Master's Thesis of Joo Hyun Park

# Empirical Analysis of Add-on Price Elasticities in the Freemium Mobile Game Market

- Eliminating unobserved heterogeneity, using hierarchical models -

프리미엄(Freemium) 모바일 게임 시장의애 드온(Add-on) 가격 탄력성에 대한 실증적 연구

February 2023

College of Business Administration
Seoul National University
Marketing Major

Joo Hyun Park

# Empirical Analysis of Add-on Price Elasticities in the Freemium Mobile Game Market

## - Eliminating unobserved heterogeneity, using hierarchical models -

Examiner Jun B. Kim

Submitting a master's thesis of
Business Administration

October 2022

College of Business Administration
Seoul National University
Marketing Major

Joo Hyun Park

Confirming the master's thesis written by
Joo Hyun Park
December 2022

| | | |
|---|---|---|
| Chair | 주 우 진 | (Seal) |
| Vice Chair | 김 상 훈 | (Seal) |
| Examiner | 김 준 범 | (Seal) |

#### Abstract

Freemium strategies contain both ‘free’ and ‘premium' options, offering some products or services for free as a sample to encourage paid option sales and expand their user base (Kumar, 2014; Liu et al., 2014; Gu et al., 2018). Distributing basic app downloads for free as a sample and selling paid options, usually through in-app purchases (IAP), has become a prevalent freemium strategy among mobile apps.

This paper empirically analyzes the freemium mobile game users' reaction to the price of add-ons using mobile game transaction data provided by an app store. The observed add-on price in the data is constant over time. Since the add-on information is not included in the dataset (e.g., the characteristics of add-ons, or the quality level of add-ons), the categorical information is limited. There are insufficient game characteristics to capture all the game-level variations, and the add-on price is the only add-on-level variable. The freemium mobile game users can download and experience the games before they purchase add-on options and infer the quality of the games and add-ons. Therefore, it is crucial to include add-on level intercepts to separate the impact of game-level and add-on-level heterogeneity and correctly specify the impact of the add-on price.

First, this research aims to determine mobile game users’ reactions to the add-on price of apps that use freemium strategies. Second, this study aims to find a categorical intercept that efficiently captures the time-invariant bias. Since the add-on price is time-consistent, including add-on level intercepts in the linear demand models is impossible. This study includes profit-maximizing firm assumptions to obtain a 2-stage model with add-on level intercepts. The categorical heterogeneity can be captured by including fixed or random intercepts. Third, reflecting the multi-level structure of this data is the objective of this paper. The price coefficient can be specified at the genre level. Since the data used in this paper is hierarchical, the Bayesian inference method was implied to improve the understanding of the multi-level structure of the models.

**Keywords: Mobile Game, Freemium, Price Elasticity, In-App Purchase, Pricing, Conversion**

**Student Number:** 2020-25878

# Table of Contents

# Table Index

# Figure Index

# I. Introduction

## 1.1. Study Background

The game industry has grown significantly over the previous few years (Wijman, 2020; Business of Apps, 2021; Statista, 2022b; Dobrilova, 2023). [①] The mobile-game application market, which is easy to access through smartphones, is accelerating its growth. According to Statista (2022b), the number of mobile game users in 2021 was over 1.8 billion, which has increased by 11% compared to the previous year. Mobile game application downloads account for 25% of all downloads in IOS application stores and 21% of the Android app stores (Business of Apps, 2021; Dobrilova, 2022). Also, the mobile game segment is the most profitable among mobile apps, generating around 60% of the total application revenue worldwide (Statista, 2022a).

Freemium strategies contain both ‘free’ and ‘premium' options. As we can get the hint from the name, firms using a freemium strategy provide some portion of their product or service for ‘free,’ usually at the early stage of their business, and sell ‘premium’ options afterward (Kumar, 2014; Liu et al., 2014; Gu et al., 2018; Holm & Gunzel-Jenson, 2017). The freemium strategy is popular among online · mobile-based firms and can be commonly

[①] According to Statista (2022a), the report in website Statista, total revenue of mobile games category is expected to reach 267 billion USD in 2022. In addition, Dobrilova (2023) in website Techjury shows that the size of the mobile games market USD by the end of 2021 is 175 billion USD.

seen in mobile application stores, including Google Play and Apple Appstore (Liu et al., 2014). Freemium strategies are also famous for mobile game applications. The total revenue of mobile game apps is projected to be 267 billion USD in 2022 (Statista, 2022a). The income from paid applications is projected to be 1.25 billion USD in 2022, less than 0.5% of the predicted total revenue (Statista, 2022a). In-app purchase (IAP) revenue is expected to be the primary source of the mobile game apps' total income, creating 54.4% (Statista, 2022a). The advertising revenue is predicted to be the secondary source of mobile game apps' total revenue in 2022, generating 45.1% (Statista, 2022a).

Freemium firms design the structure of the freemium strategy to maximize their profit, considering the balance between the expansion of their business and the profit from the premium options. According to Lee et al. (2017), consumers choose to upgrade from the free version to the paid version considering their utility function. Therefore, firms need to understand the choice drivers of the consumers to design the freemium structure that gives the appropriate level of incentives for consumers to achieve the balance between growth and monetization (Lee, Kumar & Gupta, 2017; Holm & Gunzel-Jenson, 2014).

The price of freemium add-ons provided through in-app purchases is one of the critical factors impacting the conversion rate. Free-base app users converge to the premium option when the additional value they expect to get from the premium option is positive, considering the price and quality of each option. Therefore, investigating the consumers' response to the price of in-app purchase add-ons would help improve the understanding of freemium pricing strategies.

## 1.2. Research Objectives

First, this research aims to determine mobile game users' reactions to the add-on price of apps that use freemium strategies. This paper empirically analyzes mobile game users' reactions to the price of add-on options in the freemium mobile game application market, using the add-on level aggregated panel data. Analyzing the price sensitivity of mobile game users can be helpful to the pricing decisions of mobile game apps. The price coefficient can also be estimated by the genre of the game. Since the data used in this paper is hierarchical, the Bayesian inference method was implied to improve the understanding of the multi-level structure of the models.

Second, this study aims to include a time-invariant intercept that efficiently captures the categorical heterogeneities. The users can infer the quality of the game and add-ons from their experience during the free usage. The quality of the product will mainly impact the users' conversion behavior. To efficiently capture the impact of the games and add-on' observed quality for the users, cross-sectional fixed or random coefficients can be added to the demand models. The users can infer the quality of the game and add-ons from their experience during the free usage. The quality of the product will mainly impact the users' conversion behavior. To efficiently capture the impact of the observed qualities of the games and add-ons for the users, cross-sectional fixed or random coefficients are added to the demand models. However, including j-level fixed effects as a linear regression demand model is impossible because of the j-level price variable. Therefore, the supply-side assumption that the company sets the price to maximize profits was made to derive price equations and obtain a 2-stage model with add-on level intercepts.

This paper can contribute to the field of marketing in two parts. First, this paper can contribute to the research field of freemium pricing strategies. There have been some studies on freemium pricing, but not much research has been done on consumers' responses to the price of add-ons of freemium mobile apps that provide base apps download for free and get profit from in-app purchases. This study empirically analyzes the demand price elasticities of add-on options in freemium mobile game apps providing a free base.

Second, this research includes a multi-level structure with more than three levels. Since the data used in this paper is hierarchical, the Bayesian inference method was implied to improve the understanding of the multi-level structure of the models. By using both the demand and supply-side structural models, the estimation of the multi-level structures can be much easier. From the first stage, using the transformed demand equation, the unobserved characteristics of the cross-sectional groups can be captured by including add-on level intercepts.

The research on freemium pricing strategies and the effect of free samples in freemium apps will be introduced as the research background of this study. Section 3 shows the structural modeling process of the demand models, including the supply-side assumption. Data and variables will be explained in Section 4. The demand models with various fixed effects will be compared in Section 5, and the genre-specific price coefficients estimated from Bayesian estimation will be shown in Section 6. Section 7 will contain the conclusions and discussions of this paper.

# II. Literature Review

## 2.1. Freemium Strategies in the Mobile game Industry

Freemium strategies are commonly used in the mobile game industry.[②] According to the Business of Apps, 98% of Google Play revenue comes from free apps (Tafradzhiyski, 2022). Also, more than 99% of the total mobile game application revenue is predicted to come from freemium game apps

---

[②] The link below is a figure that shows the proportion of paid apps from Statista. According to this figure, 3.3% of apps are paid apps in Google Play, and 7.3% of apps in iOS Appstore are paid apps.
https://infogram.com/global-share-of-free-apps-vs-paid-apps-1h7j4dv9m1eq94n

(Statista, 2022a). In-app purchase (IAP) revenue is the primary source of the mobile game apps' total revenue creating 54.4%, and the advertising revenue is the second source creating 45.1% of the total revenue (Statista, 2022a).

For freemium firms, the primary marketing question is what to provide for free. Setting an optimal target conversion rate considering the target market is essential in designing the freemium structure (Kumar, 2014). The optimal conversion rate may differ across the product categories. Sur (2022) states that the average app conversion rate is 31% in US Appstore and 32.7% in the US Google Play. The average conversion rate of each game app category is summarized in Table A1 in Appendix A (Sur, 2022). The differences in the conversion rates across Appstore and Google Play are significant. This paper used Android-based app store data, and the conversion rate in US Google Play varies a lot across categories. Therefore, we can assume that the freemium structure should be designed considering the characteristics of the products and market.

Consumers converge to the paid version when they can get additional values from the paid version, considering the costs. The price of add-on options directly affects mobile game users' in-app purchase intentions (Hsiao & Chen., 2016). It has been argued that the effect of price has an informative side and an allocative role. The informative part of price indicates the quality of the product, and the allocative function of price allocates restricted resources (McConnell, 1968; Gardner, 1971; Monroe, 1973). Jang and Chung (2021) analyzed the impact of absolute and relative add-on prices on sales using the data of paid applications using freemium strategies. The influence of absolute add-on price on sales was positive in their paper, indicating price's informative role. When there is not enough information about the product, consumers infer the quality level of the product from its price. Also, the coefficient of the relative price of an add-on over the price of base apps and other add-on products in the same app was negative, capturing the allocative role of price (i.e., the law of demand).

Application developers should determine the proper level of the free base app quality to achieve the target conversion rate (Deng & Liu, 2022). If the quality of the free version is too high, users' incentive to converge would be low. If the quality of the free version is too low, users will not download the base app. Since the apps are easily downloaded in the mobile app store, attracting consumers to download the base app is a prevalent strategy among mobile app firms to expand their user base. Monetization from the continuous engagement of the application users is possible through in-app advertising (Rutz et al., 2019). Thus, lowering the quality of the free version will increase the conversion rate, but the number of entering users will decrease, which results in decreased in-app ad profits acquired from the advertisers. In addition, the play frequency and social interaction positively impacted the purchase intention of in-app add-ons (Jang et al., 2021). The fun factor indirectly affects the in-app purchase intention through increased loyalty to the mobile game apps (Hsiao et al., 2016). Also, reducing the variety of in-app store items is ideal for reducing the choice overload of users (Ajmera et al., 2022). Around 34.6% of mobile game firms specified that the paid versions are differentiated by providing more levels, according to Deng and Liu (2022).

Surprisingly, 43.3% of the mobile game apps did not have much difference between the free and paid version apps on the app description page, which indicates that the paid version app can be promoted by the free version app (Deng et al., 2022). Therefore, the appropriate quality level of the free version is also essential to attract new users to download the base app and keep the users' engagement level with the app.

## 2.2. Sampling Effect in Freemium Apps

The effect of providing free samples can be explained by the acceleration-cannibalization-expansion (ACE) model, which includes three potential effects: (1) the acceleration effect, (2) the cannibalization effect, and (3) the expansion effect (Bawa & Shoemaker, 2004). First, the acceleration effect happens when consumers buy the brand's product ahead of their schedule when they receive free samples. Second, the cannibalization effect occurs when the consumers try the paid version of the brand less when they receive free samples. Lastly, the expansion effect is the increase of the new users who did not have plans to purchase the brand if they did not receive free samples (e.g., word-of-mouth).

