assumptions that were mentioned before, are not met.

Regarding the assumption of the normally distributed residuals and as we can see in Figure 3, the null hypothesis of normality is rejected for any level of statistical significance (Prob. = 0.00). The histogram of residuals supports this conclusion, as well.

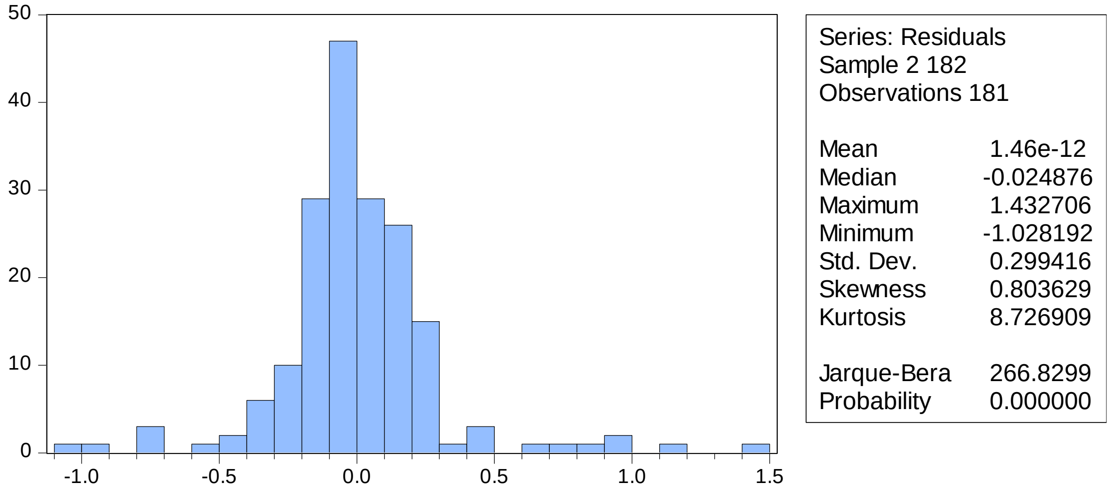


Figure 3. Histogram – Normality test of AR(1) Model Estimated with ML Method

The histogram above does not look like bell-shaped. In addition, both of the measures skewness and kurtosis do not have a value, which would make us believe that the residuals of this model are normally distributed (skewness ≈ 0.803 > 0.00, kurtosis ≈ 8.72 > 3.00).

Regarding the assumption of autocorrelation of residuals, we find that they are closely correlated over time. This can be proved by both the correlogram of residuals and the Serial Correlation LM Test. The results of both of those methods are shown in Table 2 and Table 3, respectively.

Table 2. Correlogram of Standardized Residuals of AR(1) Model Estimated with Maximum Likelihood Method

Sample: 2 182
Included observations: 181
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob.
| **| |	| **| |	1	-0.167	-0.167	5.1115	
| |** |	| |** |	2	0.217	0.194	13.809	0.000
| **| |	| *| |	3	-0.156	-0.101	18.354	0.000
| *| |	| *| |	4	-0.029	-0.111	18.508	0.000
| |* |	| |* |	5	0.054	0.097	19.064	0.001
| *| |	| *| |	6	-0.102	-0.081	21.032	0.001
| |** |	| |** |	7	0.220	0.166	30.268	0.000

For a choice of seven lags, the correlogram of residuals indicates that the null hypothesis is rejected for a significance level equal to 0.01 and 0.05.

Similarly, applying the Serial Correlation LM Test, the result is the same.

Table 3. Breusch - Godfrey Serial Correlation LM Test of AR(1) Model Estimated with Maximum Likelihood Method

F-statistic	3.517215	Prob. F(7,172)	0.0015
Obs*R-squared	22.66453	Prob. Chi-Square(7)	0.0019

Test Equation:
Dependent Variable: Residuals
Method: ML
Sample: 2 182
Included observations: 181
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.014010	0.421711	0.033221	0.9735
AR(1)	0.003832	0.026213	0.146207	0.8839
RESID(-1)	-0.098076	0.081171	-1.208256	0.2286
RESID(-2)	0.186589	0.080393	2.320975	0.0215
RESID(-3)	-0.135094	0.081553	-1.656522	0.0994
RESID(-4)	-0.063022	0.081365	-0.774558	0.4397
RESID(-5)	0.049853	0.079886	0.624058	0.5334
RESID(-6)	-0.065356	0.079323	-0.823927	0.4111
RESID(-7)	0.164012	0.078273	2.095395	0.0376

R-squared	0.125218	Mean dependent var	1.46E-12
Adjusted R-squared	0.084531	S.D. dependent var	0.299416
S.E. of regression	0.286482	Akaike info criterion	0.386162
Sum squared resid	14.11636	Schwarz criterion	0.545204
Log likelihood	-25.94769	Hannan-Quinn criter.	0.450641
F-statistic	3.077563	Durbin-Watson stat	2.004846
Prob(F-statistic)	0.002881		

As the results in Table 3 indicate, the null hypothesis of no serial correlation among residuals is rejected again for both of $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance, just as in the case of the table before, as Prob. = 0.001.

The test for heteroskedasticity of residuals, now, shows that this assumption is violated, too. The results of the ARCH Heteroskedasticity Test are presented in the Table 4:

Table 4. Heteroskedasticity Test: ARCH of AR(1) Model Estimated with Maximum Likelihood Method

F-statistic	33.67459	Prob. F(1,178)	0.0000
Obs*R-squared	28.63559	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: Residuals^2
Method: ML
Sample (adjusted): 3 182
Included observations: 180 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.053749	0.018148	2.961644	0.0035
RESID^2(-1)	0.398841	0.068730	5.802981	0.0000

R-squared	0.159087	Mean dependent var	0.089347
Adjusted R-squared	0.154362	S.D. dependent var	0.249194
S.E. of regression	0.229155	Akaike info criterion	-0.097789
Sum squared resid	9.347124	Schwarz criterion	-0.062312
Log likelihood	10.80100	Hannan-Quinn criter.	-0.083404
F-statistic	33.67459	Durbin-Watson stat	2.089039
Prob(F-statistic)	0.000000		

As we mentioned before, the null hypothesis of no serial correlation among the squared residuals or heteroskedasticity among residuals, in other words, is rejected for any level of significance (Prob. = 0.00).

To encounter with the violation of the autocorrelation assumption, we added a MA(*l*) term. More specifically, we first tried to only insert the term MA(3), with the coefficients of MA(1) and MA(2) terms being equal to zero ($w_1 = w_2 = 0$), as we observed that in this specific lag, the residuals of the main model estimated seem to be autocorrelated for the first time. Furthermore, to encounter with the violation of the heteroskedasticity assumption, the ARMA(1,3) model was estimated with the ML – ARCH Method.

