An Empirical Analysis of Consumption Patterns for Mobile Apps and Web: A Multiple Discrete-Continuous Extreme Value Approach

Completed Research Paper

Sang Pil Han
Arizona State University
W.P. Carey School of Business
Tempe, Arizona
shan73@asu.edu

Sungho Park
Arizona State University
W.P. Carey School of Business
Tempe, Arizona
Sungho.Park.1@asu.edu

Wonseok Oh
Korea Advanced Institute
of Science and Technology
85 Hoegi-ro, Dongdaemoon gu,
Seoul, Korea
wonseok.oh@kaist.ac.kr

Abstract

Using a unique panel data set detailing individual-level mobile app consumption, this study develops a utility theory-based structural model for multiple discrete/continuous choices in app use. We identify the dynamics and inter-dependencies between mobile apps and jointly quantify consumers' app choice and satiation simultaneously. The results suggest that mobile users' baseline utility is the highest for communication apps, while the lowest for personal financing apps. In addition, users' satiation level is the highest for the personal financing apps and the lowest for the game apps. However, a substantial heterogeneity in baseline utility and satiation is observed across diverse users. Furthermore, both positive and negative correlations exist in the baseline utility and satiation levels of mobile web and app categories. Consequently, the proposed frameworks could open new perspectives for handling large-scale, micro-level data, serving as important resources for big data analytics in general and mobile app analytics in particular.

Keywords: Mobile analytics, time use modeling, satiation, interdependence, econometrics

Introduction

The mobile revolution has fundamentally changed the way we understand data. According to Flurry, a global leader in mobile analytics solutions, the average US mobile consumer spends 2 hours and 42 minutes per day on mobile devices in 2013 and approximately 86 % of the user's mobile time (i.e., 2 hours and 19 minutes per day) is consumed on apps that run on smartphones and tablets (Khalaf, 2014). The predominance of apps in the emerging mobile paradigm has been empowered by their colossal growth and expansion. The bulk of these countless apps generate the mammoth amount of data in the form of social media exchanges, purchase transactions, music downloads, car navigations, stock investments, and search queries. In the app-based economy and social environment, every text, every transaction, every digital process, and literally every touch and move through apps can become data points.

Despite the pervasiveness of mobile apps in everyday economic and social exchanges, not all apps are created equal. Many apps (e.g., communication messengers, games, news/stock feeds) are consumed with great frequency, while many others are used only once or twice in their lifespans and then quickly disappear into the void forever. As mobile users are increasingly inundated with the massive influx of apps, they adopt choice mechanisms through which to winnow the ocean of alternatives and manage their app consumption effectively to maximize their utility given time constraints. However, little is known about the changing dynamics between app users' choice decisions and their utility maximization.

Understanding the underlying mechanisms of users' app consumption has important implications for firms' strategies and in particular mobile app monetization, mobile consumer engagement and mobile media planning. According to Flurry, between 2010 and 2012 80 percent of apps were free, but by 2013, 90 percent of apps in use were free (Gordon, 2013). As app developers and publishers migrate from paid downloads to in-app purchase and ad-supported business models, they started shedding more light on how much time their users spend in their apps. More time spent in an app is likely to generate additional in-app advertising revenues and purchases. In addition, brands use mobile apps as a communication and transaction touch-point to interact with their customers. Extra time spent in a brand's app enhances consumer engagement with the brand. Furthermore, as advertisers spend more on mobile advertising, they aim to select optimal mobile media vehicles such as mobile apps and websites to maximize their advertisement exposures and therefore to derive consumers' economic behaviors. To quote Flurry, "in the world of advertising, time-spent by consumers is the timeless currency (Khalaf, 2014)," thus it is critical for businesses in the app-based economy to understand how users determine their time use on apps.

Understanding users' app time-use is further complicated and constrained because of user heterogeneity. In particular, user demographics affect intrinsic preferences for apps and the marginal utility derived from app consumption. Since behavioral heterogeneities on IT usage have been frequently observed across diverse demographic attributes (Taylor and Todd, 1995), the choice and consumption patterns regarding apps are expected to vary across age, gender, income, and educational levels. However, no comprehensive, systematic analytics has yet to explore how demographic characteristics influence the consumption of numerous mobile apps and its implications on individual utility trajectories. Furthermore, some apps are often utilized in chorus in order to increase user experience and utility, while others replace each other, competing for volume and frequency of use in the battle of the substitutes.

