Article

Exploring the Effect of In-Game Purchases on Mobile Game Use with Smartphone Trace Data

Kristof Boghe [1,*], Laura Herrewijn [2], Frederik De Grove [1], Kyle Van Gaeveren [1] and Lieven De Marez [1]

[1] imec-mict-UGent, Ghent University, 9000 Ghent, Belgium; E-Mails: kboghe@gmail.com (K.B.),
frederik.degrove@gmail.com (F.D.G), kyle.vangaeveren@ugent.be (K.V.G.), lieven.demarez@ugent.be (L.D.M.)
[2] Center for Persuasive Communication, Ghent University, 9000 Ghent, Belgium; E-Mail: laura.herrewijn@ugent.be

* Corresponding author

Abstract

Microtransactions have become an integral part of the digital game industry. This has spurred researchers to explore the effects of this monetization strategy on players' game enjoyment and intention to continue using the game. Hitherto, these relationships were exclusively investigated using cross-sectional survey designs. However, self-report measures tend to be only mildly correlated with actual media consumption. Moreover, cross-sectional designs do not allow for a detailed investigation into the temporal dimension of these associations. To address these issues, the current study leverages smartphone trace data to explore the longitudinal effect of in-game purchase behavior on continual mobile game use. In total, approximately 100,000 hours of mobile game activity among 6,340 subjects were analyzed. A Cox regression with time-dependent covariates was performed to examine whether performing in-game purchases affects the risk of players removing the game app from their repertoire. Results show that making an in-game purchase decreases this risk initially, prolonging the survival time of the mobile gaming app. However, this effect significantly changes over time. After the first three weeks, a reversal effect is found where previous in-game purchase behavior negatively affects the further survival of the game. Thus, mobile games without previous monetary investment are more prone to long-term continual game use if they survive the first initial weeks. Methodological and theoretical implications are discussed. As such, the current study adds to those studies that use computational methods within a traditional inferential framework to aid theory-driven inquiries.

1. Introduction

The mobile gaming market is expected to reach $77,2 billion in 2020, accounting for 48% of the global gaming revenue and seeing a year-on-year growth rate of 13.3% (Wijman, 2020). A contributing factor in this success is the increasing reliance on monetization through the use of microtransactions. Microtransactions refer to in-game purchases of additional downloadable content (DLC; e.g., adding new game modes) and/or virtual goods that can enhance the player's experience in the game (e.g., items that give players a winning edge in the game; Alha, Koskinen, Paavilainen, Hamari, & Kinnunen, 2014; Luton, 2013). The monetization strategy is being lucratively applied across game genres and platforms, and within both free-to-play games (e.g., Pokémon Go) and games that also involve an initial purchase price (e.g., Minecraft).

As a consequence, academic research has started to turn its attention to the study of in-game purchase be-

havior in recent years. In particular, research has started to investigate in-game purchase behavior in free-to-play games (both on mobile and on other platforms) by examining attitudes towards the microtransaction model (e.g., Alha et al., 2014; Hamari, 2015), which players decide to make in-game purchases and why (e.g., Balakrishnan & Griffiths, 2018; Hamari, 2015; Hamari, Alha, Järvelä, Kivikangas, & Koivisto, 2017; Hamari & Keronen, 2017; Hsiao & Chen, 2016), and what the relationship is between game enjoyment, the making of in-game purchases, and the intention to continue playing the game (e.g., Hamari, 2015; Hamari & Keronen, 2017; Hsiao & Chen, 2016).

Consistently, these studies have underlined the fundamental impact that the microtransaction model has on the game design philosophy. Specifically, when game developers decide to implement the use of microtransactions, they have to find ways to justify and create value for the virtual goods that they offer in order to motivate players to make in-game purchases as frequently as possible (Hamari, 2015; Hamari et al., 2017; Hamari & Keronen, 2017; Hamari & Lehdonvirta, 2010). Game developers themselves indicate that this balancing act has been a matter of great difficulty, since players have no reason to pay money for virtual goods if they are already having a great time with the core game (Alha et al., 2014). As such, artificial barriers are often integrated into the gameplay that make the game mechanics bothersome (e.g., by limiting the amount of lives, resources, or time that players have to play the game every day), after which virtual goods are offered that can break them down (Hamari, 2015; Hamari et al., 2017; Hamari & Keronen, 2017; Hamari & Lehdonvirta, 2010).

This suggests that in-game purchase decisions of players are no longer only influenced by their existing general attitudes and consumption motivations, but also by the developers' design decisions; and that in-game purchase behavior may in turn affect whether a person will continue playing a game (Hamari, 2015; Hamari et al., 2017; Hamari & Keronen, 2017; Hamari & Lehdonvirta, 2010; Hsiao & Chen, 2016). Indeed, the findings of prior research have shown that in-game purchase decisions are highly motivated by the appeal of 'unobstructed play' (i.e., circumventing the barriers that make the gameplay inconvenient), in addition to purchases being motivated by factors such as 'social interaction' (e.g., gift giving), 'economic rationale' (e.g., capitalizing on good deals) and 'unlocking content' (i.e., DLC; Hamari et al., 2017; Hsiao & Chen, 2016). Moreover, studies have found that there is a negative association between enjoyment from playing the game and in-game purchase intention (Hamari, 2015; Hamari & Keronen, 2017), and at the same time, a positive association between in-game purchase intention and continued playing intention (Hamari, 2015; Hamari & Keronen, 2017; Hsiao & Chen, 2016). Both results seem to imply that obstructing players' gameplay (generating frustration) may result in increased in-game purchasing (alleviating the barriers for enjoyment), which

may in turn lead to continual game use, at least in the short term.

It is important to note, though, that prior research has measured the association between in-game purchases and players' intention for continual game use exclusively with cross-sectional survey data. Self-report measures of media use, however, are notoriously poor proxies for actual consumption, showing only moderate to low correlations with measures obtained from log data or experience sampling (Araujo, Wonneberger, Neijens, & de Vreese, 2017; Boase & Ling, 2013; Ellis, Davidson, Shaw, & Geyer, 2019; Scharkow, 2016). Biases in retrospective measurements might be especially prevalent for heavily fragmented and short media consumption patterns such as playing mobile games or making in-game purchases (Naab, Karnowski, & Schlütz, 2018). The current study aims to remedy this limitation by analyzing smartphone trace data to shed light on players' actual purchase and gaming behavior. Thus, the following research question is posed:

RQ1: What is the relationship between performing in-game purchases and continual mobile game use?