4.4. ARMA(1,3) Model Estimation with ML – ARCH Method

The estimation output, with ML - ARCH Method, after the addition of the MA(3) term in our AR(1) model, is presented in Table 5.

Table 5. Output of ARMA(1,3) Model Estimation with ML - ARCH Method

Dependent Variable: Log Daily Users
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 2 182
Included observations: 181 after adjustments
Convergence achieved after 45 iterations
MA Backcast: -1 1
Presample variance: backcast (parameter = 0.7)
GARCH = C(4) + C(5)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	4.758028	0.309980	15.34948	0.0000
AR(1)	0.959922	0.011570	82.96949	0.0000
MA(3)	-0.204784	0.041130	-4.978998	0.0000
	Variance Equation			
C	0.029156	0.002895	10.07138	0.0000
RESID(-1)^2	0.603307	0.168611	3.578101	0.0003
R-squared	0.921812	Mean dependent var		4.803435
Adjusted R-squared	0.920934	S.D. dependent var		1.053340
S.E. of regression	0.296186	Akaike info criterion		-0.068026
Sum squared resid	15.61525	Schwarz criterion		0.020330
Log likelihood	11.15640	Hannan-Quinn criter.		-0.032205
Durbin-Watson stat	2.276891			
Inverted AR Roots	.96			
Inverted MA Roots	.59	-.29-.51i	-.29+.51i	

The written form of the ARMA(1,3) model, estimated with ML – ARCH Method, is the following:

$$Y_t = 4.758 + e_t\,,$$
$$where\ e_t = 0.959e_{t-1} - 0.204\varepsilon_{t-3} + \varepsilon_t$$
$$\varepsilon_t = z_t\sigma_t$$
$$z_t \overset{i.i.d.}{\sim} N(0,1)$$
$$\sigma_t^2 = 0.029 + 0.603\varepsilon_{t-1}^2.$$

The constant term, the first lag of the dependent variable and the MA(3) term are all statistically significant for any level of significance (Prob. = 0.00). In addition, the R^2 value has been slightly increased (from 0.919 to 0.921).

We apply, again, the normality, autocorrelation and heteroskedasticity tests to examine if the addition of the MA(3) term and the estimation with ML – ARCH Method resulted in the assumptions to be met. As we will see in the tables below, the residuals of the ARMA(1,3) are not correlated over time for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance and the squared residuals are not characterized by serial correlation for any level of significance, as well. However, the residuals are still not distributed normally.

Figure 4 indicates that residuals are not normally distributed, despite the amendments we have made.

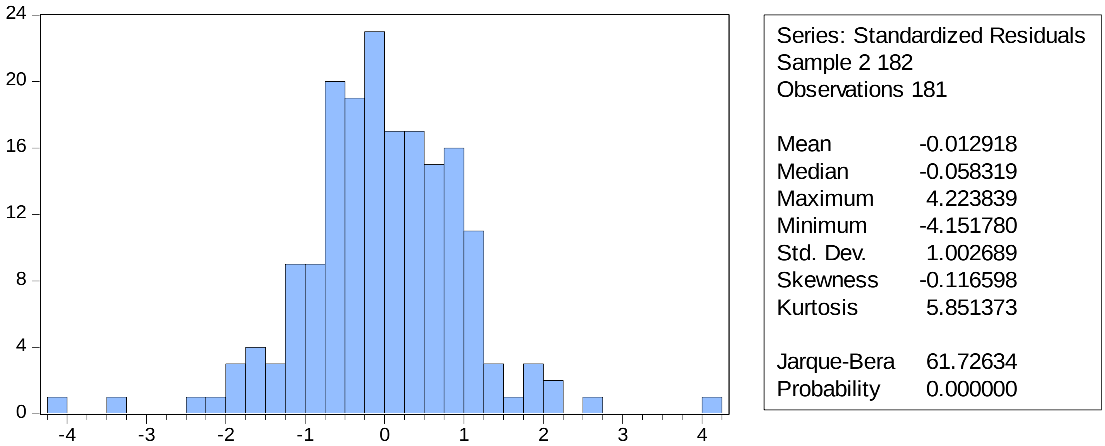


Figure 4. Histogram – Normality test of ARMA(1,3) Model Estimated with ML – ARCH Method

Figure 4 indicates the non-normality of the ARMA(1,3) model residuals. The probability of the Jarque – Bera statistical quantity is equal to zero (Prob. = 0.00), which means that the null hypothesis of the normality test is rejected for any level of significance. Similarly with the case of the AR(1) model residuals, the shape of the histogram does not look like bell-shaped. In parallel, the value of the skewness measure is quite close to zero (skewness = -0.111 ≈ 0.00), whereas kurtosis is greater enough than three (kurtosis = 5.851 > 3.00). The non-normality of the residuals of the models estimated in this study is a characteristic we did not manage to encounter with.

Furthermore, the null hypothesis of no serial correlation among the residuals is not rejected anymore, for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance and for seven lags, as the Correlogram of Residuals in Table 6 indicates.

Table 6. Correlogram of Standardized Residuals of ARMA(1,3) Model Estimated with ML - ARCH Method

Sample: 2 182
Included observations: 181
Q-statistic probabilities adjusted for 2 ARMA term(s)

<table>
<tr><th>Autocorrelation</th><th>Partial Correlation</th><th></th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Prob.</th></tr>
<tr><td>| *| |</td><td>| *| |</td><td>1</td><td>-0.058</td><td>-0.058</td><td>0.6295</td><td></td></tr>
<tr><td>| |* |</td><td>| |* |</td><td>2</td><td>0.075</td><td>0.072</td><td>1.6784</td><td></td></tr>
<tr><td>| |* |</td><td>| |* |</td><td>3</td><td>0.058</td><td>0.067</td><td>2.3096</td><td>0.129</td></tr>
<tr><td>| *| |</td><td>| *| |</td><td>4</td><td>-0.094</td><td>-0.094</td><td>3.9776</td><td>0.137</td></tr>
<tr><td>| |* |</td><td>| |* |</td><td>5</td><td>0.051</td><td>0.032</td><td>4.4652</td><td>0.215</td></tr>
<tr><td>| *| |</td><td>| | |</td><td>6</td><td>-0.012</td><td>0.004</td><td>4.4916</td><td>0.344</td></tr>
<tr><td>| |** |</td><td>| |** |</td><td>7</td><td>0.145</td><td>0.153</td><td>8.4942</td><td>0.131</td></tr>
</table>

Similarly, the ARCH Heteroskedasticity Test indicates that the residuals of the ARMA(1,3) model, estimated with ML – ARCH Method, are now homoscedastic, as the null hypothesis is not rejected anymore for any level of significance. The above are presented in Table 7.