The free sample effects can explain the result of various promotion strategies (e.g., coupons). Since firms using a freemium system provide some portion of the service for free before they sell premium options (Kumar, 2014; Liu et al., 2014; Gu et al., 2018; Li et al., 2019), consumers can experience the basic service or product for free before they decide to purchase add-ons. Therefore, freemium strategies can be explained by the effect of free samples. The users can get information about the original product by experiencing the quality and attractiveness of the free sample (Li et al., 2019). The extent to which the free sample resembles the paid product directly impacts the revenue of the original product, which can be considered as the paid version of the free sample.

In the early stage of the mobile application industry, before in-app purchases became the dominant freemium strategy, free and paid version apps were launched separately. Previous empirical research on freemium strategies in app stores has focused on analyzing multiple versioned apps that include free and paid versions (e.g., Deng et al., 2022). Consumers can get information about the product from the free version apps, which impacts the demand for the paid version because of the free sample effect (Liu et al., 2014; Deng et al., 2022). Deng et al. (2022) used Apple Appstore data for the analysis and found that the launch of free version apps increases the demand, proxied by the application rank, for paid version apps due to the free sample effect. This can be classified as the expansion effect since the market for the paid version increased after providing free sample versions to the users who would not have tried nor purchased the paid version app without the free samples. Deng et al. (2022) maintain that providing horizontal differences (i.e., more modes or themes) in the paid version is not practical, whereas vertical differences (i.e., more levels) are significantly effective. Providing a free version app with fewer themes may have a cannibalization effect on the paid version app.

Improved chance to discover the application in the app store due to the free version can also increase the demand for the paid version (Deng et al., 2022). Liu, Au, and Choi (2014) analyzed the relationship between the review ratings of the free version apps and the revenue of the paid version apps, using the app panel data of the apps registered in Google Play, which turned out to have a positive connection. They showed that the number of reviews and good ratings of the reviews positively impact the probability of a higher app ranking, which has more chance of exposure to potential users. Free sample promotion can have a long-term impact on the brand's sales and may not stabilize to a certain point, considering both the acceleration effect and cannibalization effect in sampling theory (Bawa & Shoemaker, 2004). Also, the expansion effect can cause a long-term carryover effect. The repeat purchase rate influences the potential long-term impact.

The launch of the free apps had a more significant positive impact on the sales of popular products than less popular products (Li et al., 2019). This can be explained by the expansion effect and the acceleration effect. Also, the free version app had a more considerable impact on the paid version app sales than providing a limited period free trial (Liu et al., 2014). If the repeat purchase probabilities are small, the free sample can cause a cannibalization effect on the paid version (Bawa & Shoemaker, 2004). Again, constructing a proper incentive structure to maintain a certain level of conversion rate is crucial in freemium strategies.

The relevant studies related to the freemium strategies are summarized in Table 1 below. The mobile game application category is commonly used as the focus product in the studies of freemium strategies (e.g., Rutz et al., 2019; Jang and Chung, 2021; Deng et al., 2022). This paper also used mobile game apps as the focal product. This study analyzes the price sensitivity of users in freemium game apps that provide free base apps and sell paid options by in-app purchase add-ons.

**Table 1.** Summary of Relevant Freemium Studies.

| Research | Research Focus | Relevant Findings | Research Method | Focus Product | Freemium Structure |
|---|---|---|---|---|---|
| Liu et al. (2014) | The impact of the free version app on the paid version app sales. | The freemium strategy of providing the free version apps positively impacts the sales of the paid version apps. Positive trial experience on the free app leads to an increase in the paid version app's sales. Also, the impact of these review ratings is lower for the hedonic apps. | Empirical analysis using the ranking of free and paid apps as a proxy of sales for the apps registered in Google Play. | Mobile Apps. | Paid version apps with or without the free version app. |
| Rutz et al. (2019) | To analyze the factors that impact the user engagement level and forecast the usage of mobile game apps. | The heterogeneity of the engagement level across mobile games is significant. Therefore, mobile apps must collect and analyze users' usage data to forecast engagement and profit from in-add advertising. | Empirical analysis using mobile game-level aggregated usage data collected by market research firms tracking the users' usage behavior. | Mobile Games. | Free apps and ongoing in-app advertising. |

| | | | | | |
|---|---|---|---|---|---|
| Li et al. (2019) | A theoretical framework for the optimal sample quality level for freemium firms. | Freemium firms should offer higher quality free samples when the paid version quality is higher, users' attention is greater, and the functionality of the free version is lower. | Theoretical framework of sample quality/ Field experiment. | Publications. | Separate market for books and access-free PDFs. |
| Jang & Chung (2021) | Impact of add-on price on sales of the add-on products for the freemium firms. | The impact of price can be divided into absolute and relative prices. The absolute price positively impacts the sales of add-ons due to the preference for the add-on quality. The relative price over the base app and the other add-on in the same game harms the add-on sales. | Empirical analysis using Asia's leading app store data. | Mobile Games. | Paid base apps and paid add-ons. |
| Deng et al. (2022) | The impact of the launch of the free version app on the existing paid version app. | The launch of free version apps increases the sales of the paid version app due to the sampling effect and the improved chance to discover the app in the app store. | Empirical analysis using app store data collected from Apple Appstore by a data gathering company. | Mobile Games. | Paid version apps with or without the free version app. |
| This research | Estimating the price-elasticity of the free apps' IAP add-ons. | Including add-on level fixed effects by adopting supply-side assumption and Bayesian sampling method can when the price is constant during the observed period. | Empirical analysis using app store data provided by ONE Store. | Mobile Games. | Free apps and paid add-ons. |

**Table 1 (Continued)**

# III. Model

## 3.1. Demand Model

To estimate the consumers' reaction to the price of game in-app purchase add-ons in the freemium game application industry, we need to construct demand models that include add-on price as an independent variable. Log-linear demand functions are regularly used to estimate price elasticity since the coefficient of the log-log model can be considered as the percentage point change, which is the definition of price elasticity. However, during the profit-maximizing process, which will be explained later in this section, the supply-side restrictions cannot be sustained using log-linear models. Therefore, linear demand models were used in this paper to adopt supply-side assumptions despite the convenience of the log-linear model.

The demand function can be written as equation 1 below using the add-on-level monthly aggregated panel data. The $\mathrm{q_{jt}}$ term represents the number of add-on j's sales quantity in month t. The model includes time-specific dummy variables, $\beta_t$, to capture the sales fluctuation caused by time-specific circumstances. The price of add-on j can be marked as a time-invariant denotation $P_j$, since the add-on price stays constant during the observed period. Several time-invariant game characteristics variables $X_g$ and time-variant add-on-level $X_{jt}$ variables are included in the models. The residual of the model is denoted as $u_{jt}$ .

$$q_{jt} = \beta_t - \alpha \cdot P_j + \gamma_1 \cdot X_g + \gamma_2 \cdot X_{jt} + u_{jt}, \tag{1}$$

$j: add - on\ j$ .
$t: month\ number\ t, from\ 3\ to\ 11.$
$g: game\ g, where\ j \in g$ .
$X_g: Game\ g - specific\ characteristics, constant\ during\ all\ t$ .
$X_{jt}: Add - on\ j's\ characteristic\ on\ time\ t$ .
$u_{jt}: estimated\ error\ term\ of\ the\ model.$

Since there are not enough variables to explain the heterogeneity within games and add-ons, the model's error term embodies the unobserved variations within cross-sectional groups. The error term, $u_{jt}$ can be separated into three parts in equation 2 below. The unobserved impact of time is not included in equation 2 since it can be explained by $\beta_t$, monthly dummy variables. The unobserved aspect that causes the add-on-level variation is included as $\nu_j$ in the model. The game-level term $\omega_{g(j)}$ represents the unobserved heterogeneity among the games. The game g, where add-on j belongs, can be denoted as $g(j)$, since the add-ons are fully nested within games. The model is $git$ level 3-dimensional model, but here in this paper will be denoted as $jt$. The $\varepsilon_{jt}$ term expresses the actual errors of the demand model.

$$u_{jt} = \omega_{g(j)} + \nu_j + \varepsilon_{jt}\ , \tag{2}$$

The demand model, including equation 2 in equation 1, is demonstrated as equation 3. The actual errors, $\varepsilon_{jt}$ , are independent and identically distributed (i.i.d), following a normal distribution with zero mean and jt level variance. One of the simplest ways to absorb the unobserved categorical influence is to improve the model specification by including the fixed effects in the model. Since the fixed effects work as a group of dummy variables, including fixed effects clears the unobserved aspect included in the error variance. However, the models with cross-sectional group-level variables cannot include the same-level fixed effects. Considering that the price variable is indispensable in the demand model and the main purpose of this study is to measure the users' reaction to the add-on price, add-on level fixed effects cannot be included in the demand model based on equation 3.

$$\mathrm{q_{jt}} = \beta_t - \alpha \cdot P_j + \gamma_1 \cdot X_g + \gamma_2 \cdot X_{jt} + \omega_{g(j)} + \nu_j + \varepsilon_{jt}\,, \tag{3}$$

$$where \quad \varepsilon_{jt} \sim i.i.d\; N(0, \sigma_{jt}^2)$$

The game-specific fixed effects can be included in the model when the game-specific $X_g$ variables are excluded as equation 4 below. The game-level fixed effects are denoted as $\beta_{FE(g)}$. Since the fixed effects (cross-sectional dummies) can efficiently capture the categorical variations, the $\omega_{g(j)}$ term in equation 3 can be eliminated as in equation 4. Since the data structure is strictly hierarchical, this model containing the game-specific fixed effects can explain a certain proportion of add-on characteristics. Add-ons in the same game share some common grounds since it is impossible to buy add-ons unless one downloads the game application first. However, the fluctuations within add-ons in the same game remain unexplained. The unobserved within-game divergences are expressed as $\nu_j^g$. The challenge of this data is figuring out a method to capture add-on level differences in the model without interrupting the price term.

$$\mathrm{q_{jt}} = \beta_{FE(g)} + \beta_t - \alpha \cdot P_j + \gamma_2 \cdot X_{jt} + \nu_j^g + \varepsilon_{jt}\,, \tag{4}$$

This study aims to include add-on level constant terms to capture the add-on level characteristics in the model. Equation 5 shows the demand model that fits this paper's goal. There are no add-on level variables except for the fixed price, and it is crucial to include add-on level fixed effects in the model to improve the estimated models' robustness by reflecting the add-ons' true qualities. Since add-on prices do not vary over time, we cannot combine add-on level coefficients in the linear demand models. The class of game g, where add-on j belongs, also does not vary over time, which means that the game level is a perfect subset of the game's constant characteristics class variables (e.g., genres). The data structure of the models is super-nested and needs to be untangled by the structural construction of the variables.

$$\mathrm{q_{jt}} = \beta_{FE} + \beta_t - \alpha \cdot P_j + \gamma_2 \cdot X_{jt} + \varepsilon_{jt}\,, \tag{5}$$

## 3.2. Supply-Side Assumptions

To design structural models for the multi-level data, some papers included supply-side assumptions in the model construction to draw additional equations needed for the estimation. Assumptions of the profit-maximizing firms are generally proposed and implemented by obtaining the price equation earned from First Order Condition (FOC) derivatives of the profit function Thomadsen (2005) applied the supply-side assumption that fast food franchise companies set the prices of their products at the outlet level, which maximizes aggregated franchise firms' profit. Gentzkow (2007) considered both the demand and supply sides of the online and offline newspaper market. Since the price of the newspaper did not vary over the observed period, the authors adopted the profit-maximizing price assumption.

As we discussed, the price of add-on j does not change, and it is impossible to include add-on fixed effects in this model. There may be unobserved parts of add-on j that may affect the result of this model. Therefore, we can presume that free mobile game apps maximize profit by setting optimal prices. This paper uses FOC restrictions to build the price equation for the model. After combining equations from the FOC 2-stage model, we can use the add-on specific fixed effects.