In addition to increasing the validity of behavioral measures as such, leveraging log data also enables the modeler to discern granular and temporal patterns which are indistinguishable using cross-sectional survey data. More specifically, this approach makes it possible to establish whether the association between in-game purchases and continual game use might change over time. This seems especially relevant, as results from prior research making use of interview and survey research additionally suggest that implementing microtransactions in games might constitute a double-edged sword in the long term. Notwithstanding the initial proposed positive effects, players seem to argue that having to buy in-game goods with real money in order to be able to continue playing the game the way they want to weakens the game experience in the long run (Alha et al., 2014; Hamari, 2015; Hamari et al., 2017). Furthermore, being able to buy virtual goods that give the owner certain advantages in the game is also believed to skew the competition with other players, potentially resulting in unbalanced gameplay and some games getting called 'pay-to-win' (Alha et al., 2014; Hamari, 2015; Hamari et al., 2017; Hamari & Lehdonvirta, 2010). These previous studies therefore imply that these negative experiences might lead to the formation of negative player attitudes towards the microtransaction model over time (Alha et al., 2014; Hamari, 2015; Hamari et al., 2017). Since attitudes are positively associated with continued playing intention (Hamari, 2015; Hamari & Keronen, 2017), this may eventually result in an increased risk of abandoning the game. No research has actually investigated the existence of such a negative effect over time, however. Therefore, the following research question is posed:

RQ2: How does the relationship between performing mobile in-game purchases and continual game use change over time?

Notably, we frame this research as an exploratory inquiry, serving as a potential starting point for future researchers in sketching out the psychological mechanisms that could explain our findings. In this sense, our data-driven angle is an exemplar of what Margolin (2019) calls the symbiotic approach between computationally-intensive observational studies and more traditional methods within communication sciences. Within this framework, computational approaches can serve as a fertile ground for generating new hypotheses based on the observation of behavioral patterns. This is one of the key advantages of what Margolin (2019) calls "the computational niche." Thus far, methodological limitations have steered the field into inquiries that neglect the temporal nature of mobile gaming, i.e., the fact that purchase behavior should be situated within a game's total 'life span' and as part of an individual gamer's repertoire. It is exactly this temporal dynamic that could yield new valuable hypotheses. In this sense, we hope to stimulate other researchers to disentangle said causal mechanisms within carefully-controlled research settings (e.g., experiments).

2. Methods

We had access to a database containing log data from a free-to-install Android application (Mobile DNA) that gives consumers insight into their own smartphone behavior. The software and the data it generates is proprietary of our institution and is exclusively designed to serve academic research. Users agreed to potential use of their data for this purpose by the authors' research unit before opting in. No other organization or individual outside the research group has access to the database. A subset of our subjects was recruited by our organization, although most users installed the app on their own initiative thanks to a considerable amount of media exposure (e.g., the app has been featured on popular current affairs programs on national television). Subjects received no incentive for installing the application and were free to uninstall it at any time. For our purposes, we extracted all application-data from 01/01/2018 until 02/09/2019. These logs contain information on the specific app used by the subject, including start and end timestamps (precision: 1 millisecond). Moreover, all logs include an anonymized subject-key. Although the application collects geospatial data, these variables were not requested from the database. For this reason, the extracted data contain no identifiable information on our subjects whatsoever.

2.1. Sample

The database holds information on 14,426 subjects, totaling 202 million logs. Of special interest here is that the archive contains 287,789 hours (4,4 million logs) of mobile game use. 9,039 subjects opened at least one game during the data collection period and were included in our initial sample. To identify faulty data logs caused by an early software bug, we identified subjects that supposedly had spent more than 24 hours in a single day on their smartphone. As a result, six individuals were removed from the sample. Next, we accounted for non-human activity (e.g., test devices, bots) in our dataset by inspecting ids that appear in the 99th percentile on both of the following variables: median time spent on smartphone in a single day (median = 2.35 hours, 99th percentile = 7.38 hours) and median duration of a single smartphone session (i.e., opening and closing of smartphone; median = 53 seconds, 99th percentile = 10.21 minutes). Although a subject could legitimately obtain extreme scores on these indicators when the pool of observations is relatively small (e.g., the subject only logged his activity for two days), we would expect a reversion to the median when the number of days under study increases. For this reason, to be eligible for deletion subjects had to be included in the dataset for at least seven days. We deleted eight potential non-human subjects due to this procedure. The density functions of the abovementioned criteria can be found in the Supplementary File (see Figure A1 in the Supplementary File).

2.2. Defining Relevant Survival Periods

Crucially, from a conceptual viewpoint, we are not primarily interested in the total time spent on a mobile game. Instead, the so-called survival time is of key interest. The goal here is to make abstraction of the intensity of game use and rather capture consistent gaming behavior or how long an app remains in one's repertoire. However, unlike available approximations of app survival time in the literature (e.g., Jung, Baek, & Lee, 2012), we argue that the time interval between the first and last day of app usage is a poor approximation of how long an app actually remained in a subject's repertoire. In many cases, only a subset of the total time interval is relevant for measuring actual user interaction due to the many and long pauses in individual gaming behavior. To illustrate this, consider the fact that in our 2018 sample 71% of all days between the first and last active date of a specific game are dates without any gaming activity. At the same time though, it is unreasonable to define continued game usage as a completely non-interrupted streak of gaming activity. Thus, the challenge here is to define what could be considered a maximal allowed tolerance (in days) or time gap between play days.

For this purpose, we formulate an elementary gain metric. The gain metric aims to balance two co-occurring tendencies when one increases this so-called tolerance: Although the amount of captured play days invariably increases with more liberal windows, the interval will include—proportionally—more and more non-play days

as well. The metric aims to establish at which tolerance-level the increase in non-play days far outweighs the gain in captured play days. A specific, though fictitious, example of how different tolerant levels impact the survival time of an app can be found in the Supplementary File (see Table A1 in the Supplementary File).