Table 7. Heteroskedasticity Test: ARCH of ARMA(1,3) Model Estimated with ML – ARCH Method

F-statistic	0.248289	Prob. F(1,178)	0.6189
Obs*R-squared	0.250729	Prob. Chi-Square(1)	0.6166

Test Equation:
Dependent Variable: WGT_RESID^2
Method: ML
Sample (adjusted): 3 182
Included observations: 180 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.966703	0.181526	5.325422	0.0000
WGT_RESID^2(-1)	0.037320	0.074897	0.498286	0.6189

R-squared	0.001393	Mean dependent var	1.003893
Adjusted R-squared	-0.004217	S.D. dependent var	2.215382
S.E. of regression	2.220049	Akaike info criterion	4.443984
Sum squared resid	877.2937	Schwarz criterion	4.479462
Log likelihood	-397.9586	Hannan-Quinn criter.	4.458369
F-statistic	0.248289	Durbin-Watson stat	2.001422
Prob(F-statistic)	0.618898		

As Prob. = 0.618 > $\alpha = 0.01$ and $\alpha = 0.05$, we accept that the residuals of our model are homoscedastic for a 0.99 and a 0.95 confidence interval, respectively.

So, we have concluded in an ARMA(1,3) Model, which has been estimated with ML – ARCH Method, the residuals of which are not correlated over time, they are homoscedastic, but not normally distributed for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance. In addition, as we have mentioned before, the dependent variable (Log Daily Users) is positively affected to a great extent by its first lag, as the coefficient value is equal to 0.959, while the term MA(3) affects it in a negative way, since the value of its coefficient is equal to -0.204.

In the next sections, we extended our model by adding the dummy variables as independent ones, as we wanted to indicate the absence or presence of the campaign effect that may shifts the outcome, regarding the number of daily users, and investigate the effect of the campaign on user retention, as well. For this reason, we firstly added D_1 and D_2 variables, one at a time and finally, we estimated our model with both of the D_2 and D_3 variables added simultaneously. The results are interesting enough.

4.5. ARMA(1,3) Model with the First Dummy Variable Used as an Independent One

As we have mentioned before, we have created the first dummy variable (D_1), in order to examine the effect of the marketing campaign implementation on the number of daily users. For this reason, this variable takes on the value "1" during the campaign and "0" otherwise. In essence, it only affects the values of the dependent variable, in case of a date that the marketing campaign is effective. Its writing form is the following one:

$$D_{1t} = \begin{cases} \text{1, if the campaign is effective.} \\ \text{0, otherwise.} \end{cases}$$

We estimate our model again and the results are presented in Table 8.

Table 8. Output of ARMA(1,3) Model Estimation with ML – ARCH Method after D_1 Variable Was Added

Dependent Variable: Log Daily Users
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 2 182
Included observations: 181 after adjustments
Convergence achieved after 36 iterations
MA Backcast: -1 1
Presample variance: backcast (parameter = 0.7)
GARCH = C(5) + C(6)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	4.630729	0.083957	55.15579	0.0000
D1	1.269742	0.075895	16.73023	0.0000
AR(1)	0.873091	0.015078	57.90397	0.0000
MA(3)	-0.183297	0.032835	-5.582434	0.0000
	Variance Equation			
C	0.019850	0.003751	5.291578	0.0000
RESID(-1)^2	1.025591	0.224681	4.564644	0.0000

R-squared	0.926773	Mean dependent var	4.803435
Adjusted R-squared	0.925531	S.D. dependent var	1.053340
S.E. of regression	0.287445	Akaike info criterion	-0.120750
Sum squared resid	14.62460	Schwarz criterion	-0.014723
Log likelihood	16.92788	Hannan-Quinn criter.	-0.077764
Durbin-Watson stat	2.329049		

According to the table above, the written form of our model is the following:

$$Y_t = 4.630 + 1.269D_{1t} + e_t\,,$$
$$\text{where } e_t = 0.873e_{t-1} - 0.183\varepsilon_{t-3} + \varepsilon_t$$

$$\varepsilon_t = z_t\sigma_t$$
$$z_t \overset{i.i.d.}{\sim} N(0,1)$$
$$\sigma_t^2 = 0.019 + 1.025\varepsilon_{t-1}^2.$$

The coefficient of D_1 variable can be interpreted as a measure of sensitivity of the dependent one to a change in D_1 values. That means that D_1 coefficient can be interpreted as elasticity for the values of the dependent variable. As we can see in Table 8, D_1 coefficient is equal to 1.269 > 1.00. For this reason, "Log Daily Users" variable can be characterized as an elastic one, according to the definition of which it actually responds more than proportionally to changes in D_1 values. So, we actually capture a change in log daily users, if we consider as a benchmark the date marketing campaign started. Furthermore, all the independent variables are statistically significant for any level of significance (Prob. = 0.00), while the R^2 is presented as slightly greater than in the cases of the two previous estimation outputs ($R^2 = 0.926$).

The model meets both of the assumptions of no serial correlation and homoscedasticity of residuals. However, as we have mentioned before, the assumption of normally distributed residuals is still violated.

The figure and tables below present the results regarding the three tests of normality, autocorrelation and heteroskedasticity of residuals, respectively.

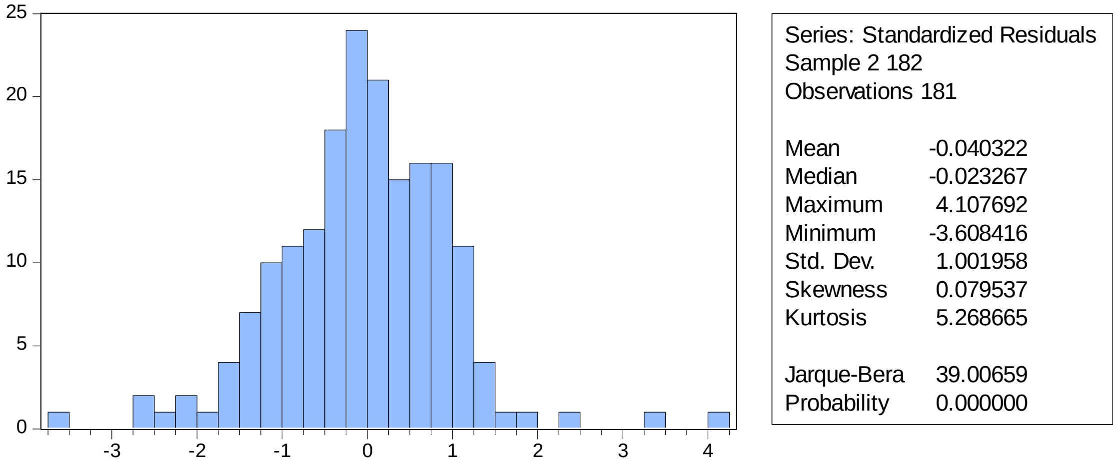


Figure 5. Histogram – Normality test of ARMA(1,3) Model Estimated with ML – ARCH Method after D_1 Variable Was Added

The probability of Jarque – Bera statistical quantity remains equal to zero (Prob. = 0.00), resulting in the rejection of normality assumption for any level of significance, while the shape of the histogram confirms this outcome, as well.