The mobile game app companies' profit model can be expressed as equation 6 below. The profit of game g can be subscripted as $\Pi_g$, the sum of the profits of all add-ons provided by the firm, which is subscripted as $\sum_{j \in g} \Pi_j$ . The add-on level profit is aggregated among all periods. Since the price does not change over time, we can assume that the game applications decide their optimal price considering the total profit across the periods. The first order condition is the derivative of the profit function with $P_j$, since the assumption that the firms set the optimal price for their profit is made. It is essential to check whether each term is related to the $P_j$.

Equation 6 shows the replacement of $q_{jt}$ in the profit equation with the demand equation 5. Since some games are launched during the observed period, and some add-ons may have been launched during the period, the number of the included t differs among the add-ons. The time-variant variables are aggregated by each add-on level. The time-consistent terms can be aggregated by multiplying the number of the periods included by add-on j.

$$\Pi_{\mathrm{g}} = \sum_{\mathrm{j} \in \mathrm{g}} \Pi_{\mathrm{j}} = \sum_{\mathrm{j} \in \mathrm{g}} \sum_{\mathrm{t} \in \mathrm{t_j}} \Pi_{\mathrm{jt}} = \sum_{\mathrm{j} \in \mathrm{g}} \sum_{\mathrm{t} \in \mathrm{t_j}} \left( (\mathrm{P_j} - \mathrm{c_j}) \cdot \mathrm{q_{jt}} - FixedCost_{\mathrm{jt}} \right)$$

$$= \sum_{j \in g} P_j \cdot \left( T_j \cdot (\beta_j - \alpha \cdot P_j) + \sum_{t \in t_j} (\beta_t + \gamma_2 \cdot X_{jt} + \varepsilon_{jt}) \right), \qquad (6)$$

$\Pi_g$ *is the profit of game* $g$, *which is the sum of all add on profits in game* $g$.

$\Pi_j$ *is the profit of add on* $j$, *which is the sum of profits during all periods.*

$t_j$ *is the set of all time period* $t$ *included by add on* $j$.

$T_j$ *is the number of* $t$ *included in the data by add on* $j$.

If the firms maximize their profit, the first-order condition can be applied. The profits of the add-ons are not related in this model structure, so the partial of $P_j$ of the game profit has the same value as the part of the add-on profit. Then, we can get the additional equation of $P_j$, shown as equation 7. Marginal cost $c_j$ can be considered as zero since the transaction cost is low for online · mobile-based products. Also, $OtherProfit_{jt}$ and $FixedCost_{jt}$ are assumed not to be related to the product j's add-on price $P_j$. Therefore, the partial of profit $P_j \cdot q_{jt}$ over the partial of the price j, partial of is the value to be calculated during the FOC process.

$$\partial \Pi_g / \partial P_j = T_j \cdot (\beta_j - \alpha \cdot P_j) + \sum_{t \in t_j} (\beta_t + \gamma_2 \cdot X_{jt} + \varepsilon_{jt}) - \alpha \cdot T_j \cdot P_j = 0 \,, \tag{7}$$

$$for\ all\ add\ on\ j\ in\ game\ g\ (j \in g).$$

Rearranging the subscripts of the average time aggregated terms, the price equation can be expressed as equation 8. The price equation's error term, $\psi_j$ follows normal distributions including the error variations in the demand model. The error term follows a normal distribution since the original error term follows i.i.d normal restriction. This price equation is an add-on level equation since the time-variant terms were aggregated by the add-on level.

$$P_j = \frac{1}{2 \cdot \alpha} \cdot (\beta_j + \overline{\beta_t} + \gamma_2 \cdot \overline{X_{J.}} + \overline{\varepsilon_{J.}}) = \frac{1}{2 \cdot \alpha} \cdot (\beta_j + \overline{\beta_t} + \gamma_2 \cdot \overline{X_{J.}}) + \psi_j \,, \tag{8}$$

$$\frac{1}{T_j} \cdot \Sigma_t \beta_t = \overline{\beta_t}$$

$$\frac{1}{T_j} \cdot \Sigma_t X_{jt} = \overline{X_{J.}}$$

$$\frac{1}{T_j} \cdot \Sigma_t \varepsilon_{jt} = \overline{\varepsilon_{J.}} \,,$$

$$\psi_j = \frac{1}{2\alpha} \cdot \overline{\varepsilon_{J.}}$$

The price term $P_j$ in equation 5 can be replaced by equation 8 above, and the resulting demand equation is expressed as equation 9 below. The updated demand model does not contain cross-sectional terms. Thus, add-on specific fixed effects can now be included in the demand model estimation. The estimation process is divided into the 2-stages. In the first stage of the estimation process, we estimate the parameters in equation 9. Then in the second stage, we use the estimated parameters from the first stage as a proxy for the estimation of the price in equation 8. The price coefficient alpha can be measured during the second stage. The error term in equation 5 follows i.i.d normal distribution under the condition that the unobserved cross-sectional variations are correctly captured. If the error terms are identically and independently distributed across the time and cross-sectional levels, then the error term $\xi_{jt}$ in equation 9 follows a normal distribution.

The FOC function considers the firms' dynamic decisions over time. In previous econometric research, time-variant variables' average term over time is used as instrumental variables to separately examine the impact of multiple cross-sectional levels (e.g., Hausman and Taylor, 1981; Balazsi et al., 2018; Yang and Schmidt, 2021). In the dynamic model, using nested panel data, including proper terms that effectively capture the cross-sectional variations of all levels, is an important issue. Models including time-consistent variables, will be estimated and compared in this paper.

$$q_{jt} = \frac{1}{2} \cdot \beta_{FE} + \left(\beta_t - \frac{1}{2} \cdot \overline{\beta_t}\right) + \gamma_2 \cdot \left(X_{jt} - \frac{1}{2} \cdot \overline{X_{j.}}\right) + \xi_{jt}, \quad (9)$$

$$where\ \xi_{jt} = \varepsilon_{jt} - \overline{\varepsilon_{j}}., if\ not\ biased.$$

# IV. Data and Descriptive Analysis

## 4.1. Mobile-game App Store Data

Mobile game apps are the leading products in the app store, in which the mobile game category is the most profitable among mobile app categories, generating around 60% of the total application revenue worldwide (Statista, 2022a). The data used in this paper is mobile-game application download and in-app purchase panel data provided by ONE Store, one of the most prominent application transaction sites in South Korea.[③]

The data contains about 1.5M app store transaction data during 2018. Mobile game in-app purchases and base-app download data were used in this paper. Most apps use freemium strategies, meaning the base-app downloads are mostly free. Among the included mobile game apps, 91% were free-based app products, which shows the popularity of the freemium pricing strategies in the mobile game category.

③ According to CISION PR Newswire (2022) article, ONE store is the second largest app market in Korea. ONE store had the second largest market share of South Korea's application market of 13.8% in 2021. This exceeds the market share of Apple Appstore's market share of 11.6%.

Considering the lagged variables, the first observations of each add-on were dropped. Add-on level aggregated panel time series data is used in this paper. Games with more than nine monthly observations were included in the final dataset. The number of game apps included in the final dataset is 216, and these games had 4,108 add-ons.

## 4.2. Independent Variables

The price variable is denoted as $P_j$ in the models. The unit of the price variable $P_j$ is 1,000 KRW. Time-specific monthly dummies are denoted as $\beta_t$, where t = 3, 4, … ,12. The number of observations included in the game differs since the number of add-ons is different for each game, and also, the number of periods included for each add-on differs. Game-level variables, denoted as $X_g$, were used to analyze the linear demand model without the fixed effects based on equation 1. The $X_g$ variables include game-level categorical dummy variables of $genre$, $productgrade$, and $seller$. Since some games do not have past download information, $DL0unknown$ is included as a dummy variable. Without this dummy, the cumulated download variable will be correlated with game-specific characteristics. Considering the given game-level cross-sectional variables, $filesize$, $numaddons$, of $genre$, $productgrade$, and $seller$ the quality of games may not be correctly specified.

$$X_g \begin{cases} genre\text{:}genre\ dummies\ of\ games. \\ \quad \begin{cases} genre\ RPG: 1\ if\ genre\ is\ RPG,\ \ o.w.\ 0. \\ genre\ simulation: 1\ if\ genre\ is\ simulation,\ \ o.w.\ 0. \\ genre\ puzzle: 1\ if\ genre\ is\ puzzle,\ \ o.w.\ 0. \\ genre\ sports: 1\ if\ genre\ is\ sports,\ \ o.w.\ 0. \\ genre\ action: 1\ if\ genre\ is\ action,\ \ o.w.\ 0. \\ genre\ others: if\ all\ genre\ dummies\ are\ 0. \end{cases} \\ productgrade\text{:}product\ grade\ dummies\ of\ games. \\ \quad \begin{cases} grade\ adult\text{:}\ 1\ if\ p.g.\ adult,\ \ o.w.\ 0. \\ grade\ OVER15\text{:}\ 1\ if\ p.g.\ is\ over\ 15,\ \ o.w.\ 0. \\ grade\ OVER12\text{:}\ 1\ if\ p.g.\ is\ over\ 12,\ \ o.w.\ 0. \\ grade\ others\text{:}if\ all\ grade\ dummies\ are\ 0. \end{cases} \\ seller\text{:}seller\ type\ dummies\ of\ games. \\ \quad \begin{cases} seller\ foreign\text{:}\ 1\ if\ seller\ is\ foreign\ o.w.\ 0. \\ seller\ corporate\text{:}\ 1\ if\ seller\ is\ corporate\ \ o.w.\ 0. \\ seller\ others\text{:}\ if\ all\ seller\ dummies\ are\ 0. \end{cases} \\ DL0unknown\text{:}\ dummy\ of\ g\ with\ unknown\ past\ downloads. \\ filesize\text{:}filesize\ of\ each\ game. \\ numaddons\text{:}\ number\ of\ add{-}ons\ for\ each\ game. \end{cases}$$

There are some time-variant variables across add-ons, denoted as $X_{jt}$. Since the number of observations vary among the add-ons, jt level variables are included $where\ t \in T_j$. The $age_{gt}$ variable denotes the month number after the launch of each game g at t. This variable is constructed as a sum of the previous launched periods plus the number of periods passes by. The age variable can be considered as the sum of the game level variable and time level variable because age increases by one every month for every game. Thus, $age_{gt}$ cannot be included in the models with fixed effects. The number of downloads at t-1 is expressed as $DL_{g,t-1}$. The cumulated downloads from the launch to the t-2 period are denoted as $cumulDL_{g,t-2}$. Both download terms' unit is 1,000 downloads.

Four promotion variables and four promotion lag variables are included in the analysis. The proportion of payments made by a specific promotion during period t is used as a proxy of the level of promotion at period t. The dummy for zero lag quantity was included for the promotion lag variables since the payment proportions are not observed when the sales quantity is zero. The carryover effects in the linear regression model are captured by including the lag promotion and download variables.

$$X_{jt}\begin{cases} age_{g,t}\text{:}number\ of\ months\ after\ launch\ for\ each\ game\ g\ at\ time\ t. \\ DL_{g,t-1}\text{:}numberof\ game\ g\ downloads\ at\ period\ t-1. \\ cumulDL_{g,t-2}\text{:}cumulated\ total\ downloads\ at\ period\ \ t-2. \\ q0_{j,t-1}\text{:}dummy\ of\ whether\ q\ was\ 0\ for\ promotion\ variables. \\ membership_{j,t-1}\text{:}proportion\ of\ membership\ payment\ for\ j\ in\ t-1. \\ membership_{j,t}\text{:}proportion\ of\ membership\ payment\ for\ j\ in\ t. \\ coupon_{j,t-1}\text{:}proportion\ of\ coupon\ payment\ for\ j\ in\ t-1. \\ coupon_{j,t}\text{:}proportion\ of\ coupon\ payment\ for\ j\ in\ t. \\ gamep_{j,t-1}\text{:}proportion\ of\ game\ point\ payment\ for\ j\ in\ t-1. \\ gamep_{j,t}\text{:}proportion\ of\ game\ point\ payment\ for\ j\ in\ t. \\ storep_{j,t-1}\text{:}proportion\ of\ game\ point\ payment\ for\ j\ in\ t-1. \\ storep_{j,t}\text{:}proportion\ of\ game\ point\ payment\ for\ j\ in\ t. \end{cases}$$

## 4.3. Summary Statistics of the Data

First, the entire data is the dataset consisting of 216 games and 4,081 add-ons in those games, which contains 28,983 observations when the non-purchased observations are included. From the transaction data, the number of the sales volume is observed. Only purchase data includes only purchased observations, where q is non-zero. Since the add-on's launch period is unsure, non-purchased observations were included at the intervals of each add-on's first and last purchase occasions. Some games were launched during the observed period, and the number of observations for each add-on diverges. The datasets in this study have an unbalanced structure.