The gain metric employs the survival time defined by tolerance$_{(1)}$ as the baseline, which is the most strict tolerance level (i.e., no pauses are allowed) and compares it with the play days/total days ratio captured by applying a more liberal tolerance level. It is thus defined as follows:

$$\text{Gain}_{tx} = \frac{\text{play days}_{tx} - \text{play days}_{t1}}{\text{total days}_{tx} - \text{total days}_{t1}}$$

Here, $t(x)$ represents the results obtained by applying tolerance level x, while $t(1)$ represents the number of (play) days captured by the baseline tolerance$_{t(1)}$. We calculated the gain for each game played by a specific individual when employing a tolerance level between two and 14 days. Only tolerance-levels which succeeded in increasing the total interval were retained (as indicated by the crossed-out tolerances in Table A1 in the Supplementary File). We subsequently calculated the mean gain (and standard error) for each tolerance-level.

Similar to how eigenvalues are used for determining the optimal number of components in principal component analysis (Schönrock-Adema, Heijne-Penninga, Van Hell, & Cohen-Schotanus, 2009), the preferred solution is determined by taking the tolerance-level just before the point of inflection, which seems to be seven days (see Figure 1). In other words, mobile gamers are allowed to take a break for six consecutive days from playing a specific mobile game. If the game remains untouched on day seven, the last play day before this seven (+) days gap is determined as the end date of the gamer's continual app usage. This procedure captured 76% of all play days while reducing the amount of dates without gaming ac-

tivity with 90%. In total, 228,035 hours (around 80% of all gaming activity) is retained after applying the maximally allowed tolerance level.

2.3. Left Truncation

Apps were left truncated and thus removed for analysis if the mobile game appears during the first seven days of an individual's data collection. This minimizes the risk of including apps that were already part of one's mobile game repertoire before the subject actually started logging. After all, our gain metric reveals that gamers tend to remove a game from their repertoire if they pause their game activity for more than six days. This procedure diminished the amount of gaming hours in the final dataset to around 96,000 hours, a reduction of 58%.

2.4. Extracting In-Game Purchases

We defined in-game purchase processes by employing an algorithm that looks for specific sequences within an individual's logs set. Users who instantaneously (< one second) switched from playing a mobile game to visiting the Google Play Store (for at least four seconds) and subsequently switched back (< one second) to playing the very same mobile game were considered to have purchased an item in-game. This specific log-chain proved to be most predictive for actual Google Play vending processes after investigating log data of in-game purchases performed by the researchers. We aimed for a conservative estimate to limit the inclusion of accidental or other non-purchase related switches to the Play Store, such as clicking on an advertisement by accident. For this reason, we incorporated two additional decision rules within our purchase detection algorithm. First, all first switches to the Play Store (for each game) are ignored by default. Similarly, switches to the Play Store within the initial five
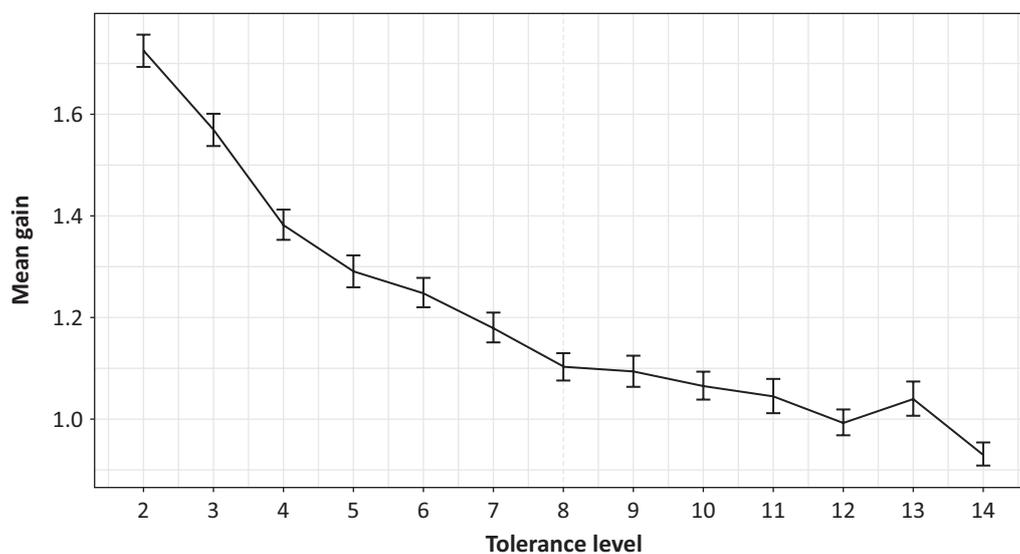

Figure 1. Relation tolerance level and mean gain. Note: Error bars represent standard errors (+/−).

minutes during the first gaming session are disregarded. This allows the user to fall prey to predatory advertising techniques once per game before a vending-chain is categorized as an in-game purchase.

2.5. Scraping Application Data

Since our database as such only includes the application name without any additional info, we developed our own web scraper to obtain relevant metadata. This allowed us to differentiate between mobile games and other apps. In total, we crawled five online app repositories using a sequential scraping method. In order, we scraped the following repositories: Google Play Store, APKMonk, APK Support, APKsHub and APK Pure. If an app was unavailable in the Google Play Store, we opted for the second most reliable online repository, and so on. Next to the general app category (i.e., 'mobile game,' 'social app') and the availability of in-game purchases, we collected other relevant variables which serve as covariates in our model. For more information on the web scraper we refer to Boghe, De Grove, Herrewijn, and De Marez (2020). Since all of these scraped variables are covariates within our model, we removed apps which were unavailable in the Google Play repository from our final dataset to avoid missing values. This led to an exclusion of 3,746 games (7% of all games) from our analysis. However, to check whether this exclusion had an impact on our findings, we ran the model again for all mobile games without the aforementioned covariates.

2.6. Measures

The following descriptive statistics count each and every unique id/application combination as a single mobile game unless mentioned otherwise. Control variables, however, are aggregated on game or id-level, since they do not vary within the same user or app level (e.g., game rating, median time spent gaming for each subject).