Table 9. Correlogram of Standardized Residuals of ARMA(1,3) Model Estimated with ML – ARCH Method after D_1 Variable Was Added

Sample: 2 182
Included observations: 181
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob.
| | |	| | |	1	0.008	0.008	0.0110	
| |* |	| |* |	2	0.096	0.096	1.7154	
| |** |	| |** |	3	0.131	0.131	4.9301	0.026
| *| |	| *| |	4	-0.027	-0.037	5.0653	0.079
| |* |	| |* |	5	0.116	0.093	7.6031	0.055
| |* |	| |* |	6	0.042	0.032	7.9389	0.094
| |* |	| |* |	7	0.118	0.111	10.574	0.061

For seven lags, the residuals are not autocorrelated, for α = 0.01 level of significance.

Table 10. Heteroskedasticity Test: ARCH of ARMA(1,3) Model Estimated with ML – ARCH Method after D_1 Variable Was Added

F-statistic	0.321123	Prob. F(1,178)	0.5716
Obs*R-squared	0.324146	Prob. Chi-Square(1)	0.5691

Test Equation:
Dependent Variable: WGT_RESID^2
Method: ML
Sample (adjusted): 3 182
Included observations: 180 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.047620	0.171690	6.101804	0.0000
WGT_RESID^2(-1)	-0.042433	0.074881	-0.566677	0.5716

R-squared	0.001801	Mean dependent var	1.005404
Adjusted R-squared	-0.003807	S.D. dependent var	2.071384
S.E. of regression	2.075323	Akaike info criterion	4.309160
Sum squared resid	766.6402	Schwarz criterion	4.344637
Log likelihood	-385.8244	Hannan-Quinn criter.	4.323545
F-statistic	0.321123	Durbin-Watson stat	2.001294
Prob(F-statistic)	0.571648		

Table 10 indicates that, for one lag, residuals are homoscedastic, for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance, as Prob. = 0.569.

In the next section, D_1 variable, which can be characterized as the marketing campaign effect on the dependent one, will be replaced by D_2 variable, in order to examine the effect of the campaign on user retention. As we will see, marketing campaign implementation in apps has a positive effect on the number of daily users and this attempt seems to be effective enough in terms of creating loyal users, as well.

4.6. ARMA(1,3) Model with the Second Dummy Variable Used as an Independent One

The second dummy variable (D_2) has been created in order for the retention rate of daily users in the trivia game platform to be evaluated. In particular, we will examine how marketing campaign implementation affected the number of daily users after the campaign ended. For this reason, D_2 variable takes on the value "1" in the latter case and "0" otherwise, while its writing form is the following:

$$D_{2t} = \begin{cases} 1, \textit{ if the campaign has ended.} \\ 0, \textit{ otherwise.} \end{cases}$$

The results of the estimation of the ARMA(1,3) model, with D_2 variable as an independent one, are presented in the Table 11.

Table 11. Output of ARMA(1,3) Model Estimation with ML – ARCH Method after D_2 Variable Was Added

Dependent Variable: Log Daily Users
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 2 182
Included observations: 181 after adjustments
Convergence achieved after 64 iterations
MA Backcast: -1 1
Presample variance: backcast (parameter = 0.7)
GARCH = C(5) + C(6)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	5.317804	0.232524	22.86993	0.0000
D2	-0.605938	0.200274	-3.025544	0.0025
AR(1)	0.950597	0.009979	95.26338	0.0000
MA(3)	-0.218618	0.033216	-6.581638	0.0000
	Variance Equation			
C	0.026664	0.003017	8.836723	0.0000
RESID(-1)^2	0.647550	0.184275	3.514050	0.0004

R-squared	0.924310	Mean dependent var	4.803435
Adjusted R-squared	0.923027	S.D. dependent var	1.053340
S.E. of regression	0.292238	Akaike info criterion	-0.107992
Sum squared resid	15.11637	Schwarz criterion	-0.001964
Log likelihood	15.77328	Hannan-Quinn criter.	-0.065006
Durbin-Watson stat	2.298851		

Inverted AR Roots	.95		
Inverted MA Roots	.60	-.30+.52i	-.30-.52i

According to Table 11, the written form of our model is the following:

$$Y_t = 5.317 - 0.605D_{2t} + e_t\,,$$
$$\text{where } e_t = 0.95e_{t-1} - 0.218\varepsilon_{t-3} + \varepsilon_t$$
$$\varepsilon_t = z_t\sigma_t$$
$$z_t \overset{i.i.d.}{\sim} N(0,1)$$
$$\sigma_t^2 = 0.026 + 0.647\varepsilon_{t-1}^2.$$

As Prob. = 0.0025, D_2 variable is statistically significant for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance. Its coefficient is equal to -0.605, which means that although the campaign ended, a 39.5% of daily users still remains. The value of R^2 is equal to 0.924, a little lower than in the case of D_1 variable.

Both of the assumptions of no serial correlation and homoscedasticity of residuals are met. Similarly to the case of D_1 variable, however, the assumption of normality of residuals is violated.

The next figure and tables present the results regarding the three tests of normality, autocorrelation and heteroskedasticity of residuals, respectively.