Data from the top 100 games are also analyzed in this research. The criterion of the ranking is the number of sales quantities during all periods, not sales, to determine the impact of the number of transactions. Four datasets are used in the analysis considering the most-selling games and non-zero observations. Table 2 shows the summary statistics of several variables used in the models. The data set of the top 50 data was also included in table 2 for comparison. The average quantity among the entire dataset, including non-purchased observation, is 20.43, and the average price is 32.68 (1,000 KRW). The average price and quantity will be used to calculate the price elasticities from the linear demand models in Section 7. For downloads at period t-1 and the cumulated download variables, the most-selling apps had significantly larger current and cumulated downloads, reflecting the larger user base and a higher degree of interest. Promotion variables show the average purchased proportion of each promotion. Among the promotion variables, coupons and game points are mainly used with around 5~7 percent on average. Considering the differences in the number of observations between the non-purchased and only purchased datasets, about 20% of the total data is non-purchased observations. The age of the products is similar, regardless of their popularity.

**Table 2.** Summary Statistics of the Data.

| | Unit | q | Price (1,000KRW) | DL_lag (1,000 DL) | cum_DL (1,000 DL) | age | membership (%) | coupon (%) | gamep (%) | storep (%) | n_addons |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **All Games** | mean | 20.43 | 32.68 | 7.65 | 449.51 | 26.02 | 0.02 | 0.05 | 0.05 | 0.00 | 59.16 |
| **(Non-Purchased)** | std | 71.04 | 39.35 | 16.50 | 983.08 | 18.91 | 0.03 | 0.11 | 0.16 | 0.02 | 64.77 |
| (28983 obs.) | median | 3.00 | 12.10 | 1.92 | 62.81 | 21.00 | 0.00 | 0.00 | 0.00 | 0.00 | 33.00 |
| **All Games** | mean | 25.40 | 32.32 | 8.04 | 466.07 | 26.27 | 0.02 | 0.06 | 0.07 | 0.00 | 58.11 |
| **(Only Purchased)** | std | 78.41 | 39.88 | 16.79 | 987.03 | 18.98 | 0.03 | 0.12 | 0.17 | 0.02 | 65.21 |
| (23317 obs.) | median | 5.00 | 11.00 | 2.09 | 64.93 | 21.00 | 0.02 | 0.02 | 0.00 | 0.00 | 32.00 |
| **TOP 100 Games** | mean | 28.34 | 33.50 | 10.69 | 577.86 | 25.37 | 0.02 | 0.05 | 0.05 | 0.00 | 75.42 |
| **(Non-Purchased)** | std | 84.73 | 39.36 | 19.13 | 1128.99 | 17.28 | 0.03 | 0.11 | 0.14 | 0.01 | 70.82 |
| (19783 obs.) | median | 5.00 | 16.50 | 3.67 | 101.01 | 22.00 | 0.01 | 0.01 | 0.00 | 0.00 | 49.00 |
| **TOP 100 Games** | mean | 33.59 | 32.98 | 10.78 | 574.01 | 25.55 | 0.02 | 0.06 | 0.06 | 0.00 | 72.06 |
| **(Only Purchased)** | std | 91.29 | 40.11 | 19.07 | 1110.34 | 17.40 | 0.03 | 0.12 | 0.16 | 0.01 | 70.89 |
| (16689 obs.) | median | 8.00 | 12.10 | 3.97 | 104.85 | 22.00 | 0.02 | 0.02 | 0.00 | 0.00 | 46.00 |
| **TOP 50 Games** | mean | 39.17 | 31.97 | 14.62 | 768.26 | 28.60 | 0.02 | 0.05 | 0.05 | 0.00 | 93.08 |
| **(Non-Purchased)** | std | 104.15 | 37.27 | 22.25 | 1283.56 | 18.08 | 0.02 | 0.11 | 0.14 | 0.02 | 79.35 |
| (12389 obs.) | median | 8.00 | 12.10 | 7.07 | 180.96 | 26.00 | 0.01 | 0.02 | 0.00 | 0.00 | 69.00 |
| **TOP 50 Games** | mean | 45.24 | 31.16 | 14.62 | 768.26 | 28.71 | 0.02 | 0.06 | 0.06 | 0.00 | 88.58 |
| **(Only Purchased)** | std | 110.69 | 37.51 | 22.25 | 1283.56 | 18.01 | 0.02 | 0.12 | 0.15 | 0.02 | 79.17 |
| (10728 obs.) | median | 12.00 | 11.00 | 7.07 | 180.96 | 26.00 | 0.02 | 0.03 | 0.00 | 0.00 | 59.00 |

Figure 1 shows the relationship between the price variable and the sales quantities, which approximately shows the 2-dimensional demand curve for the add-ons. The figure shows the plot where the price is the y-axis, and the sales quantity is the x-axis. The price ranges between 100~297,000 KRW, and the quantity ranges between zero to 2232 quantities. The shape of the scatter plot is convex and matches the law of demand (i.e., sales increase when the price is higher) without the influences from other variables, since the shape of the plot looks like it goes down in the right direction.

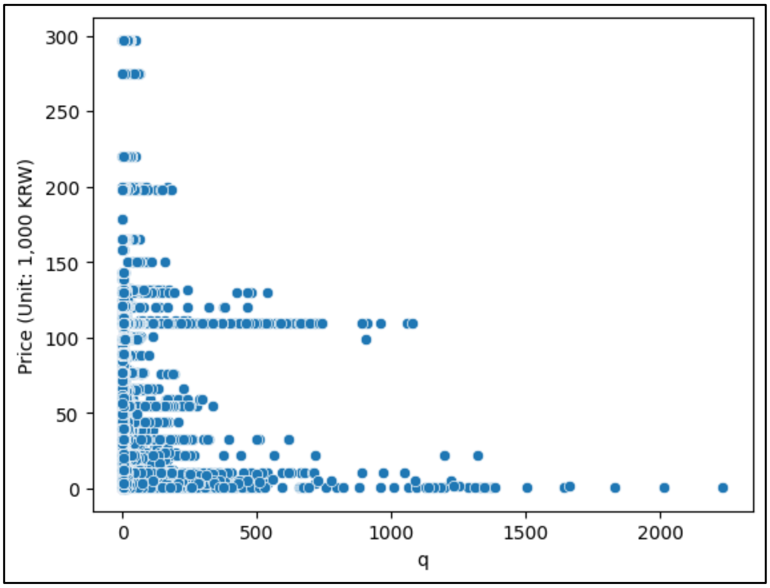


**Figure 1.** Scatter Plot of the Full Data, Including Non-Purchased Observations.

# V. Empirical Analysis with Fixed Effects

## 5.1. Fixed Effect Models with Promotion and Download Lag.

Models with different levels of fixed effects will be compared in this section. The time-consistent fixed effects, $\beta_{FE}$, will be included in the model to estimate the unobserved time-invariant heterogeneity as in equation 3. The cross-sectional dummy variables will capture the characteristics of add-ons beyond price. Since the users experience the game before they purchase add-on options, the users' observed quality level should be included in the model.

The optimal linear demand models with fixed effects can be expressed as equation 5, where the fixed effects well capture the cross-sectional bias. If the unobserved heterogeneity across or within groups is not sufficiently removed from the included variables, the cross-sectional bias will be encompassed in the residual term and interrupt estimating the actual error variance.

First, all data from 216 games and 4801 add-ons were used for the models in Table 3. Table 3 shows the models' coefficients and results with and without the fixed effects. Models A and B are the models without any fixed effects. Models C and D show the estimation results with game-fixed effects. At last, models E and F include the add-on level fixed effects using the FOC 2-stage model described above in Section 3.2.

The price coefficients of models A, B, C, and D, without the supply-side assumptions, are insignificant. Also, the price coefficients of the 2-stage FOC models E and F are more than 300 times larger than all the other models. Since the adjusted R-squared of models A and B are extremely low, it is plausible that the cross-sectional unobserved variations are significant when the cross-sectional effects are only captured by several game-level variables. Even models C and D with game-fixed effects had similar price coefficients to those without fixed effects. If the add-on level heterogeneity is not adequately treated, the estimates of the price variable will also be biased.

For the 2-stage FOC model, the model fit of stage 1 is very high, but the fit of stage 2 is shallow. Here, the game-level variations are not well specified in the model. Stage 2 is the j-level model, and the total observation is 4081, which is the number of j. The FOC model with add-on specific fixed effect captures most of the add-on level variations, but the game-level variations, independent of the add-on level variations, must be captured to increase the model's fitness.

Promotion variables usually have a carryover effect over time. Since we use monthly data of 2018 transactions, the maximum number of observations per add-on is 11, considering one lag variable. Therefore, it is difficult to include multiple lag variables. In this section, the impact of including a one-period lag variable will be specified.