2.6.1. Purchases

In total, 46,184 mobile games were included in the sample among 6,340 mobile gamers. 1,082 mobile games included in-game purchases, with a total of 3,884 purchases. Evidently, some games were shared among multiple subjects. When aggregated on game-level, 7,901 unique mobile games were included in the dataset, of which 527 included (a) purchase(s) of at least one subject.

2.6.2. Control Variables

2.6.2.1. Rating and Number of Downloads

Both the average rating of the app on the Google Play Store (1–5) and the number of downloads (ordinal scale from 100–1 billion) were incorporated as proxies of mo-

bile game quality. The assumption here is that highly entertaining mobile games tend to dominate the (free) market. Indeed, popular games tend to receive higher ratings ($r_s = 0.18$, $p < 0.001$). Unsurprisingly, the apps in our dataset tend to be relatively popular (median: one million downloads) and highly rated (median: 4.20, min: 1.30, max: 5.00). For modeling purposes, we aggregated the variable into three categories using the 1st (one million downloads) and 3rd (five million downloads) quartile as cut-offs.

2.6.2.2. Free-to-Play Versus Paid Apps

The differentiation between free-to-play and paid apps is of key interest given our focus on in-game purchases. If in-game purchases have a determining impact on the survival time of a mobile game, this covariate might serve as an important confounding factor. After all, some mobile gamers already made a monetary investment before installing the app in the first place. In our dataset, 6% of all mobile games were paid apps.

2.6.2.3. Availability of Multiplayer Component

Previous research has uncovered that social play is an important motivational factor for continual mobile game use (e.g., Hsiao & Chiou, 2012; Teng & Chen, 2014). In total, 14% of all mobile games in our dataset contain a multiplayer component.

2.6.2.4. Median Game Session Duration

We calculated the median duration of a single game session (in minutes) for each unique application in our dataset. Sessions were defined as the opening and closing of one's smartphone. This metric serves as a proxy of the time investment needed for continual game use. It is not unreasonable to assume that more time-intensive games might exhibit different survival patterns than apps which are more easily appropriated for short game sessions. The median time spent on a mobile game during a single session is 4.10 minutes (min: 0.01, max: 622.67). Three out of four games have a median game session duration of less than 7.60 minutes.

2.6.2.5. Median Time Spent Gaming

Finally, the median time (in minutes) spent gaming in a single day was added for each individual to the model. Avid gamers might exhibit different consumption patterns than sporadic gamers and therefore might show different survival curves. Only days with at least one game log for a specific individual were considered. The distribution is heavily right skewed, with a median time of 16 minutes spent on games in a single day (min: 0.01, max: 625.00). Around one out of four gamers tend to play more than 34 minutes in a single day.

2.6.2.6. Game Repertoire

The survival of an app might be dependent on the current game repertoire of an individual. Specifically, removing an app might be a reasonable strategy to reduce a cluttered game repertoire. We measured the amount of gaming apps opened by a specific individual in a specific day (min: 1, median: 2, max: 32). For days without any gaming activity (12% of all days), we imputed the most recent game repertoire for each individual. On most days (23%), gamers played a single mobile game.

The density functions (for interval variables) or bar plots (for categorical variables) of the abovementioned covariates can be found in the Supplementary File (see Figures A1 and A2 in the Supplementary File).

2.7. Analysis

2.7.1. Time-Dependent Cox Regression

To determine whether in-game purchases have an impact on continual gaming behavior, we performed a Cox regression with time-dependent covariates (Therneau, Crowson, & Atkinson, 2017). The Cox model, unlike traditional regression models, is a survival model which accommodates to censoring techniques necessary for most time-to-event data (Prinja, Gupta, & Verma, 2010; Therneau & Grambsch, 2000). In essence, a survival model estimates the association between covariate scores and the average hazard or failure rate, which constitutes the event rate (e.g., removing an app from the repertoire) at time t given that the subject (i.e., mobile game) survived up until time t. Censored apps only have an impact on the hazard function until their censored survival time. After that, they are removed from the risk set entirely and therefore have no impact on subsequent hazard estimates. For our purposes, apps were censored (i.e., coded as having survived the period under study) if the last activity log lies within less than seven days (the maximal tolerance-level) from the last log of the subject.

Moreover, the longitudinal variant of the Cox model employed here estimates the average hazard (also called risk score) while correcting for the fact that values of (some) predictors might be correlated with survival time (Suissa, 2008). In our case, the probability of having made an in-game purchase rises steadily as survival time increases. The probability of having made an in-game purchase after ten days of playing a mobile game is a mere 10%, which rises to 26% after 20 days. Since in a cross-sectional survival analysis these purchases are observed only at the end of said life span, any observed relationship between in-game purchases and survival outcomes might be an effect of the variable 'time under study' as such. To avoid this, a key advantage of the longitudinal model is that it compares covariate-scores of apps with a survival time of t with the covariate-scores of other apps up until time t. In other words, only current values have an impact on the estimated hazard and the model is—on purpose—ignorant of any future state of the subject.

To account for correlated events within-subject, we used the cluster variance by id to estimate a robust standard error (Therneau et al., 2017). Since more than half (51%) of all video games were not shared among more than three subjects, we did not account for the shared variance in survival times within the same game-cluster. However, we reran our analyses while excluding games shared among more than three subjects to check for the potential influence of specific survival times on game-level (see Section 3.3).

2.7.2. Proposed Model

We ran four blockwise Cox regressions with time-dependent covariates. The first model aims to establish the relationship between having performed at least one in-game purchase up until time t, before adding other relevant app characteristics (Model 2), proxies of app quality (Model 3) and player characteristics (Model 4). Importantly, the purchase variable could take on three different values depending on the specific game played and the actions undertaken by the user up until time t. More specifically, the user could play a game where: a) microtransactions were simply unavailable; b) microtransactions were available, but no in-game purchases were made up until time t; or c) microtransactions were available, and the user has made at least one in-game purchase in the past. The time-dependent covariates in our model allowed the user to switch from category b) (purchase available but not yet performed) to c) (purchase performed) on any specific day. We entered category b) as the reference category in the model. Moreover, we entered the median game session duration (Model 2), app rating (Model 3), and both covariates on id-level (Model 4) on a logarithmic scale due to its heavily skewed distribution.