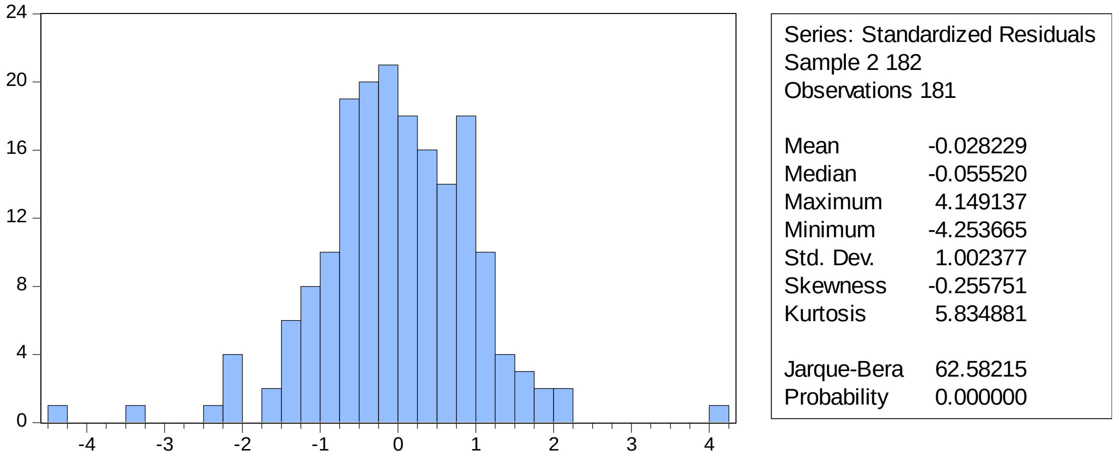


Figure 6. Histogram – Normality test of ARMA(1,3) Model Estimated with ML – ARCH Method after D_2 Variable Was Added

The probability of Jarque – Bera statistical quantity is equal to zero (Prob. = 0.00). For this reason, the null hypothesis of the normality test is rejected for any level of significance. The shape of the histogram supports this result, as it does not remind us a bell shape, with a mean equal to zero and a constant variance.

Table 12. Correlogram of Standardized Residuals of ARMA(1,3) Model Estimated with ML – ARCH Method after D_2 Variable Was Added

Sample: 2 182
Included observations: 181
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob.
| *| |	| *| |	1	-0.045	-0.045	0.3725	
| |* |	| |* |	2	0.066	0.064	1.1849	
| |* |	| |* |	3	0.061	0.067	1.8689	0.172
| *| |	| *| |	4	-0.085	-0.085	3.2301	0.199
| |* |	| |* |	5	0.063	0.048	3.9683	0.265
| *| |	| *| |	6	-0.024	-0.011	4.0731	0.396
| |* |	| |** |	7	0.129	0.133	7.2471	0.203

For seven lags, the residuals are not autocorrelated, for α = 0.01 and α = 0.05 levels of significance.

Table 13. Heteroskedasticity Test: ARCH of ARMA(1,3) Model Estimated with ML – ARCH Method after D_2 Variable Was Added

F-statistic	0.111173	Prob. F(1,178)	0.7392
Obs*R-squared	0.112352	Prob. Chi-Square(1)	0.7375

Test Equation:
Dependent Variable: WGT_RESID^2
Method: ML
Sample (adjusted): 3 182
Included observations: 180 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.979638	0.181613	5.394093	0.0000
WGT_RESID^2(-1)	0.024981	0.074922	0.333427	0.7392

R-squared	0.000624	Mean dependent var	1.004523
Adjusted R-squared	-0.004990	S.D. dependent var	2.215815
S.E. of regression	2.221337	Akaike info criterion	4.445145
Sum squared resid	878.3125	Schwarz criterion	4.480622
Log likelihood	-398.0630	Hannan-Quinn criter.	4.459529
F-statistic	0.111173	Durbin-Watson stat	2.000244
Prob(F-statistic)	0.739205		

Table 13 presents that, for one lag testing, residuals are homoscedastic, for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance, as Prob. = 0.737.

In this section, we examined how D_1 and D_2 variables, one at a time, affect the dependent one. We found that, as D_1 coefficient is equal to 1.269, log daily users are positively increased during the marketing campaign implementation. Furthermore, we found that, as D_2 coefficient is equal to -0.605, a 39.5% of the log daily users remains after the campaign ended. Marketing executives are concerned both for the campaign effect and the retention rate of users, in an attempt return of investment (ROI) to be satisfying. For this reason, both of the two findings above are interesting enough, as they suggest that apps are a means of successful advertising for two reasons. Firstly, an increase of daily users is being observed and secondly, a quite large percentage of them (39.5%) is presented as being loyal even after the campaign is not effective anymore.

What if, however, D_2 and D_3 variables were added in the model at the same time?

4.7. ARMA(1,3) Model with the Second and Third Dummy Variables Used as Independent Ones

In the previous sections, we estimated models in which D_1 and D_2 variables were included, one at a time. According to the estimation outputs, they were statistically significant, a result which indicates that a campaign effect actually exists and an effect on user retention, as well. Below, we will estimate our last model, in which D_2 and D_3 variables will be included in the main model simultaneously.

The third dummy variable (D_3) has been created in order for the retention rate of daily users in the trivia game platform to be evaluated, in case the period of the campaign is considered as a benchmark. For this reason, D_3 variable takes on the value "1" for dates before the campaign launch and "0" otherwise, while its writing form is the following:

$$D_{3t} = \begin{cases} 1, \text{ before the campaign started.} \\ 0, \text{ otherwise.} \end{cases}$$

Table 14 presents the estimation output of the ARMA(1,3) model, in which the second and third dummy variables we have created are included as independent ones.

Table 14. Output of ARMA(1,3) Model Estimation with ML – ARCH Method after D_2 and D_3 Variables Were Added

Dependent Variable: Log Daily Users				
Method: ML – ARCH				
Sample (adjusted): 9/02/2015 2/29/2016				
Included observations: 181 after adjustments				
Convergence achieved after 120 iterations				
MA Backcast: 8/30/2015 9/01/2015				
Presample variance: backcast (parameter = 0.7)				
GARCH = C(6) + C(7)*RESID(-1)^2				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	5.489905	0.189468	28.97534	0.0000
D2	-0.762472	0.128089	-5.952693	0.0000
D3	-1.437670	0.502030	-2.863712	0.0042
AR(1)	0.928208	0.019748	47.00162	0.0000
MA(3)	-0.193758	0.043261	-4.478858	0.0000
	Variance Equation			
C	0.027157	0.002759	9.842172	0.0000
RESID(-1)^2	0.604197	0.179078	3.373934	0.0007
R-squared	0.929611	Mean dependent var		4.803435
Adjusted R-squared	0.928011	S.D. dependent var		1.053340
S.E. of regression	0.282619	Akaike info criterion		-0.123377
Sum squared resid	14.05772	Schwarz criterion		0.000322
Log likelihood	18.16563	Hannan-Quinn criter.		-0.073227
Durbin-Watson stat	2.530883			
Inverted AR Roots	.93			
Inverted MA Roots	.58	-.29-.50i	-.29+.50i	