Although the number of apps is increasing at an unprecedented rate, the lack of validated empirical schemes, robust computational frameworks and analytics, and actual usage data in large quantity has plagued our understanding of app usage patterns and their interdependence and competition. To explore systematically the patterns of consumption dependency for a wide selection of apps, this study develops, building on the work of Bhat (2005), a utility theory-based structural model for multiple discrete/continuous choices in app use. The underlying mechanism assumes that a consumer's marginal utility diminishes as the level of consumption of any particular app increases – a phenomenon known as *satiation*. To provide empirical insights into our utility-based choice paradigm, we use a unique panel data set detailing individual user-level app consumption. Specific research questions include: What characterizes the baseline marginal utility of a particular app category when no app is consumed? To what extent does marginal return diminish as the consumption of a particular category of apps increases? How do the baseline utility and satiation levels vary across different user demographics? Answering these questions within a rigorous theoretical framework with empirical validation can enhance our understanding of the baseline utility and satiation levels of numerous mobile apps in diverse categories.

The frameworks and methods derived in this study may have several important contributions to the emerging field of mobile analytics, which has yet to coalesce upon a clearly defined approach that incorporates simultaneously both consumer choice and utility. Unlike usual discrete choice models where a single option is chosen, the multiple discrete/continuous analytical approach allows multiple alternatives to be selected in tandem. The diminishing marginal utility provides a horizontal motivation for multiple discrete purchases. That is, if consumers face diminishing marginal utility for each of several potential apps, and the attractiveness of the apps is comparable to each other, then several alternatives will be chosen and evaluated simultaneously. We regard the diminishing marginal utility as a manifestation of satiation.

Furthermore, our approach jointly models the app choice and consumption time decisions within a single utility framework. For mobile app users, decisions regarding which apps to use and how extensively consume them are closely related since both questions are influenced by the same factors (e.g. time constraints and app developer's marketing activity like one-week free trial). Modeling the two decisions in isolation, for instance, using a multivariate Logit model for the choice decision and regression models for the quantity decisions, may not be accurate if the quantity decision is not statistically independent of the choice decision and vice versa. The separate modeling approach is prone to resulting in biased and inconsistent regression parameter estimates and inefficient Logit parameter estimates (Tellis 1988; Krishnamurthi and Raj 1988). The bias occurs because the regression models omit a relevant variable and the inefficiency in the Logit model arises because the information contained in the data on the continuous use time is ignored. Therefore, these modeling challenges reinforce the need for a new methodological paradigm for big data analytics in the context of mobile app markets where consumers face numerous choices and need to make effective consumption decisions.

One of the key contributions our approach makes to mobile analytics is that it incorporates factor analytic structures into the multiple discrete/continuous framework. The proposed structural techniques and processes enable us to estimate correlations in both baseline utilities and satiation levels of mobile app categories in a parsimonious and flexible way. In addition, the framework can be applied to modeling individual user-level time-use of various IT artifacts and their interdependence using large-scale, micro-level data. Furthermore, this modeling approach can reduce computational inefficiencies inherent in estimating unobserved heterogeneity and interdependence when the number of app categories is exceedingly overwhelming (e.g., dimensionality issue). Consequently, our methods and processes could serve as important resources for big data analytics and open new perspectives for mobile app analytics.

In addition to estimating the baseline utility and satiation related to app consumption, we examine app time-use dynamics and interdependence among apps in diverse categories. For example, will consuming social network apps increase or decrease the use of communication apps or vice versa? Recently, Facebook (SNS category) acquired WhatsApp (Communication category) in a landmark deal for $19 billion to cement its position and create a positive synergy. This corporate coupling further expands and diversifies Facebook's existing strategic app portfolio that includes, among others, Instagram (Photo category), Spaceport (Game category), and Jibbigo (Utility category). Furthermore, we provide methodological insights into how app category (dis)similarities in unobserved attributes can be estimated.