**Table 3.** Comparison of Different Fixed Effect Models – Non-Purchased

| | Fixed Effect | No FE ($\beta_0$) | | | | Game FE ($\beta_g$) | | | | FOC - Add-on FE ($\beta_j$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | promotion lag | w/o | | with | | w/o | | with | | w/o | | with | |
| | Model | **A** | | **B** | | **C** | | **D** | | **E** | | **F** | |
| | (Intercept) | **1.03** | | **5.45** | | **-0.0675** | | **0.88** | | **251.42** | *** | **251.88** | *** |
| | | (0.839) | | (0.284) | | (0.994) | | (0.924) | | (0.000) | | (0.000) | |
| $\alpha$ | **1/alpha** | | | | | | | | | **[STEP 2]:** **0.1838** (0.000) | | **[STEP 2]:** **0.1627** (0.000) | |
| | **price** | **0.014** | | **0.009** | | **0.0188** | . | **0.0131** | | **alpha =** **5.4407** (=1/0.1838) | | **alpha =** **6.1463** (=1/0.1627) | |
| | | (0.186) | | (0.402) | | (0.057) | | (0.184) | | | | | |
| $\gamma_1$ | **genre_RPG** | **8.65** | ** | **7.48** | * | | | | | | | | |
| | | (0.004) | | (0.012) | | | | | | | | | |
| | **genre_simulation** | **14.62** | *** | **13.92** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| | **genre_puzzle** | **10.92** | ** | **11.63** | ** | | | | | | | | |
| | | (0.001) | | (0.001) | | | | | | | | | |
| | **genre_sports** | **24.58** | *** | **24.60** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| | **genre_action** | **7.13** | * | **5.75** | . | | | | | | | | |
| | | (0.035) | | (0.088) | | | | | | | | | |
| | **grade_adult** | **6.04** | ** | **4.14** | * | | | | | | | | |
| | | (0.001) | | (0.027) | | | | | | | | | |
| | **grade_OVER15** | **15.05** | *** | **16.08** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| | **grade_OVER12** | **13.05** | *** | **12.62** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| | **seller_foreign** | **25.56** | *** | **23.37** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| | **seller_corporate** | **8.99** | ** | **6.61** | . | | | | | | | | |
| | | (0.019) | | (0.083) | | | | | | | | | |
| | **DL0_unknown** | **-7.14** | *** | **-6.22** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| | **filesize** | **-0.02** | *** | **-0.01** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| | **num_addons** | **-0.10** | *** | **-0.09** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| $\gamma_2$ | **age** | **-0.17** | *** | **-0.17** | *** | | | | | | | | |
| | | (0.000) | | (0.000) | | | | | | | | | |
| | **DL_lag** | **0.49** | *** | **0.46** | *** | **0.24** | *** | **0.23** | *** | **0.27** | *** | **0.27** | *** |
| | | (0.000) | | (0.000) | | (0.000) | | (0.000) | | (0.000) | | (0.000) | |
| | **cumul_DL** | **0.004** | *** | **0.004** | *** | **0.012** | | **0.011** | | **0.018** | *** | **0.017** | *** |
| | | (0.000) | | (0.000) | | (0.138) | | (0.168) | | (0.000) | | (0.000) | |
| | **q0_lag** | | | **-19.31** | *** | | | **-6.57** | *** | | | **2.01** | * |
| | | | | (0.000) | | | | (0.000) | | | | (0.015) | |
| | **membership_lag** | | | **-78.64** | *** | | | **-23.34** | | | | **-8.17** | |
| | | | | (0.000) | | | | (0.152) | | | | (0.467) | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **coupon_lag** | | **-6.93** . | | **1.16** | | | **-7.48** ** | |
| | | (0.081) | | (0.747) | | | (0.004) | |
| **gamep_lag** | | **10.349** ** | | **15.896** *** | | | **11.255** *** | |
| | | (0.001) | | (0.000) | | | (0.000) | |
| **storep_lag** | | **22.681** | | **-16.328** | | | **4.767** | |
| | | (0.348) | | (0.460) | | | (0.748) | |
| **membership_t** | **59.851** *** | **37.951** * | **51.321** *** | **43.193** ** | **22.121** * | | **22.136** * | |
| | (0.000) | (0.015) | (0.000) | (0.002) | (0.025) | | (0.025) | |
| **coupon_t** | **28.421** *** | **25.211** *** | **22.222** *** | **19.884** *** | **6.768** ** | | **7.220** ** | |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.009) | | (0.005) | |
| **gamep_t** | **29.305** *** | **22.945** *** | **30.989** *** | **24.522** *** | **15.374** *** | | **13.620** *** | |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | | (0.000) | |
| **storep_t** | **44.641** . | **31.695** | **-16.672** | **-18.745** | **-0.13** | | **0.258** | |
| | (0.074) | (0.207) | (0.469) | (0.416) | (0.993) | | (0.987) | |
| | | | | | **Step1** | **Step2** | **Step1** | **Step2** |
| **Adj. R-squared** | 0.047 | 0.057 | 0.250 | 0.252 | 0.718 | 0.030 | 0.718 | 0.030 |
| **-2LL** | 3.279 E+05 | 3.276 E+05 | 3.209 E+05 | 3.207 E+05 | 2.883 E+05 | 4.412 E+04 | 2.882 E+05 | 4.362 E+04 |
| **AIC** | 3.280 E+05 | 3.277 E+05 | 3.213 E+05 | 3.212 E+05 | 2.965 E+05 | 4.412 E+04 | 2.964 E+05 | 4.362 E+04 |
| **BIC** | 3.283 E+05 | 3.280 E+05 | 3.232 E+05 | 3.232 E+05 | 3.304 E+05 | 4.412 E+04 | 3.304 E+05 | 4.363 E+04 |

**Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1**

**Table 3 (continued)**

Since $cumul\_DL$ variable is missing for some games, the variable will include some of the game-level fixed effects unless the dummy variable $DL0\_unknown$ is included. In Table 3, $cumul\_DL$ variable was specified for models A and B, where the past download unknown dummy was included. The variable is also specified when add-on level fixed effects are included, but not specified for the models containing game-level fixed effects. Also, Figure 2 shows that $cumul\_DL$ is highly correlated with the $DL\_lag$ variable. Therefore, cumulated download variables will not be included in the linear regression models with fixed effects.

Table 3 also shows the impact of including the promotion lag variables. A, C, and E are the models without the lag promotion term, only including current promotions in period t, whereas B, D, and F are the models with both current and lag promotion variables. Considering the log-likelihood criteria, the model's fit was slightly larger after including the promotion lag variables. From the correlation plot of Figure 2, the promotion variables are less correlated between the lag and current level variables than the download variables. Game point variables have the largest correlations among the promotion variables, with 0.43. The FOC models with promotion lag variables had higher price coefficient estimates, and the other models had lower price coefficients, including the promotion lag variables. The estimate of membership lag was negative for all models, and the estimate of the current membership variable had positive coefficients. The current and lag game point variables were all estimated to have positive numbers. The promotion lag variables seem to impact the price coefficients and will be included in the linear regression models since the correlation is less than the cumulated download variable.

In summary, models without using the 2-stage model estimate of price coefficients are very low, even with the game-specific fixed effects. This seems to be biased by the unobserved add-on variations. Also, the model with add-on level fixed effects well specified the add-on level variations but seems to have limitations in capturing the game-level variations.

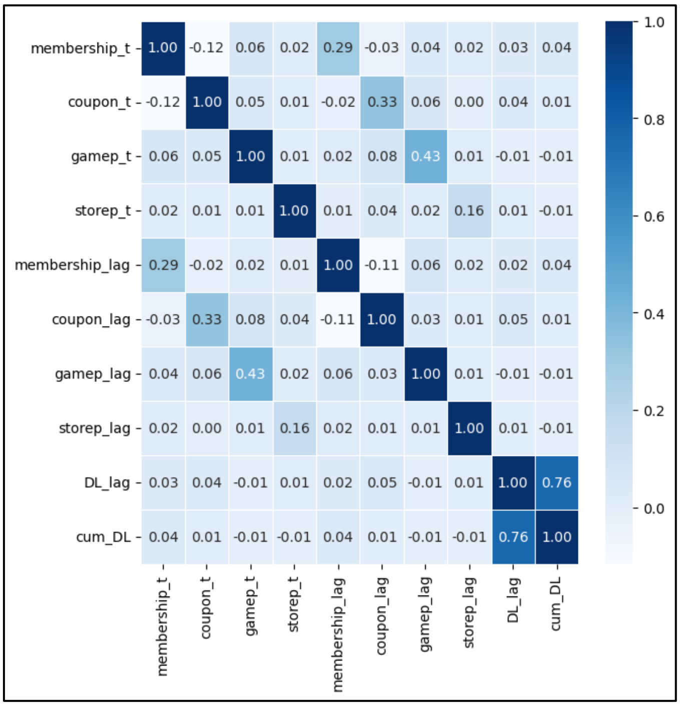


**Figure 2**. Correlation Heat Map of the Download and Promotion Variables.

## 5.2. Comparison of the Fixed Effect Models

In this section, 2-stage models with game-level fixed effects will also be included in the models. Table 4 compares the fixed effect models with game-level or add-on level fixed effects using data with or without non-purchased data where sales quantity is zero. The price coefficient alpha is larger in non-purchased data for all models since the variation will be larger if the zero data is included in the model.

The estimates of the models using the game-level fixed effect did not change much, and the fit of the model is almost the same. The price coefficients were still low and insignificant for the OLS models with the game-fixed effects. The estimated alpha for the FOC models with add-on level fixed effects is large, as in section 5.1. The adjusted R-squared of stage 1 is larger than 0.7, but the model fit for stage 2 is very small.

The 2-stage model with game-fixed effects had much less model fit for stage 1, but a much larger fit for stage 2 than the add-on fixed effect model. Considering the structure of equation 8, the stage 2 price equation for the FOC models, ruling out game-level variation is more important since the time-variant variables are averaged by j-level. The FOC model with game-specific fixed effects can eliminate the game-level fixed effect of stage 2. However, the structure of stage 1 is expressed in equation 9, which is the jt level equation. Since the add-ons are nested into games, including add-on level fixed effects will explain more of the unobserved characteristics of the games and add-ons. The add-on fixed effects in the 2-stage model capture the add-on level heterogeneity and some of the game-specific variations when the game-fixed effects can only capture the shared game level characteristics. The OLS model with game-level fixed effect has difficulty capturing the add-on level variations compared to the FOC game fixed effect model.

**Table 4.** Fixed Effect Model Comparison – Non-Purchased and Only Purchased

| | | Non-Purchased | | | | | Only Purchased | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | OLS | FOC Models | | | | OLS | FOC Models | | | |
| | | Game FE | Game FE | | Add-on FE | | Game FE | Game FE | | Add-on FE | |
| | **(Intercept)** | **0.62** | **3.25** *** | | **252.39** *** | | **-0.37** *** | **4.76** *** | | **254.43** *** | |
| | | (0.924) | (0.860) | | (0.000) | | (0.977) | (0.837) | | (0.000) | |
| $\alpha$ | $1/\alpha$ | | **0.4659** | | **0.1627** | | | **0.4880** | | **0.1961** | |
| | **Price** | **0.0130** | alpha = **2.1464** | | alpha = **6.1463** | | **0.0104** | alpha = **2.0492** | | alpha = **5.0994** | |
| | | (0.188) | (=1/0.4659) | | (=1/0.1627)) | | (0.391) | (=1/0.4880) | | (=1/0.1961) | |
| $\gamma_2$ | **DL_lag** | **0.22** *** | **0.23** *** | | **0.25** *** | | **0.26** *** | **0.27** *** | | **0.29** *** | |
| | | (0.000) | (0.000) | | (0.000) | | (0.000) | (0.000) | | (0.000) | |
| | **q0_lag** | **-6.57** *** | **-3.40** ** | | **1.99** * | | **-5.57**. *** | **-2.52** | | **2.42** * | |
| | | (0.000) | (0.008) | | (0.017) | | (0.000) | (0.171) | | (0.041) | |
| | **membership_lag** | **-23.27** | **-12.47** | | **-8.08** | | **-10.44** | **-8.08** | | **-7.60** | |
| | | (0.154) | (0.482) | | (0.472) | | (0.613) | (0.721) | | (0.591) | |
| | **coupon_lag** | **1.21** | **2.12** | | **-7.33** ** | | **4.67** | **1.05** | | **-14.86** *** | |
| | | (0.737) | (0.597) | | (0.005) | | (0.314) | (0.841) | | (0.000) | |
| | **gamep_lag** | **15.94** *** | **18.39** *** | | **11.39** *** | | **22.10** *** | **24.11** *** | | **14.57** *** | |
| | | (0.000) | (0.000) | | (0.000) | | (0.000) | (0.000) | | (0.000) | |
| | **storep_lag** | **-16.45** | **-5.75** | | **4.62** | | **-21.35** | **-8.22** | | **8.81** | |
| | | (0.457) | (0.808) | | (0.756) | | (0.467) | (0.797) | | (0.658) | |
| | **membership** | **43.29** ** | **41.74** ** | | **22.43** * | | **-3.62** | **-1.81** | | **0.04** | |
| | | (0.002) | (0.007) | | (0.023) | | (0.835) | (0.926) | | (0.997) | |
| | **coupon** | **19.86** *** | **19.42** *** | | **7.24** ** | | **8.71** * | **9.32** . | | **-0.58** | |
| | | (0.000) | (0.000) | | (0.005) | | (0.040) | (0.054) | | (0.853) | |
| | **gamep** | **24.51** *** | **25.18** *** | | **13.67** *** | | **20.09** *** | **22.60** *** | | **11.92** *** | |
| | | (0.000) | (0.000) | | (0.000) | | (0.000) | (0.000) | | (0.000) | |
| | **storep** | **-18.77** | **-9.77** | | **0.25** | | **-21.30** | **-12.82** | | **-0.88** | |
| | | (0.416) | (0.689) | | (0.987) | | (0.405) | (0.642) | | (0.959) | |
| $\beta_{FE}$ | **mean** | 40.31 | 35.72 | | 34.10 | | 46.54 | 43.47 | | 38.71 | |
| | **std** | 59.30 | 60.24 | | 60.24 | | 62.56 | 63.38 | | 63.38 | |
| | **min** | -6.47 | -3.24 | | -4.87 | | -7.76 | -2.45 | | -7.21 | |
| | **median** | 20.74 | 14.59 | | 12.96 | | 25.75 | 22.60 | | 17.84 | |
| | **max** | 795.12 | 792.74 | | 791.11 | | 798.69 | 796.59 | | 791.83 | |
| | | | **Step1** | **Step2** | **Step1** | **Step2** | | **Step1** | **Step2** | **Step1** | **Step2** |
| | **Adj. R-squared** | 0.252 | 0.249 | 0.153 | 0.718 | 0.030 | 0.253 | 0.251 | 0.151 | 0.729 | 0.051 |
| | **-2LL** | 3.207 E+05 | 3.209 E+05 | 4.328 E+04 | 2.883 E+05 | 4.362 E+04 | 2.626 E+05 | 2.626 E+05 | 4.307 E+04 | 2.347 E+05 | 4.353 E+04 |
| | **AIC** | 3.212 E+05 | 3.213 E+05 | 4.328 E+04 | 2.965 E+05 | 4.362 E+04 | 2.630 E+05 | 2.631 E+05 | 4.308 E+04 | 2.429 E+05 | 4.353 E+04 |
| | **BIC** | 3.232 E+05 | 3.233 E+05 | 4.329 E+04 | 3.304 E+05 | 4.363 E+04 | 2.649 E+05 | 2.650 E+05 | 4.308 E+04 | 2.759 E+05 | 4.354 E+04 |

**Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

The estimated alpha is smaller for the models with game-fixed effects, and the models with add-on fixed effects had three times larger price coefficients. For the FOC models, game-level or add-on-level fixed effects have their strengths and drawbacks. Therefore, determining the level of fixed effect is crucial in this 2-stage model.