Data cleaning and the calculation of general summary statistics for each user was done in Python using the Dask library as it supports parallel processing. The actual modelling was performed in R using the survival package (Therneau, 2020).

3. Results

3.1. Effect of Purchases on Survival Time

Performing an in-game purchase significantly decreases the risk estimate for app deletion (hazard ratio [HR]: 0.72; 95% confidence interval [CI]: 0.68–0.77). More specifically, apps where in-game purchases are available but none have been made up until time t experience a 39% increase in risk of app removal. Interestingly, games without any available in-game purchases experience a 1.22 fold increase in risk (HR: 1.22; 95% CI: 1.17–1.26) when compared to games where in-game purchases are available but not yet performed. Thus, the mere availability

of in-game purchases decreases the likelihood of app removal on time t, even without having actually performed said purchase(s). However, actually performing at least one in-game purchase further decreases the risk. Notably, these relationships remain relatively robust regardless of the control variables added to the model.

When it comes to app characteristics, there is no significant association between playing a pay-to-play game and the estimated hazard (HR: 1.00; 95% CI: 0.93–1.08). Contrary to this, playing a game with a multiplayer component does decrease the risk score slightly but significantly (HR: 0.96; 95% CI: 0.93–0.98). The relationship between the median duration of a single play session and app survival is less clear-cut. Although time-intensive games initially seem to have a lower risk for app removal, the final model suggests a small but significant increase in the hazard ratio as the median time spent per session increases (HR: 1.04, 95% CI: 1.02–1.05). Both proxies for app quality—rating and the amount of downloads— have a negative association with the risk estimate. First, for each log unit increase in game rating, the hazard decreases with 11% (HR: 0.89, 95% CI: 0.86–0.92). Second, games between 0.5 and ten million (HR: 0.93, 95% CI: 0.90–0.97) and more than ten million downloads (HR: 0.82, 95% CI: 0.79–0.86) experience a lower risk of app removal when compared with less popular (< 0.5 million downloads) games. Finally, our model suggests that gamers who tend to spend more time on mobile games in a single day retain apps for a longer time period (HR: 0.79, 95% CI: 0.78–0.81). At the same time, though, apps are more likely to be deleted as an individual's game repertoire on time t increases (HR: 1.70, 95% CI: 1.65–1.75). Detailed parameter estimates for all four models can be found in Table 1.

3.2. Time-Dependent Coefficients

Finally, the presence of time-dependent coefficients was explored by calculating the Spearman rank correlation between survival time and scaled Schoenfeld's residuals. This analysis shows that the parameter of 'in-game purchases performed' changes significantly over time ($\chi^2 = 27.30$, $p < 0.001$).

Therefore, we reran the fourth model containing a step function for $\beta(t)$ divided over multiple time periods (Zhang, Reinikainen, Adeleke, Pieterse, & Groothuis-Oudshoorn, 2018). The direction of the relationship seems to reverse after playing a game for three weeks (see Figure 2), which served as a cut-off point for our new model. The step function shows a clear reversal effect. While the risk for app removal decreases for games with an in-game purchase during the first 21 days (HR: 0.64; 95% CI: 0.59–0.70), in-game purchases are associated with a 38% increased risk after said time period (HR: 1.38; 95% CI: 1.19–1.59). Thus, while at first glance the model suggests that performing (an) in-game purchase(s) stimulates continual game use, it is exactly this type of previous monetary investment that increases the risk of app removal later in a game's life span.

3.3. Model Robustness

We validated our model by checking for influential observations, outliers, and non-normality. Based on a cross-sectional Cox model, we calculated DfBetas (DfBeta/SE) for each observation. Following recommendations by Belsley, Kuh, and Welsch (2004), values larger than two should raise our attention. There were no influential observations present in our data, with a maximal DfBeta

Table 1. Estimated parameters of Cox regression with time-dependent covariates.

Variable	Model 1 HR (95% CI)	Model 2 HR (95% CI)	Model 3 HR (95% CI)	Model 4 HR (95% CI)
purchases: [a]				
purchase unavailable	1.34 (1.30–1.38) ***	1.31 (1.27–1.36) ***	1.22 (1.18–1.26) ***	1.22 (1.17–1.26) ***
purchase performed	0.68 (0.64–0.73) ***	0.68 (0.63–0.73) ***	0.69 (0.64–0.73) ***	0.72 (0.68–0.77) ***
paid app		1.00 (0.93–1.08)	0.97 (0.90–1.05)	1.00 (0.93–1.08)
multiplayer		0.89 (0.87–0.92) ***	0.91 (0.89–0.94) ***	0.96 (0.93–0.98) **
median session game		0.98 (0.97–0.99) **	1.00 (0.98–1.01)	1.04 (1.02–1.05) ***
rating			0.83 (0.80–0.86) ***	0.89 (0.86–0.92) ***
downloads: [b]				
0.5–ten million			0.97 (0.94–1.01)	0.93 (0.90–0.97) **
+ ten million			0.87 (0.83–0.91) ***	0.82 (0.79–0.86) ***
time spent on games				0.79 (0.78–0.81) ***
game repertoire				1.70 (1.65–1.75) ***
Wald-score (p)	443 ($p < 0.001$)	518 ($p < 0.001$)	751 ($p < 0.001$)	1909 ($p < 0.001$)