The written form of the model is the following:

$$Y_t = 5.489 - 0.762D_{2t} - 1.437D_{3t} + e_t\,,$$
$$\text{where } e_t = 0.928e_{t-1} - 0.193\varepsilon_{t-3} + \varepsilon_t$$

$$\varepsilon_t = z_t\sigma_t$$

$$z_t \overset{i.i.d.}{\sim} N(0,1)$$

$$\sigma_t^2 = 0.027 + 0.604\varepsilon_{t-1}^2.$$

As we can see, both of D_2 and D_3 variables are statistically significant for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance, as Prob. = 0.00 and Prob. = 0.004, respectively. More specifically, the coefficient of D_2 variable is equal to -0.762, while the one of D_3

variable is equal to -1.437. In essence, we have found that in case we control for the fact that log daily users are less for the dates before the campaign launch than after (D_3 is negative and significant), there is still a retained percentage after the campaign ending, which is equal to 23.8%. It is quite remarkable the fact that the addition of the third dummy variable has affected the coefficient of the second one, as retained users are less in this case than in the previous one or, in other words, the slope for D_2 dummy variable in this case is larger than in the case before ($|-0.605| < |-0.762|$).

Conclusively, the results of the analysis in this part of the study are generally consistent with those in the previous one, as both of the estimated models have shown that there is a percentage of users that actually can be characterized as loyal, as they actually return to join the trivia game platform even in case the marketing campaign is not effective anymore.

At this point, we will apply the tests of normality, autocorrelation and heteroskedasticity of residuals.

According to Figure 7, the null hypothesis of the normality test is rejected, as Prob. = 0.00, while the shape of the histogram supports this result. The values of skewness and kurtosis are equal to -0.134 and 5.879, respectively.

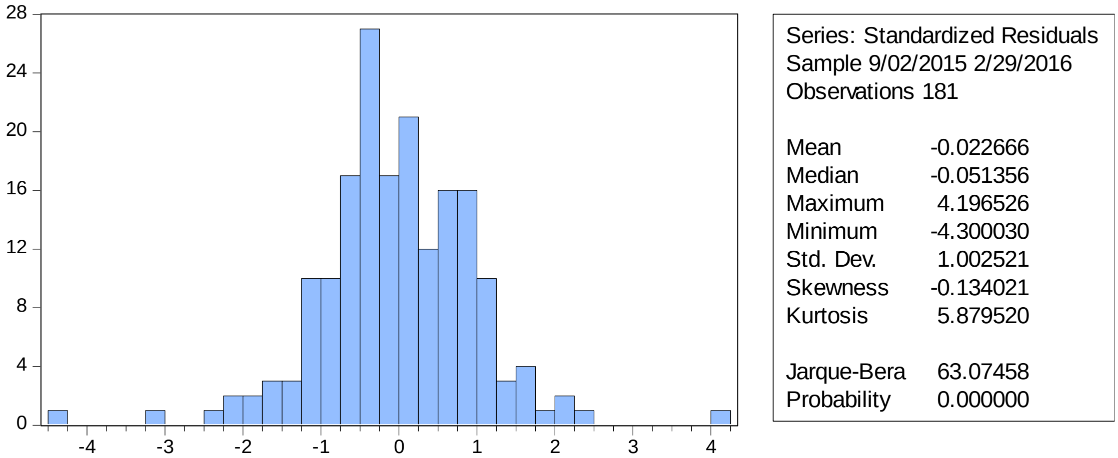


Figure 7. Histogram – Normality test of ARMA(1,3) Model Estimated with ML – ARCH Method after D_2 and D_3 Variables Were Added

For seven lags, Table 15 presents the result of the Serial Correlation LM Test, according to which residuals are not correlated over time, for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance.

Table 15. Correlogram of Standardized Residuals of ARMA(1,3) Model Estimated with ML – ARCH Method after D_2 and D_3 Variables Were Added

Sample: 9/02/2015 2/29/2016
Included observations: 181
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob.
| *| |	| *| |	1	-0.076	-0.076	1.0659	
| |* |	| |* |	2	0.071	0.066	2.0030	
| |* |	| |* |	3	0.078	0.089	3.1487	0.076
| *| |	| *| |	4	-0.118	-0.112	5.7609	0.056
| |* |	| |* |	5	0.079	0.053	6.9482	0.074
| |* |	| |* |	6	0.017	0.037	7.0007	0.136
| |* |	| |* |	7	0.080	0.094	8.2155	0.145

Furthermore, for one lag, the ARCH Heteroskedasticity Test results in the acceptance of the null hypothesis for $\alpha = 0.01$ and $\alpha = 0.05$ levels of significance. That is to say that squared residuals are not correlated over time or, in other words, residuals are homoscedastic, as Prob. = 0.652.

Table 16. Heteroskedasticity Test: ARCH of ARMA(1,3) Model Estimated with ML – ARCH Method after D_2 and D_3 Variables Were Added

Heteroskedasticity Test: ARCH

F-statistic	0.200399	Prob. F(1,178)	0.6549
Obs*R-squared	0.202422	Prob. Chi-Square(1)	0.6528

Test Equation:
Dependent Variable: WGT_RESID^2
Method: ML
Date: 02/02/17 Time: 00:51
Sample (adjusted): 9/03/2015 2/29/2016
Included observations: 180 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.970627	0.182036	5.332062	0.0000
WGT_RESID^2(-1)	0.033535	0.074911	0.447659	0.6549

R-squared	0.001125	Mean dependent var	1.004028
Adjusted R-squared	-0.004487	S.D. dependent var	2.222720
S.E. of regression	2.227701	Akaike info criterion	4.450866
Sum squared resid	883.3520	Schwarz criterion	4.486344
Log likelihood	-398.5780	Hannan-Quinn criter.	4.465251
F-statistic	0.200399	Durbin-Watson stat	2.000930
Prob(F-statistic)	0.654944		

5. BRIEF DISCUSSION

5.1. Summary of the Thesis

Based on previous research, which actually suggests that mobile apps lead to engaged customers and that engagement is a requirement, in order for marketing campaign to be effective, the basic hypothesis that is examined in this study is that mobile applications are truly a quite effective means of communication, in case it is for the launch of a marketing campaign.

The effectiveness of the latter can be attributed to both of the new customer acquisition and the retention of the already existing ones. Regarding the campaign effect on daily users, the empirical results of the case study lead in the acceptance of the hypothesis mentioned in the previous paragraph, while the positive effect on user retention becomes evident as well.

The preexisting literature, though, suggests that the evaluation of the effect on the retention rate of users is much more difficult than the campaign one, as in order for the former to be more accurate the availability of qualitative variables is required, while their measurement can be performed through qualitative methods of research.