The $\beta_{FE}$ is the sum of the intercept and the fixed effects. The range of the fixed effects was similar across the models. The fixed effects were between -10 to 800 for all the models. The standard error of the fixed effects was around 60.

## 5.3. Comparison of Models with Popular Games.

Table 5 compares the FOC models with add-on level fixed effects using all data and top 100 data. The top 100 products included 2695 add-ons in the data. The price coefficient of only purchased data is the same for the two datasets, but the price coefficient of the top 100 data containing zero was larger than that of all data. The other statistics were similar between the two data sets. The FOC add-on fixed effect models with top 100 data estimated larger fixed effects. The range of the fixed effects increased. The top 100 data will be used in the next session.

**Table 5.** Model Comparison with FOC j-level – Top Selling Apps

| | | All Data | | | | TOP 100 Games | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Non-Purchased | | Only Purchased | | Non-Purchased | | Only Purchased | |
| | **(Intercept)** | **252.39** | *** | **254.43** | *** | **254.19** | *** | **267.87** | *** |
| | | (0.000) | | (0.000) | | (0.000) | | (0.000) | |
| $\alpha$ | $1/\alpha$ | **0.1627** | | **0.1961** | | **0.1514** | | **0.1961** | |
| | **Price** | **alpha = 6.1463** (=1/0.1627)) | | **alpha = 5.0994** (=1/0.1961) | | **alpha = 6.6050** (=1/0.1514) | | **alpha = 5.0994** (=1/0.1961) | |
| $\gamma_2$ | **DL_lag** | **0.25** | *** | **0.29** | *** | **0.24** | *** | **0.29** | *** |
| | | (0.000) | | (0.000) | | (0.000) | | (0.000) | |
| | **q0_lag** | **1.99** | * | **2.42** | * | **2.59** | * | **3.60** | * |
| | | (0.017) | | (0.041) | | (0.047) | | (0.044) | |
| | **membership_lag** | **-8.08** | | **-7.60** | | **-13.86** | | **-15.27** | |
| | | (0.472) | | (0.591) | | (0.435) | | (0.472) | |
| | **coupon_lag** | **-7.33** | ** | **-14.86** | *** | **-10.58** | ** | **-20.61** | *** |
| | | (0.005) | | (0.000) | | (0.009) | | (0.000) | |
| | **gamep_lag** | **11.39** | *** | **14.57** | *** | **18.55** | *** | **23.27** | *** |
| | | (0.000) | | (0.000) | | (0.000) | | (0.000) | |
| | **storep_lag** | **4.62** | | **8.81** | | **9.03** | | **17.47** | |
| | | (0.756) | | (0.658) | | (0.733) | | (0.594) | |
| | **membership** | **22.43** | * | **0.04** | | **32.82** | * | **-3.15** | |
| | | (0.023) | | (0.997) | | (0.030) | | (0.865) | |
| | **coupon** | **7.24** | ** | **-0.58** | | **12.02** | ** | **1.23** | |
| | | (0.005) | | (0.853) | | (0.002) | | (0.790) | |
| | **gamep** | **13.67** | *** | **11.92** | *** | **22.05** | *** | **19.01** | *** |
| | | (0.000) | | (0.000) | | (0.000) | | (0.000) | |
| | **storep** | **0.25** | | **-0.88** | | **1.12** | | **-4.13** | |
| | | (0.987) | | (0.959) | | (0.969) | | (0.892) | |
| $\beta_{FE}$ | **mean** | 34.10 | | 38.71 | | 50.16 | | 63.54 | |
| | **std** | 60.24 | | 63.38 | | 126.40 | | 132.18 | |
| | **min** | -4.87 | | -7.21 | | -20.13 | | -42.66 | |
| | **median** | 12.96 | | 17.84 | | 17.26 | | 33.71 | |
| | **max** | 791.11 | | 791.83 | | 2998.0 | | 3005.7 | |
| | | **Step1** | **Step2** | **Step1** | **Step2** | **Step1** | **Step2** | **Step1** | **Step2** |
| | **Adj. R-squared** | 0.718 | 0.030 | 0.729 | 0.051 | 0.714 | 0.038 | 0.728 | 0.080 |
| | **-2LL** | 2.883 E+05 | 4.362 E+04 | 2.347 E+05 | 4.353 E+04 | 2.041 E+05 | 2.885 E+04 | 1.733 E+05 | 2.873 E+04 |
| | **AIC** | 2.965 | 4.362 | 2.429 | 4.353 | 2.095 | 2.885 | 1.788 | 2.873 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | E+05 | E+04 | E+05 | E+04 | E+05 | E+04 | E+05 | E+04 |
| **BIC** | 3.304 E+05 | 4.363 E+04 | 2.759 E+05 | 4.354 E+04 | 2.310 E+05 | 2.886 E+04 | 1.997 E+05 | 2.874 E+04 |

**Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

**Table 5 (continued)**

# VI. Bayesian Estimation

## 6.1. Bayesian Structure for Genre-Specific Price Coefficients

Bayesian inference methods are widely used in marketing literature and are especially helpful when the data has a large number of unknown factors or when the data is aggregated from various sources (Rossi & Allenby, 2003). In Part 6, Bayesian inference methods are used to specify the multi-level heterogeneity structure of the demand model.

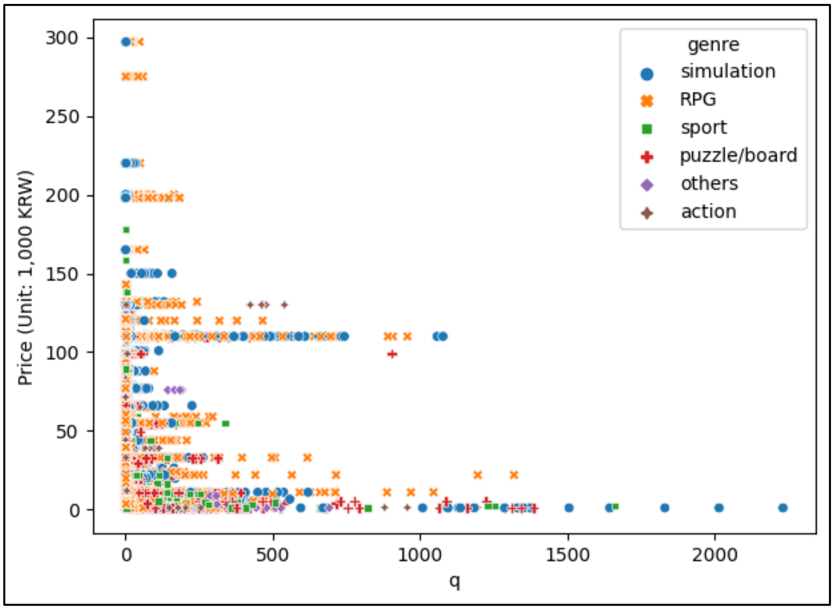


**Figure 3.** Scatter Plot of Price and Sales Quantity of Add-ons by Genre.

Figure 3 shows the relationship between the price and sales quantity of all add-ons by genre. From Figure 3, the distribution of each genre looks different, and the price response may differ between the genres. In this section, the Bayesian structure can be used to measure the different levels of price coefficients among the genres.

The Bayesian structure of genre-specific price coefficients in stage 2 of the FOC model is expressed in Figure 4 below. The prior of the alpha inverse is considered to follow a normal distribution for each genre. The sample size of action is small compared to the other genres, and action will be included as others. The distribution of alpha inverse differs among the genres. The proxy in Figure 4 is obtained by equation 10, using the estimates of stage 1. The alpha coefficient is measured based on the proxy and the genre.

$$P_j = \frac{1}{2 \cdot \alpha} \cdot \left(\beta_j + \overline{\beta_t} + \gamma_2 \cdot \overline{X_{J.}}\right) + \psi_j = \frac{1}{\alpha} \cdot proxy, \tag{10}$$
$$where, proxy = \left(\frac{1}{2} \cdot \left(\beta_j + \overline{\beta_t} + \gamma_2 \cdot \overline{X_{J.}}\right)\right)$$

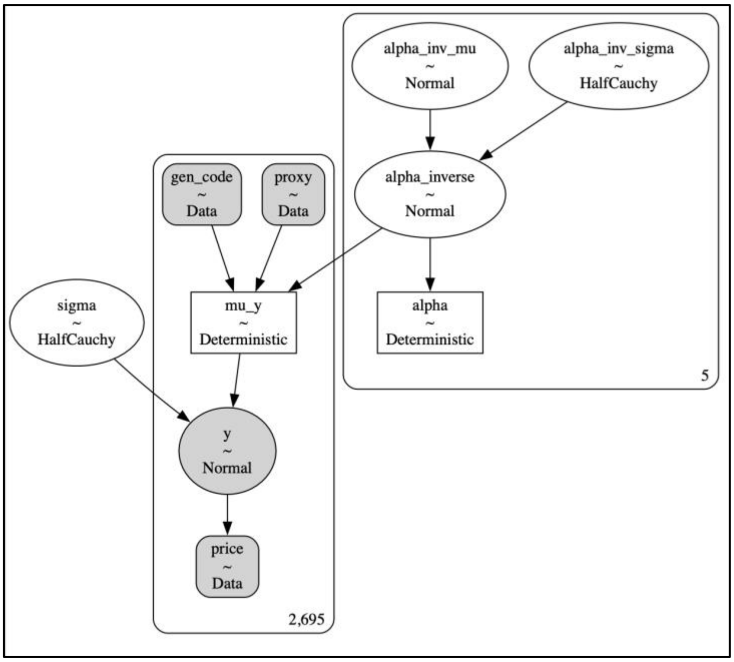


**Figure 4**. Bayesian Structure of Genre-Specific Price Coefficients in Stage 2

## 6.2. Sampling the Genre-Specific Price Coefficients

Top 100 data, including non-purchase options, are used for the estimation. For the proxy from stage 1, the models' estimates with add-on fixed effects are used. The NUTS (No U-Turn Sampling) method was used in the sampling process, which is one of the MCMC (Markov Chain Monte Carlo) sampling methods. Two chains with 1,000 samples and 1,000 tuning samples were used for the sampling process. The result of the estimated distribution of alpha inverse for each genre. The prior distribution for alpha inverse was normal, and the posterior was inferred in Figure 5.