Notes: [a] reference category ('purchases available, but not performed'); [b] reference category ('less than or equal to 0.5 million downloads'). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
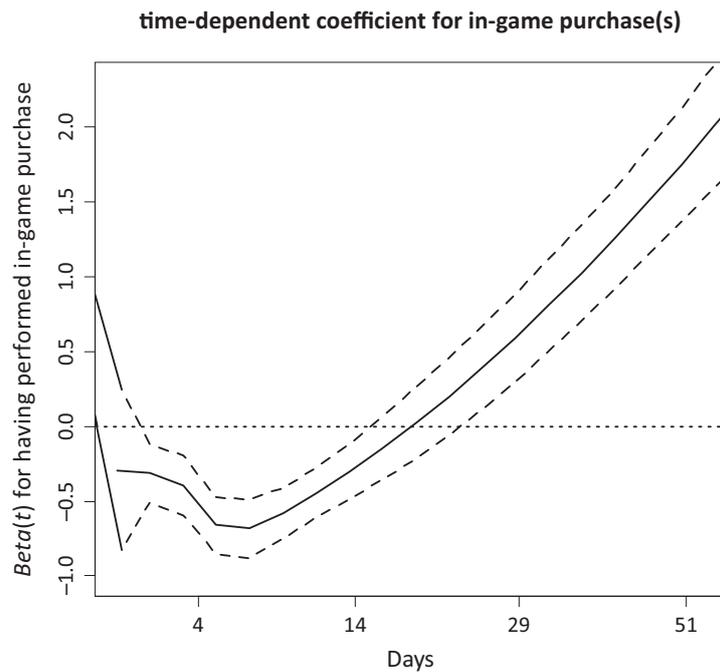
Figure 2. Estimated *Beta*(*t*) (*df* = 8) for having performed at least one in-game purchase up until time *t* over time (log[days]).

of 0.2 (see Figure A3 in the Supplementary File). Similarly, an inspection of the deviance residuals of the longitudinal Cox model show no clear deviations from normality nor any indication of outliers, with residuals centered around zero (see Figure A4 in the Supplementary File). To check whether specific mobile games shared by multiple users had an undue influence on our parameter estimates, we reran our analyses twice; once with a corrected standard error estimate by game-variance (instead of id-variance) and once by only including games that were shared by at most three subjects. Both alterations to our model had no considerable impact on the parameter estimates nor on the time-dependent effects previously reported (see Table A2 and Figure A4 in the Supplementary File). The time-dependent coefficient for the second model mentioned here shows greater uncertainty, which is to be expected given the smaller sample size. Some app characteristics such as number of downloads are less predictive, which can be explained by the fact that we are pooling less popular games here. Next, we inspected for general model robustness by: (a) running the model with a less conservative estimate of the amount of in-game purchases by categorizing the first vending process as an in-game purchase; and (b) running the model while including mobile games which were unavailable on Google Play (excluding the game-level covariates). These model adjustments had no impact whatsoever on the trends already reported (see Table A2 and Figure A4 in the Supplementary File). As one could expect, the time-dependent coefficient for the less conservative estimate of purchase detection shows a less strong effect when compared with the more stringent definition, but the reversal effect is still clearly present.

4. Discussion and Conclusion

This study contributes to the field of media entertainment studies (and game studies more specifically) on multiple grounds. First, on a methodological level, this is to our knowledge the first study that applies log data to explore time-dependent relations between actual in-game purchase behavior and continual mobile game use. While there have been a couple of investigations into the impact of several app characteristics on survival time (e.g., Lee & Raghu, 2014), no study has specifically sought to establish a clear time-dependent relation between a particular behavioral antecedent and consequence. Moreover, we developed a detection algorithm to discern in-game purchases in log data. We urge future studies to examine the validity and robustness of this formulation of vending processes. Available validated algorithms might stimulate the field to sketch out the effect of in-game purchases in a more nuanced fashion, incorporating the longitudinal logic implied here. Finally, our proposed tolerance metric might prove to be a valuable tool for communication scholars in general. We believe that most—if not all—measures of mobile media consumption stemming from behavioral log data could profit from our tolerance-approach. Mobile media tend to be consumed in extremely short bursts and are prone to habitual activation (Oulasvirta, Rattenbury, Ma, & Raita, 2012), which makes the demarcation of detailed but relevant periods of media consumption a new challenge for the communication scientist.

Second and more fundamentally, on a theoretical level the study is the first to explore the relationship between in-game purchase behavior and continual mobile game use (RQ1), as well as how this relationship

changes over time (RQ2). Regarding our first research question, the results reveal a positive relationship between making in-game purchases and continual mobile game use. Specifically, compared to games where at least one in-game purchase is performed, games that allow microtransactions to be made but where players have not (yet) performed in-game purchases experience a 39% increase in risk of being removed from players' app repertoires. Moreover, and interestingly, games that do not make use of microtransactions and where players thus have no opportunity to make purchases see a 22% increase in risk when compared with games where purchases are available but not (yet) performed. In other words, the mere availability of microtransactions in mobile games decreases their risk of removal from players' app repertoires, and when players perform at least one in-game purchase, this risk of removal decreases even further.

There are two important notes to be made about the unveiled association. First, it is noteworthy that the mere availability of microtransactions in games decreases the risk of removal as well, when compared to games where no microtransactions are available. This might be explained by some underlying characteristic(s) shared among these games. For example, game developers who make use of this monetization strategy generally have access to a more steady stream of revenue and might therefore offer a higher initial production value, as well as updates over time (e.g., balancing their games regularly) and recurrent releases of new content (e.g., releasing new chapters of the story, character customization options), which might cause players to keep the game in their repertoire for a longer period of time. A second note pertains to the control variables included in the model. Interestingly, we found no significant association between a game being either free-to-play or pay-to-play and the risk estimates. This suggests that the behavioral effects of these two different revenue models on mobile gamers are distinct.

When it comes to our second research question, analyses show that the effect of in-game purchase behavior on continual mobile game use changes significantly over time. Specifically, a clear reversal effect is found: While the risk for removal decreases with 36% for mobile games with a performed in-game purchase during the first three weeks, it is associated with a 38% increase in risk after said time period. Results thus show that prior in-game purchase behavior has a negative impact on continual mobile game use later on in a game's life span. Mobile games in which microtransactions are available but where players have not made any in-game purchases, on the other hand, are more prone to long-term survival if they make it through the first few weeks.

At first glance, these results seem to be in line with the expectations posed previously: Performing in-game purchases might prolong the survival time of a mobile game initially (e.g., by taking away obstructions that have been artificially implemented in the gameplay, as well

as the player frustration that results from this), but after a while this effect may turn sour (e.g., because the game will keep introducing barriers and looking for monetary investments, weakening the game experience and resulting in players becoming discontented). Importantly, since our results consistently show that the estimated risk score of apps with actual performed in-game purchases differs significantly from the risk score of apps where purchases are available but not (yet) performed, the reversal effect cannot be sufficiently explained by the frustrating experiences and negative player attitudes that might potentially result from the built-in barriers in microtransaction games. In both instances, players may be confronted with barriers that invite them to make an in-game purchase, but it is only after actually performing said purchase(s) that a differential survival curve and the reversal effect take place.