5.2. Limitations

At this point, we should mentioned for one last time that residuals of the estimated models are not normally distributed.

Another limitation is that the sample used in this thesis was not adequate in terms of length, not giving us the chance to study business cycle effects, that is to say long run relationships.

REFERENCES

Bellman, S., Potter, R., Treleaven - Hassard, S., Robinson, J. and Varan, D. (2011). The Effectiveness of Branded Mobile Phone Apps. *Journal of Interactive Marketing*, 25(4), 191-200.

Bolton, R. (1998). A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction. *Marketing Science*, 17(1), 45-65.

Boulding, W., Kalra, A., Staelin, R., & Zeithaml, V. (1993). A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions. *Journal of Marketing Research*, 30(1), 7-27.

Bowden, J. (2009). The Process of Customer Engagement: A Conceptual Framework. *Journal of Marketing Theory and Practice*, 17(1), 63-74.

Breusch, T. (1978). Testing for Autocorrelation in Dynamic Linear Models. *Australian Economic Papers*, 17(31), 334-355.

Bronner, F., & Neijens, P. (2006). Audience Experiences of Media Context and Embedded Advertising: A Comparison of Eight Media. *International Journal of Market Research*, 48(1), 81-100.

Burby, J., & Atchison, S. (2007). *Actionable Web Analytics: Using Data to Make Smart Business Decisions*, 1st ed. Indianapolis, IN: Sybex.

Calder, B., Malthouse, E. and Schaedel, U. (2009). An Experimental Study of the Relationship between Online Engagement and Advertising Effectiveness. *Journal of Interactive Marketing*, 23(4), 321-331.

Camm, J., Cochran, J., Fry, M., Ohlmann, J., Anderson, D., Sweeney, D., & Williams, T. (2015*). Essentials of business analytics*, 1st ed. Stamford, CT: Cengage Learning.

Cho, C., & Cheon, H. (2004). Why Do People Avoid Advertising on the Internet? *Journal of Advertising*, 33(4), 89-97.

Degiannakis, S., & Xekalaki, E. (2004). Autoregressive Conditional Heteroscedasticity (ARCH) Models: A Review. *Quality Technology & Quantitative Management*, 1(2), 271-324.

Engel, J., Blackwell, R., & Miniard, P. (1995). *Consumer Behavior*, 8th ed. Fort Worth, TX: Dryden Press.

Engle, R. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica*, 50, 987–1008.

Fagerstrom, A., & Ghinea, G. (2010). Web 2.0's Marketing Impact on Low-Involvement Consumers. *Journal of Interactive Advertising*, 10(2), 67-71.

Fazio, R., Zanna, M., & Cooper, J. (1978). Direct Experience and Attitude-Behavior Consistency: An Information Processing Analysis. *Personality and Social Psychology Bulletin*, 4(1), 48-51.

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research.* Reading, Massachusetts: Addison-Wesley, 222.

Fornell, C. (1992). A National Customer Satisfaction Barometer: The Swedish Experience. *Journal of Marketing*, 56(1), 6-21.

Fornell, C., Johnson, M., Anderson, E., Cha, J., & Bryant, B. (1996). The American Customer Satisfaction Index: Nature, Purpose, and Findings. *Journal of Marketing*, 60(4), 7-18.

Godfrey, L. (1978). Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables. *Econometrica*, 46(6), 1293-1302.

Green, W. (2002). *Econometric Analysis,* 5th ed. Upper Saddle River, NJ:Prentice Hall, 51.

Gupta, S. (2013). For mobile devices, think apps, not ads. *Harvard Business Review*, 91(March), 71-75.

Gustafsson, A., Johnson, M., & Roos, I. (2005). The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention. *Journal of Marketing*, 69(4), 210-218.

Jarque, C., & Bera, A. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review / Revue Internationale de Statistique*, 55(2), 163-172.

Khalid, H., Shihab, E., Nagappan, M., & Hassan, A. (2015). What Do Mobile App Users Complain About? *IEEE Software*, 32(3), 70-77.

Kim, E., Lin, J., & Sung, Y. (2013). To App or Not to App: Engaging Consumers via Branded Mobile Apps. *Journal of Interactive Advertising*, 13(1), 53-65.

Lavidge, R., & Steiner, G. (1961). A Model for Predictive Measurements of Advertising Effectiveness. *Journal of Marketing*, 25(October), 59–62.

Massaro, D. (1988). Review of the book Communication and Persuasion: Central and Peripheral Routes to Attitude Change, ed. by Petty, R., & Cacioppo J. *The American Journal of Psychology*, 101(1), 155-156.

Mitchell, A., & Olson, J. (1981). Are Product Attribute Beliefs the Only Mediator of Advertising Effects on Brand Attitude? *Journal of Marketing Research*, 18(3), 318-332.

Mittal, B., & Lee, M. (1989). A causal model of consumer involvement. *Journal of Economic Psychology*, 10(3), 363-389.

Percy, L., & Rossiter, J. (1992). A model of brand awareness and brand attitude advertising strategies. *Psychology and Marketing*, 9(4), 263-274.

Petty, R., & Cacioppo, J. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York, NY: Springer-Verlag, 262.

Petty, R., Cacioppo, J., & Schumann, D. (1983). Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement. *Journal of Consumer Research*, 10(2), 135-146.

Robehmed, N. (2013). What Is a Startup? *Forbes*. Retrieved January 07, 2017. Available at: http://www.forbes.com/sites/natalierobehmed/2013/12/16/what-is-a-startup/#58c15bfb4c63

Saks, A. (2006). Antecedents and Consequences of Employee Engagement. *Journal of Managerial Psychology*, 21(7), 600-619.

Seiders, K., Voss, G., Grewal, D., & Godfrey, A. (2005). Do Satisfied Customers Buy More? Examining Moderating Influences in a Retailing Context. *Journal of Marketing*, 69(4), 26-43.

Smith, R., & Swinyard, W. (1982). Information Response Models: An Integrated Approach. *Journal of Marketing*, 46(1), 81–93.

Stahl, F., Heitmann, M., Lehmann, D., & Neslin, S. (2012). The Impact of Brand Equity on Customer Acquisition, Retention, and Profit Margin. *Journal of Marketing*, 76(4), 44-63.

Wang, A. (2006). Advertising Engagement: A Driver of Message Involvement on Message Effects. *Journal of Advertising Research*, 46(4), 355-368.