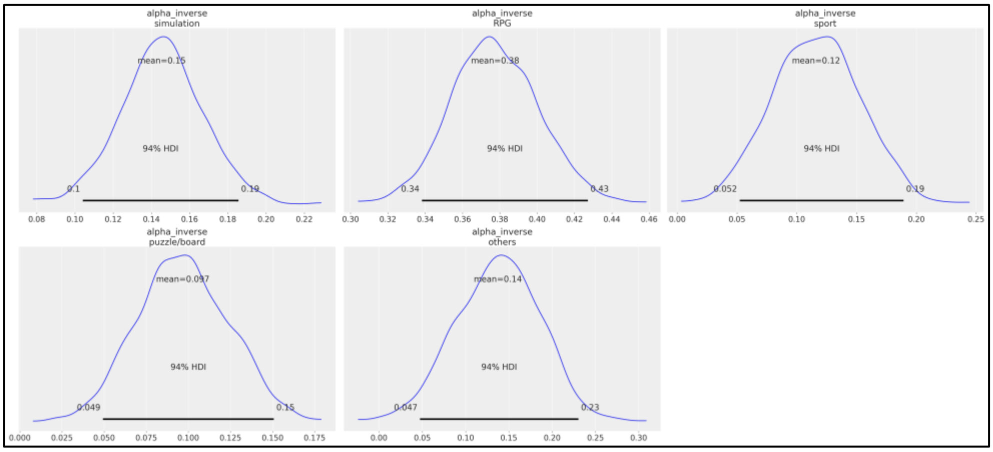


**Figure 5**. Posterior Plot of the Inverse of Genre-Specific Price Coefficients

Table 6 shows the posterior of alpha. The $\hat{R}$ is the of all alpha is less than 1.01 for all genres. The price coefficient of the RPG genre is the lowest, with a price coefficient of 2.66. The price coefficient of the puzzle/board is the highest, with a value of 11.52.

**Table 6**. Posterior of alpha

| | mean | sd | hdi_3% | hdi_97% | mcse mean | mcse_sd | ess_bulk | ess_tail | r_hat |
|---|---|---|---|---|---|---|---|---|---|
| alpha[simulation] | 7.04 | 1.12 | 5.09 | 9.10 | 0.03 | 0.02 | 2125.82 | 1422.48 | 1.00 |
| alpha[RPG] | 2.66 | 0.17 | 2.34 | 2.95 | 0.00 | 0.00 | 1412.32 | 1218.17 | 1.00 |
| alpha[sport] | 10.07 | 8.48 | 4.92 | 17.65 | 0.21 | 0.15 | 1366.03 | 948.56 | 1.00 |
| alpha[puzzle/board] | 11.52 | 5.35 | 6.08 | 18.43 | 0.17 | 0.12 | 1399.70 | 1267.25 | 1.00 |
| alpha[others] | 8.66 | 27.36 | 3.48 | 17.24 | 0.68 | 0.48 | 1834.58 | 1352.34 | 1.00 |

# VII. Add-on Price Elasticities

Since the demand model in this paper is linear, price coefficients are not price elasticities. Price elasticities can be simulated using linear models. The equation of point price elasticity is as equation 11 below. As the price changes by $\Delta P$, the demand quantity changes by $\Delta Q$. According to the linear demand model, the price coefficient can be considered as $\frac{\Delta Q}{\Delta P}$. The price coefficient in the linear demand model means that when the price changes by 1 unit, the sales quantity changes by alpha in the opposite direction. The price elasticities from the price coefficient of the linear demand models are usually calculated at the point of average price and average quantity for convenience.

$$\eta = \frac{\Delta Q}{\Delta P} \cdot \frac{\bar{P}}{\bar{Q}} = -\alpha \cdot \frac{\bar{P}}{\bar{Q}} , \tag{11}$$

Demand models are usually calculated at the point of average price and average quantity for convenience. We can calculate the price equations using the equation above. The price elasticities can be calculated from the estimated alphas in the models using the average price and quantity of the data. Table 7 shows the price elasticities of the estimated models. The models without supply-side assumptions estimated that the users are inelastic to the price of the add-ons. However, the models with the supply-side demand estimated that the users are elastic to the price. FOC models, including add-on level fixed effects, estimated much larger price coefficients than the models using game-level fixed effects.

Comparing the elasticities of the Bayesian model, the RPG game is less elastic to price than the other games, with a price elasticity of 4.17. others had the highest price elasticities among the genres. Other than others, the sports genre had the highest price elasticity of 8.08. Table A1 in the Appendix shows the conversion rates of each game genre. The sports had the largest conversion rate of 18.60 in Appstore. Considering the conversion rates, the games with high price elasticities showed higher conversion rates in the Appstore but not in Google Play Store.

**Table 7.** Estimated Price Elasticities of the Models

| | | | Non-Purchased | | | |
|---|---|---|---|---|---|---|
| | | | $\overline{P_j}$ | $\overline{q_{jt}}$ | $\alpha$ | $\eta$ |
| All Games | Fixed Effect Models | No FE | 32.68 | 20.43 | 0.0089 | -0.0142 |
| | | g-level FE | 32.68 | 20.43 | 0.0130 | -0.0208 |
| | | FOC g-level FE | 32.68 | 20.43 | 2.1464 | -3.4337 |
| | | FOC j-level FE | 32.68 | 20.43 | 6.1463 | -9.8324 |
| TOP 100 | Fixed Effect Models | No FE | 33.50 | 28.34 | 0.0100 | -0.0118 |
| | | g-level FE | 33.50 | 28.34 | 0.0163 | -0.0193 |
| | | FOC j-level FE | 33.50 | 28.34 | 6.6050 | -7.8087 |
| | Bayesian Model | RPG | 37.57 | 23.95 | 2.6599 | -4.1723 |
| | | Simulation | 39.25 | 35.61 | 7.0401 | -7.7609 |
| | | Puzzle/Board | 23.49 | 39.66 | 11.5162 | -6.8191 |
| | | Sports | 22.47 | 28.01 | 10.0719 | -8.0806 |
| | | Others | 34.63 | 30.81 | 8.6606 | -9.7344 |

# VIII. Conclusion and Discussion

The in-app purchase format of the freemium strategy is prevalent in the mobile game industry. The proportion of paid base apps is less than 10%. [④] This paper measured freemium mobile game users' responses to the price using demand and supply side structures. Since the information about the game and add-ons was limited, the models adopted fixed effect measures to separate the categorical heterogeneity from the residuals. Due to the structure of freemium apps, the game users get information about the add-ons by directly experiencing the game. Therefore, the sales of the add-ons

[④] The link below is a figure that shows the proportion of paid apps from Statista. According to this figure, 3.3% of apps are paid apps in Google Play, and 7.3% of apps in iOS Appstore are paid apps.
https://infogram.com/global-share-of-free-apps-vs-paid-apps-1h7j4dv9m1eq94n

will depend on the user's perceived value of the games and add-ons. Including the supply-side assumption increased the R-squared of the models. For the FOC models, the models with the add-on fixed effects had a better fit in stage 1, which means that it well absorbed the add-on level variations. The models with game-fixed effects had a better model fit of stage 2, since it measures the game-level variations well. Therefore, deciding the level of the fixed effects is crucial, considering the level of variations contained in the data.

There are some suggestions for future research. First, the monetization structure of the free version can be included to improve the model in this paper. Apps using freemium strategies are prevalent, and considering the growth of the mobile application market, profit from the users' engagement can be a crucial part of the freemium firms' profit structure. Therefore, the profit model in this paper can be improved by including usage per download data in the model or by implying profit per download in the profit function in this paper. Monetization from the continuous engagement of the application users is possible through in-app advertising (Rutz et al., 2019). The average revenue per download of mobile game apps in 2018 is organized by genre as Table C2 (Statista, 2022a). Including this expected value in the firms' profit function would give additional insight by improving the FOC model in this paper.

Second, variables to capture the carryover effects of the promotion variables can be added to improve the model in this paper. Promotions have a long-term impact on sales. According to Kappe et al. (2014), models containing only a finite number of lagged variables may not correctly reflect the carryover effect. Since the number of monthly observations per add-on is small for the data used in this study, only one lagged promotion variable was used for each promotion category in this paper. The model with the one-period lagged variable is flexible but may contain the danger of serial correlation and entangled carryover effects. Therefore, including more lagged variables structure in the model, using data with more observations (e.g., Koyck distributed lag model), will improve the proposed model in this paper.

Lastly, the firms' decision on sample product quality can be included in the model. Product quality is very important factor in determining the purchase occasion of the product. Consumers judge the quality of the paid version based on the quality of the free version app since the app developers provide free apps that have similar quality to the paid version with less functionality (Liu et al., 2014). Consumers also judge the quality of the paid version app by the online reviews of the app (Liu et al., 2014).

# References


Ajmera, S., Bonar, T., Scott, D., Miu, C., & Manuel, A. (2022). Market Segmentation and Recency Frequency Monetary Value Analysis for a Freemium Mobile Game. *SMU Data Science Review*, 6(2), 1.

AppsFlyer. (November 29, 2022). Retention rate on day 30 of mobile app installs worldwide in 3rd quarter 2022, by category [Graph]. In *Statista*. Retrieved January 16, 2023, from https://www.statista.com/statistics/259329/ios-and-android-app-user-retention-rate/.

Balazsi, L., Matyas, L., & Wansbeek, T. (2018). The estimation of multidimensional fixed effects panel data models. *Econometric Reviews*, *37*(3), 212-227.

Bawa, K., & Shoemaker, R. (2004). The effects of free sample promotions on incremental brand sales. *Marketing Science*, *23*(3), 345-363.

Business of Apps (March 24, 2021). Mobile gaming industry statistics and trends for 2021. In *Business of Apps*. Retrieved October 16, 2022, from https://www.businessofapps.com/insights/mobile-gaming-industry-statistics-and-trends-for-2021/.

Deng, Y., Lambrecht, A., & Liu, Y. (2022). Spillover effects and freemium strategy in the mobile app market. *Management Science*.

DIW. (February 5, 2021). Conversion rate of selected in-app advertising formats from January to October 2020, by platform [Graph]. In *Statista*. Retrieved January 16, 2023, from https://www.statista.com/statistics/1265304/conversion-rate-in-app-ads-platform/.

Dobrilova. (2022), “23+ Mobile Gaming Statistics for 2022 - Insights Into a $76B Games Market,” https://techjury.net/blog/mobile-gaming-statistics/#gref.

Dobrilova. (2023), “23+ Mobile Gaming Statistics for 2023 - Insights Into a $175B Games Market,” https://techjury.net/blog/mobile-gaming-statistics/#gref.

Gardner, D. M. (1971). Is there a generalized price-quality relationship? *Journal of Marketing Research*, *8*(2), 241-243.

Gentzkow, M. (2007). Valuing new goods in a model with complementarity: Online newspapers. *American Economic Review*, 97(3), 713-744.

Gu, Xian, P. K. Kannan, and Liye Ma (2018), “Selling the premium in freemium.” *Journal of Marketing*, 82(6), 10-27.

Hausman, J. A., & Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica: Journal of the Econometric society*, 1377-1398.

Hsiao, K. L., & Chen, C. 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..

Holm, A. B., & Günzel-Jensen, F. (2017). Succeeding with freemium: strategies for implementation. *Journal of Business Strategy*.

Jang, M., Lee, R., & Yoo, B. (2021). Does fun or freebie increase in-app purchase?. *Information Systems and e-Business Management*, *19*(2), 439-457.

Jang, S., & Chung, J. (2021). What drives add-on sales in mobile games? The role of inter-price relationship and product popularity. *Journal of Business Research*, *124*, 59-68.

PRNewswire. (2022). Korean App Market “ONE store” Announces Its Entry into the Global Market This Year. (2022, April 15). CISION PR Newswire.

Retrieved December 15, 2022, from https://www.prnewswire.com/news-releases/korean-app-market-one-store-announces-its-entry-into-the-global-market-this-year-301526505.html

Kappe, E., Stadler Blank, A., & Desarbo, W. S. (2014). A general multiple distributed lag framework for estimating the dynamic effects of promotions. *Management Science*, *60*(6), 1489-1510.

Kumar, V. (2014). Making” freemium” work. *Harvard business review*, 92(5), 27-29.

Li, H., Jain, S., & Kannan, P. K. (2019). Optimal design of free samples for digital products and services. *Journal of Marketing Research*, 56(3), 419-438.

Liu, C. Z., Au, Y. A., & Choi, H. S. (2014). Effects of freemium strategy in the mobile app market: An empirical study of google play. *Journal of management information systems*, 31(3), 326-354.