Given the exploratory nature of our inquiry, the question remains how one should interpret these findings. We call for future research to disentangle the causal mechanism behind this reversal effect. Nonetheless, we wish to give several pointers here. A potential explanation for why the act of performing in-game purchases triggers these differential patterns in continual mobile game use can be found in Self-Determination Theory (SDT; Ryan & Deci, 2000). Motivation is crucial in supporting continual behavior (Teixeira, Carraça, Markland, Silva, & Ryan, 2012), and SDT concerns itself with factors that can facilitate or undermine motivation, both intrinsic and extrinsic (Ryan & Deci, 2000). Within a digital game context, intrinsic motivation, especially, has proven to be essential (Ryan, Rigby, & Przybylski, 2006). Although games can be played for external reasons, such as receiving monetary rewards (e.g., professional gamers), most players are intrinsically motivated to do so: They play games because it is intrinsically satisfying to them (Ryan et al., 2006), because they are seeking enjoyment.

A sub theory within SDT, called Cognitive Evaluation Theory (CET; Deci & Ryan, 2000), addresses events and conditions that can reinforce or impede this intrinsic motivation. Specifically, CET states that factors that enhance a person's psychological needs of autonomy (i.e., a person's need for volition or free will when performing a task) and competence (i.e., a person's need for challenge and feelings of effectiveness) can support intrinsic motivation, while factors that thwart these needs can enfeeble it (Deci & Ryan, 2000; Ryan et al., 2006). Importantly, gaming research has shown that the satisfaction of these needs of autonomy and competence in a game also predicts players' enjoyment (Neys, Jansz, & Tan, 2014). Moreover, both players' intrinsic motivation to play games (Ryan et al., 2006) and their enjoyment of a game (Neys et al., 2014) have been shown to be positive predictors of their future play intentions.

When taking a look at in-game purchase behavior and its relation to continual mobile game use, then, it seems plausible that controlling behavior by implementing artificial barriers could impair players' sense of au-

tonomy and competence. Making a purchase may then lift these barriers and lead to unconstrained or even accelerated gameplay for a moment, which may lead to a temporary burst of enjoyment. This might explain why in-game purchase behavior (versus deciding not to make an in-game purchase) is positively related to continual game use in the short term. However, the decision to purchase may also cause players to feel as if they are forced to perform in-game purchases to be able to keep on playing the game the way they want to (diminishing their feelings of autonomy), and that their progress in the game is not due to their own skills and abilities, but rather because of their monetary investments (weakening their feelings of competence).

Additionally, in-game purchase behavior may also cause a shift in player motivation from the inherently satisfying aspects of a game (i.e., intrinsic motivation: "I play the game because I find it enjoyable") to external, monetary aspects (i.e., extrinsic motivation: "I play the game because I have already invested money in it, and I have to get my money's worth"), which research shows is much shorter-lived (e.g., Teixeira et al., 2012).

As a consequence of these two processes, the act of making purchases in a game may chip away players' intrinsic motivation for continued play one purchase at a time, resulting in a negative association with enjoyment and continual mobile game use in the long run. Players who do not make in-game purchases, on the other hand, may feel the frustrating impact of the microtransaction design initially (curbing their autonomy and intrinsic motivation in the short term), but in the long haul, they may feel that their achievements are their own (resulting in a stronger sense of competence) and that choosing not to make monetary investments allows them to play the game according to their own will (enhancing their autonomy and keeping the focus on the inherently satisfying aspects of the game instead of on its monetary facets). As such, their intrinsic motivation and enjoyment will be more durable, resulting in increased long-term survival for the mobile game(s) they play.

Two constraints, more specifically in data availability and modeling capacity, are worth noting here as they can inform future researchers to expand on these findings. First, the reliance on log data as the sole data source impaired our model significantly. Given the reported behavioral relationships in this research, one promising avenue is to use our findings to develop a testable conceptual model which not only includes behavioral measures, but also sheds light on the psychological antecedents and consequences of the reversal effect reported here. This will allow the field to contextualize the reversal effect and ultimately to formulate some sensible recommendations to strengthen players' game literacy. The combination of trace and survey data will prove to be indispensable in this regard (Stier, Breuer, Siegers, & Thorson, 2019). Second, although we are confident our analysis strategy yields valid and reliable results, the data structure employed here is ideally modeled by techniques

that take into account the cross-classified nature of the data. Given the computational limitations of the frailty packages available in R, a Bayesian approach (using Stan, for example) might prove to be valuable to optimize the parameter estimates in future research.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Alha, K., Koskinen, E., Paavilainen, J., Hamari, J., & Kinnunen, J. (2014). Free-to-play games: Professionals' perspectives. In *Proceedings of the 22nd International Academic Mindtrek Conference* (pp. 49–58). New York, NY: Association for Computing Machinery. Retrieved from https://dl.acm.org/doi/10.1145/3275116.3275133

Araujo, T., Wonneberger, A., Neijens, P., & de Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of internet use. *Communication Methods and Measures*, *11*(3), 173–190.

Balakrishnan, J., & Griffiths, M. D. (2018). Loyalty towards online games, gaming addiction, and purchase intention towards online mobile in-game features. *Computers in Human Behavior*, *87*, 238–246.

Belsley, D. A., Kuh, E., & Welsch, R. E. (2004). *Regression diagnostics: Identifying influential data and sources of collinearity*. Hoboken, NJ: John Wiley & Sons.

Boase, J., & Ling, R. (2013). Measuring mobile phone use: Self-report versus log data. *Journal of Computer-Mediated Communication*, *18*(4), 508–519.

Boghe, K., De Grove, F., Herrewijn, L., & De Marez, L. (2020). *Scraping application data from the web—Addressing the temporality of online repositories when working with trace data*. Extended abstract presented at the 70th International Communication Association Conference.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *Inter-*

national Journal of Human-Computer Studies, *130*, 86–92.

Hamari, J. (2015). Why do people buy virtual goods? Attitude toward virtual good purchases versus game enjoyment. *International Journal of Information Management*, *35*, 299–308.