APPENDICES

Appendix 1. Sample of Data

Day	Daily Users	
01/09/2015	14	Before the campaign started.
02/09/2015	20	
03/09/2015	20	
04/09/2015	10	
05/09/2015	30	
06/09/2015	12	
07/09/2015	17	
08/09/2015	13	
09/09/2015	12	
10/09/2015	14	
11/09/2015	13	
12/09/2015	9	
13/09/2015	5	
14/09/2015	25	
15/09/2015	19	
16/09/2015	53	
17/09/2015	20	
18/09/2015	35	
19/09/2015	46	
20/09/2015	29	
21/09/2015	47	
22/09/2015	98	
23/09/2015	327	The campaign is effective.
24/09/2015	419	
25/09/2015	495	
26/09/2015	539	
27/09/2015	587	
28/09/2015	539	
29/09/2015	502	
30/09/2015	493	
01/10/2015	409	
02/10/2015	421	
03/10/2015	442	
04/10/2015	460	
05/10/2015	361	
06/10/2015	426	
07/10/2015	378	
08/10/2015	784	

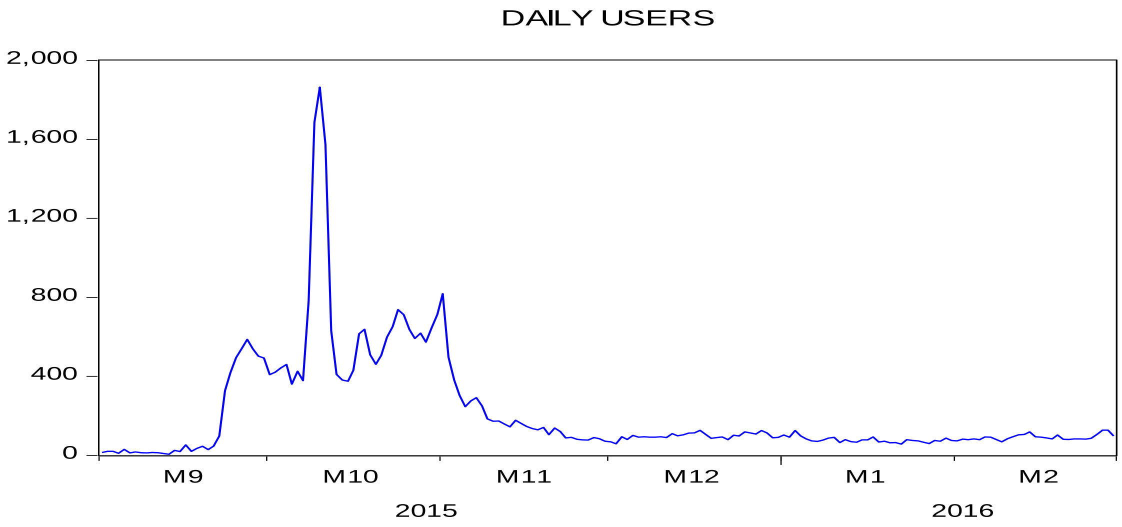
09/10/2015	1686	
10/10/2015	1865	
11/10/2015	1573	
12/10/2015	631	
13/10/2015	410	
14/10/2015	381	
15/10/2015	375	
16/10/2015	431	
17/10/2015	616	
18/10/2015	638	
19/10/2015	509	
20/10/2015	461	
21/10/2015	506	
22/10/2015	598	
23/10/2015	652	
24/10/2015	738	
25/10/2015	712	
26/10/2015	637	
27/10/2015	592	
28/10/2015	619	
29/10/2015	573	
30/10/2015	646	
31/10/2015	714	
01/11/2015	818	
02/11/2015	498	After the campaign ended.
03/11/2015	382	
04/11/2015	304	
05/11/2015	247	
06/11/2015	276	
07/11/2015	292	
08/11/2015	250	
09/11/2015	185	
10/11/2015	173	
11/11/2015	174	
12/11/2015	159	
13/11/2015	144	
14/11/2015	178	
15/11/2015	162	
16/11/2015	146	
17/11/2015	135	
18/11/2015	129	
19/11/2015	141	
20/11/2015	105	
21/11/2015	138	

22/11/2015	121	
23/11/2015	88	
24/11/2015	91	
25/11/2015	81	
26/11/2015	78	
27/11/2015	77	
28/11/2015	90	
29/11/2015	84	
30/11/2015	71	
01/12/2015	68	
02/12/2015	59	
03/12/2015	94	
04/12/2015	80	
05/12/2015	101	
06/12/2015	92	
07/12/2015	94	
08/12/2015	92	
09/12/2015	92	
10/12/2015	94	
11/12/2015	90	
12/12/2015	110	
13/12/2015	99	
14/12/2015	104	
15/12/2015	113	
16/12/2015	114	
17/12/2015	126	
18/12/2015	107	
19/12/2015	86	
20/12/2015	90	
21/12/2015	93	
22/12/2015	79	
23/12/2015	102	
24/12/2015	98	
25/12/2015	119	
26/12/2015	114	
27/12/2015	108	
28/12/2015	125	
29/12/2015	114	
30/12/2015	89	
31/12/2015	91	
01/01/2016	103	
02/01/2016	92	
03/01/2016	125	
04/01/2016	99	

05/01/2016	83	
06/01/2016	73	
07/01/2016	70	
08/01/2016	77	
09/01/2016	87	
10/01/2016	91	
11/01/2016	64	
12/01/2016	79	
13/01/2016	69	
14/01/2016	66	
15/01/2016	78	
16/01/2016	78	
17/01/2016	93	
18/01/2016	67	
19/01/2016	71	
20/01/2016	63	
21/01/2016	64	
22/01/2016	57	
23/01/2016	79	
24/01/2016	75	
25/01/2016	73	
26/01/2016	66	
27/01/2016	60	
28/01/2016	75	
29/01/2016	71	
30/01/2016	87	
31/01/2016	75	
01/02/2016	73	
02/02/2016	82	
03/02/2016	79	
04/02/2016	83	
05/02/2016	79	
06/02/2016	93	
07/02/2016	92	
08/02/2016	80	
09/02/2016	68	
10/02/2016	84	
11/02/2016	94	
12/02/2016	104	
13/02/2016	106	
14/02/2016	119	
15/02/2016	94	
16/02/2016	92	
17/02/2016	88	

18/02/2016	83	
19/02/2016	103	
20/02/2016	81	
21/02/2016	80	
22/02/2016	83	
23/02/2016	83	
24/02/2016	82	
25/02/2016	86	
26/02/2016	106	
27/02/2016	127	
28/02/2016	127	
29/02/2016	98	

Appendix 2. Time Series Plot of Daily Users (Raw Data)



Appendix 3. Time Series Plot of Log Daily Users (Dependent Variable)