McConnell, J. D. (1968). The price-quality relationship in an experimental setting. *Journal of Marketing Research*, 5(3), 300-303.

Monroe, K. B. (1973). Buyers’ subjective perceptions of price. *Journal of marketing research*, *10*(1), 70-80.

Purnami, L. D., & Agus, A. A. (2020, September). The effect of perceived value and mobile game loyalty on mobile game’s in-app purchase intention. In *2020 3rd International Conference on Computer and Informatics Engineering (IC2IE)* (pp. 224-229). IEEE.

Rossi, P. E., & Allenby, G. M. (2003). Bayesian statistics and marketing. *Marketing Science*, *22*(3), 304-328.

Rutz, O., Aravindakshan, A., & Rubel, O. (2019). Measuring and forecasting mobile game app engagement. *International Journal of Research in Marketing*, *36*(2), 185-199.

Seetharaman, P. B. (2004). Modeling multiple sources of state dependence in random utility models: A distributed lag approach. *Marketing Science*, *23*(2), 263-271.

Statista. (2022a). Mobile gaming market worldwide. In *Statista.* Retrieved January 16, 2023, from https://www.statista.com/outlook/dmo/app/games/worldwide.

Statista. (2022b). Number of digital video game users worldwide from 2017 to 2027, by segment (in millions) [Graph]. In *Statista*. Retrieved January 16, 2023, from https://www.statista.com/forecasts/456610/video-games-users-in-the-world-forecast.

Sur, S. (August 11, 2022). Average App Conversion Rate per Category [2022]. In *Apptweak*. Retrieved January 16, 2023, from https://www.apptweak.com/en/aso-blog/average-app-conversion-rate-per-category-2019.

Tafradzhiyski, N. (November 3, 2022). In-App Purchases. *Business of Apps*. Retrieved January 16, 2023, from https://www.businessofapps.com/guide/in-app-purchases/.

Thomadsen, R. (2005). The effect of ownership structure on prices in geographically differentiated industries. *RAND Journal of Economics*, 908-929.

Wijman, Tom (May 8, 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.

https://newzoo.com/insights/articles/newzoo-games-market-numbers-revenues-and-audience-2020-2023/

Yang, Y., & Schmidt, P. (2021). An econometric approach to the estimation of multi-level models. *Journal of Econometrics*, *220*(2), 532-543.

# 국문 초록

프리미엄(Freemium) 전략은 서비스나 제품의 일부를 무료로 제공하고 유료 옵션을 구매하도록 장려하여 유저 기반을 확장시키는 판매전략으로, ‘무료’와 ‘프리미엄(premium)’옵션을 모두 포함한다 (Kumar, 2014; Liu et al., 2014; Gu et al., 2018). 기본 앱을 무료로 다운로드 할 수 있도록 샘플로 제공하고 주로 인앱 결제(In-App Purchase)를 통해 유료 애드온(add-on)을 판매하는 전략은 모바일 앱 시장에서 널리 사용되고 있는 프리미엄 전략이다.

본 연구는 앱스토어에서 제공받은 모바일 게임 다운로드 및 구매 내역에 관한 데이터를 활용하여 애드온(add-on) 가격에 대한 프리미엄 모바일 게임 이용자들의 반응을 실증적으로 분석한다. 데이터에 포함된 애드온 가격은 관찰된 기간동안 일정하였다. 애드온의 특성이나 품질 등 애드온에 관한 정보가 데이터에 포함되어 있지 않으므로, 카테고리 별 정보는 제한 적이다. 또한, 게임 수준에서 발생하는 데이터의 변동을 모두 통제하기에는 포함되어 있는 게임 특성이 충분하지 않으며, 애드온 수준의 변수로는 애드온 가격이 유일하다. 따라서, 게임 수준과 애드온 수준에서 발생하는 이질성 (heterogeneity)의 영향을 담을 수 있는 변수를 포함하는 것이 중요하다.

첫째, 본 연구의 목적은 프리미엄 전략을 사요하는 무료 앱의 애드온 가격에 대한 모바일 게임 이용자들의 반응을 파악하는 것이다. 둘째, 본 연구의 목적은 시간 불변의 게임 별, 애드온 별 이질성을 효율적으로 포착할 수 있는 범주형 변수를 포함하는 것이다. 애드온의 가격은 관찰된 기간동안 일정하기 때문에 선형 수요 모형에서 애드온 수준의 고정 효과(fixed-effect)를 반영하는 것은 불가능하다. 본 연구에서는 기업이 수익을 극대화 한다는 공급 측면의 가정을 도입하여 애드온 수준의 고정 효과를 포함하는 2단계 모형을

사용하였다. 게임 또는 애드온 수준의 이질성은 고정 효과를 포함하여 관찰할 수 있다. 셋째, 이 데이터의 위계형 구조를 반영하는 것이 본 논문의 목적이다. 가격 계수는 장르 별로 측정될 수 있으며, 본 논문에서 사용한 데이터는 계층적이기 때문에 베이지안 추론 방법을 사용하여 계층적 구조를 추정하였다.

**키워드 : 모바일 게임, 프리미움 전략, 가격전략, 인앱구매, 가격 탄력성, 전환율**

**학번:** 2020-25878

# Appendix A

Figure A1 is the bar graph of the retention rate by the category of mobile apps 30 days after the mobile app installations (AppsFlyer., 2022). The f rate of mobile gaming apps in 2022 is 2.4%, which is relatively low compared to the other categories of apps.

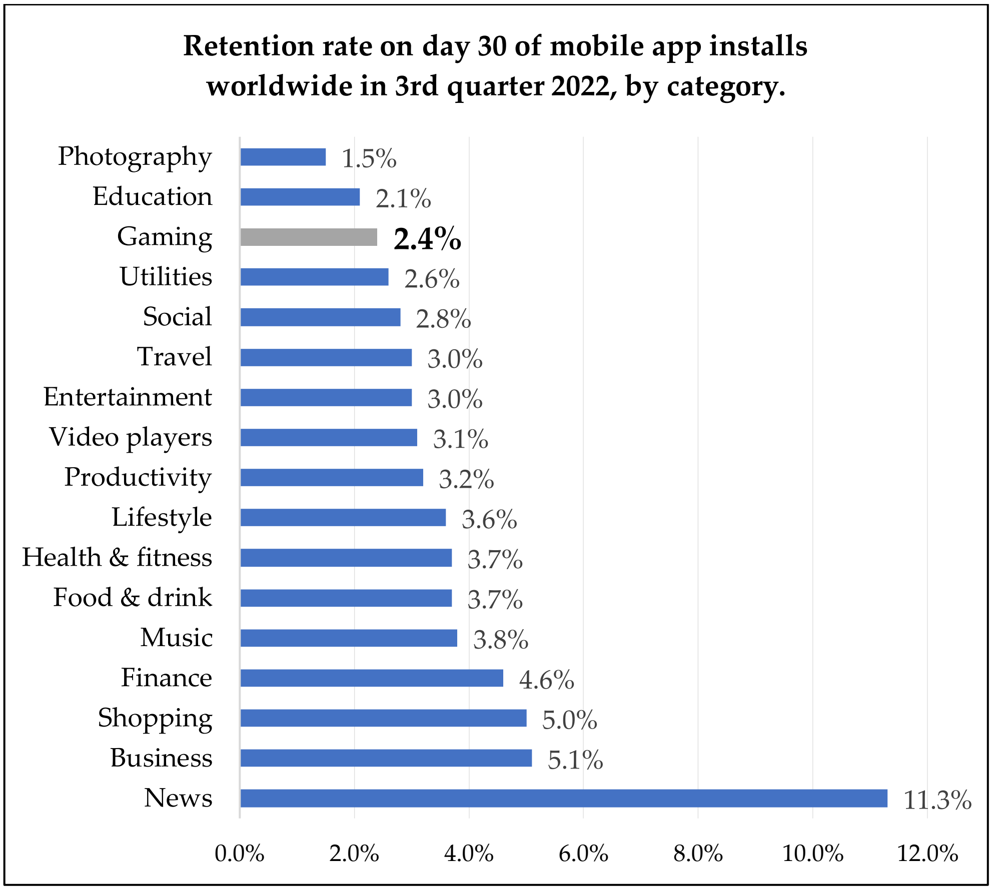


**Figure A1.** Retention rate on day 30 of mobile app installs worldwide in the 3rd quarter of 2022, by category.

*Source: Refer to AppsFlyer (2022). In Statista.*

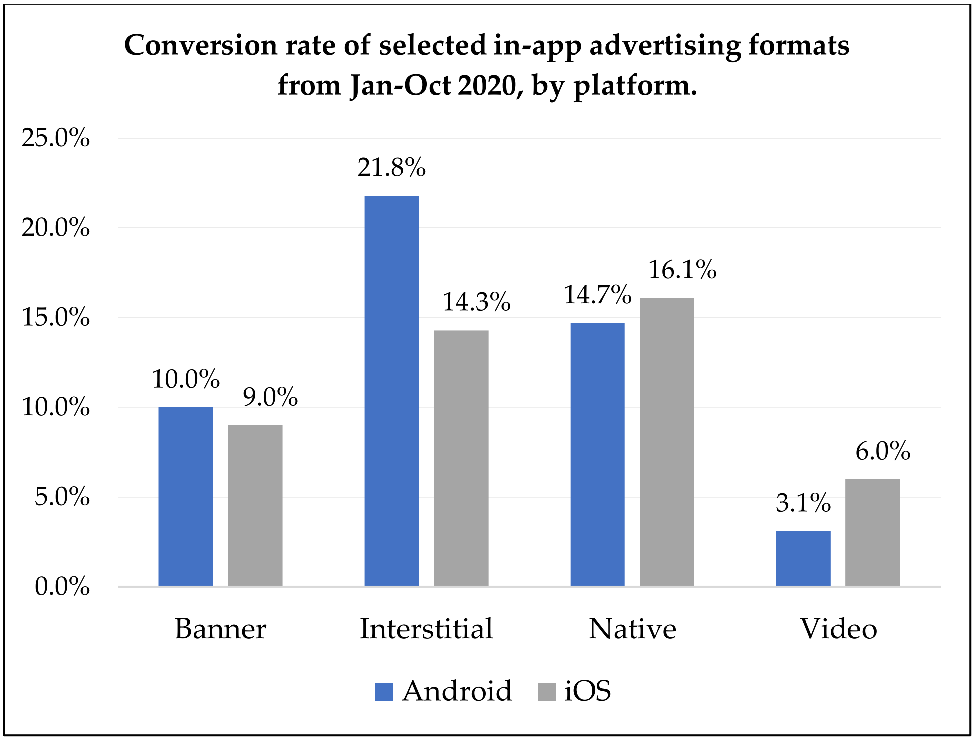


**Figure A2.** The conversion rate of selected in-app advertising formats from Jan-Oct 2020 by platform. (*DIW., February 5, 2021).*

*Source: Refer to DIW and Liftoff. (February 5, 2021) In Statista.*

According to Sur, S. (August 11, 2022), the average app conversion rate is 31% in US Appstore and 32.7% in US Google Play. Table C shows the conversion rate of the mobile game apps by category.

**Table A1. Average App Conversion Rate per Category [2022]**

| Categories | Appstore | Google Play |
|---|---|---|
| Trivia | 79.90% | 29.80% |
| Music | 55.80% | 51.20% |
| Racing | 27.70% | 28.70% |
| **Sports** | **18.60%** | **25.10%** |
| Adventure | 13.40% | 21.10% |
| **Simulation** | **12.00%** | **25.20%** |
| Casual | 9.60% | 21.00% |
| **Action** | 8.80% | 21.70% |
| Strategy | 8.50% | 14.50% |
| Games (Arcade) | 8.00% | 22.40% |
| **RPG** | **6.50%** | **26.40%** |
| Casino | 6.50% | 20.60% |
| Family | 5.30% | 41.80% |
| Word | 5.00% | 37.50% |
| **Puzzle** | **4.10%** | **19.90%** |
| **Board** | **4.10%** | **25.90%** |
| Card | 2.90% | 26.70% |

*Source: Sur, S. (August 11, 2022). https://www.apptweak.com/en.*