Hamari, J., Alha, K., Järvelä, S., Kivikangas, M. J., & Koivisto, J. (2017). Why do players buy in-game content? An empirical study on concrete purchase motivations. *Computers in Human Behavior*, *68*, 538–546.

Hamari, J., & Keronen, L. (2017). Why do people buy virtual goods: A meta-analysis. *Computers in Human Behavior*, *71*, 59–69.

Hamari, J., & Lehdonvirta, V. (2010). Game design as marketing: How game mechanics create demand for virtual goods. *International Journal of Business Science and Applied Management*, *5*(1), 14–29.

Hsiao, K., & Chen, C. (2016). What drives in-app purchase intention for mobile games? An examination of perceived values and loyalty. *Electronic Commerce Research and Applications*, *16*, 18–29.

Hsiao, C., & Chiou, J. (2012). The effects of a player's network centrality on resource accessibility, game enjoyment, and continuance intention: A study on online gaming communities. *Electronic Commerce Research and Application*, *11*, 75–84.

Jung, E., Baek, C., & Lee, J. (2012). Product survival analysis for the App Store. *Marketing Letters*, *23*(4), 929–941.

Lee, G., & Raghu, T. S. (2014). Determinants of mobile apps' success: Evidence from the App Store market. *Journal of Management Information Systems*, *31*(2), 133–170.

Luton, W. (2013). *Free-to-play: Making money from games you give away*. Boston, MA: Addison Wesley.

Margolin, D. B. (2019). Computational contributions: A symbiotic approach to integrating big, observational data studies into the communication field. *Communication Methods and Measures*, *13*(1), 1–19.

Naab, T. K., Karnowski, V., & Schlütz, D. (2018). Reporting mobile social media use: How survey and experience sampling measures differ. *Communication Methods and Measures*, *13*(2), 126–147.

Neys, J. L., Jansz, J., & Tan, E. S. (2014). Exploring persistence in gaming: The role of self-determination and social identity. *Computers in Human Behavior*, *37*, 196–209.

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, *16*(1), 105–114.

Prinja, S., Gupta, N., & Verma, R. (2010). Censoring in clinical trials: Review of survival analysis techniques. *Indian Journal of Community Medicine*, *35*(2), 217–221.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Ryan, R. M., Rigby, S. C., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, *30*(4), 344–360.

Scharkow, M. (2016). The accuracy of self-reported internet use: A validation study using client log data. *Communication Methods and Measures*, *10*(1), 13–27.

Schönrock-Adema, J., Heijne-Penninga, M., Van Hell, E. A., & Cohen-Schotanus, J. (2009). Necessary steps in factor analysis: Enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Medical Teacher*, *31*(6), 226–232.

Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2019). Integrating survey data and digital trace data: Key issues in developing an emerging field. *Social Science Computer Review*. Advance online publication. https://doi.org/10.1177/0894439319843669

Suissa, S. (2008). Immortal time bias in pharmacoepidemiology. *American Journal of Epidemiology*, *167*(4), 492–499.

Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(78). https://doi.org/10.1186/1479-5868-9-78

Teng, C., & Chen, W. (2014). Team participation and online gamer loyalty. *Electronic Commerce Research and Application*, *13*, 24–31.

Therneau T. (2020). A Package for Survival Analysis in R. R package version 3.2-3. [Computer software]. Rochester, MN: Mayo Clinic. Retrieved from https://CRAN.R-project.org/package=survival

Therneau, T. M., Crowson, C., & Atkinson, E. (2017). *Using time dependent covariates and time dependent coefficients in the cox model*. Retrieved from https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf

Therneau, T. M., Grambsch, P. M. (2000). *Modeling survival data: Extending the cox model*. New York, NY: Springer.

Wijman, T. (2020). The world's 2.7 billion gamers will spend $159.3 billion on games in 2020; the market will surpass $200 billion by 2023. *Newzoo*. Retrieved from https://newzoo.com/insights/articles/newzoo-games-market-numbers-revenues-and-audience-2020-2023

Zhang, Z., Reinikainen, J., Adeleke, K. A., Pieterse, M. E., & Groothuis-Oudshoorn, C. (2018). Time-varying covariates and coefficients in Cox regression models. *Annals of Translational Medicine*, *6*(7), 121.

About the Authors

Kristof Boghe is a PhD Student at Ghent University. His research interests lie in all things related to the computational turn in the social sciences and digital methods. He is fascinated with how the sheer abundance and availability of digital data can give us new insights on the inner workings of our (offline or online) social world. More specifically, he wants to unravel the implications of doing research in the age of big data on theory formation within communication sciences.

Laura Herrewijn is a Postdoctoral Researcher and Assistant Professor in Digital Marketing at the Department of Communication Sciences at Ghent University. Her research can be found on a crossroads between marketing communication and game studies, with a strong focus on virtual environments (such as those represented in digital games and virtual and augmented reality apps) and their potential for the integration of persuasive communication (e.g., advertising, educational content), along with consumer behavior (e.g., in-app purchases).

Frederik De Grove is an Assistant Professor at the Department of Communication Sciences (Ghent University), where he teaches Digital Methods. He successfully defended his PhD in 2014 and obtained a MA in Statistical Data Analysis in 2019 (Ghent University). In addition, he is Project-Coordinator for the GOA project NewsDNA (2018–2022). The main goal of this interdisciplinary project is to promote news diversity through algorithmic development.

Kyle Van Gaeveren started as a Junior Researcher at MICT in November 2018. He graduated as a Master in Sociology (Ghent University, 2011) and Social Work and Welfare Studies (Ghent University, 2014), and is currently obtaining his BA in Computer Science (Karel de Grote Hogeschool). He joined MICT on the Mobile DNA project with a central focus on analyzing log data of smartphone usage.

Lieven De Marez is a Professor at the Department of Communication Sciences at Ghent University. At the department, he is heading the multidisciplinary research group for Media, Innovation and Communication Technologies (MICT), affiliated to imec, the world-leading R&D and Innovation Hub in Nanoelectronics and Digital Technologies. Within imec, Lieven is the founding father of the Digimeter (http://www.imec-int.com/en/digimeter), a reference monitor on the adoption and use of digital technologies.