Theoretical Background

User Behavior in Mobile Platforms and Online Environments

Research on mobile platforms and apps has proliferated commensurate with their increasing use. Ghose and Han (2011) investigate user behavior on the mobile internet by mapping the interdependence between the generation and usage of mobile content. They find that a negative temporal interdependence exists: The more the mobile content is consumed in the previous period, the less the content is generated in the current period or vice versa. Ghose et al. (2013) report that the influence of ranking and geographical proximity are more pronounced on mobile devices than PCs. Recently, Einav et al. (2014) find that adoption of the mobile shopping application is associated with both an immediate and sustained increase in total platform purchasing. Further, our study builds on an emerging stream of literature on mobile apps. For example, using a reduced-form model and data from Apple's App Store, Carare (2012) reveals that app consumers are willing to pay an additional $4.50 for top ranked apps, but their preference towards the apps with bestseller status declines sharply for top ranked products. Recently, Xu et al. (2014) demonstrate that the introduction of a mobile app by a major national media company leads to an increase in demand at the corresponding mobile news website. Using data from Apple's App Store, Garg and Telang (2013) discover that top-ranked paid apps available for iPhone elicit 150 times more downloads than other apps ranked. Ghose and Han (2014) find that app demand increases with the in-app purchase option. Based on a thorough review, we found that the current literature on mobile apps focuses exclusively on demand in the form of downloads or paid purchases, but pays scant attention to the actual choices and consumption patterns of mobile apps.

In online and IT environments, technology acceptance and usage behaviors are often determined by demographic characteristics (Venkatesh and Morris, 2000). Bhatnagar and Ghose (2004) find that age, education level, and Internet experience are positively associated with the duration of their information-seeking and product search on the Internet. Hong and Tam (2006) show that male consumers are more likely to adopt mobile data services than female counterparts. Based on the analysis of individual users' social networks, Igarashi et al. (2005) discover that women have a higher tendency to expand their social networks and maintain interpersonal relationships through mobile phone text messages than men.

Multiple Discrete-Continuous Choice Models

Several analytics have been proposed to capture the dependence between choice and quantity decisions: the single utility (or structural) approach (Chintagunta 1993; Kim et al 2002) and the error-dependence (or reduced form) approach (Tellis 1988; Zhang and Krishnamurthi 2004). The former approach specifies a utility function and the optimal choice and quantity are derived as an equilibrium solution from the utility function. The dependence between choice and quantity decision is captured in the utility function. The latter approach handles the dependence by allowing correlations in error terms of the choice and quantity models. A major advantage of the single utility approach is that it allows researchers to estimate structural parameters and metrics of economic interest (e.g. compensating variation). The proposed model based on the multiple discrete/continuous approach belongs to the single utility approach. Mobile users' app usage decisions can be decomposed broadly into two elements - which apps to adopt and how extensively to consume them. Since multiple discrete-continuous choice models tackle both problems within a single utility maximization framework, they are appropriate for analyzing our mobile app and web time-use data. Kim et al. (2002) propose a translated nonlinear, addictive utility model in which a parsimonious specification provides both corner and interior solutions in the context of the simultaneous purchase of multiple varieties. The multiple discrete-continuous extreme value (MDCEV) model formulated by Bhat (2005, 2008) extends single discrete-continuous frameworks (e.g., Dubin and McFadden 1984; Chintagunta 1993) to handle multiple discreteness in demand and resolve the presence of heteroscedasticity and correlations that arise from unobserved characteristics. MDCEV models have frequently been used for analyzing time-use data. Bhat (2005) scrutinizes time-use allocation decisions among several discretionary activities on weekends. Building on Bhat's MDCEV time-use model, Spissu et al. (2009) study within-subject variation over the 12 week sample period for six broad activity categories along with another activity. Luo et al. (2013) incorporates dynamic components into a MDCEV model to examine how consumers allocate time to a portfolio of leisure activities over time.

In this study, we develop a unique structural model of app selection and time-use decision by incorporating a factor analytic structure into a MDCEV framework. The vectors of individual-level baseline utilities and satiation parameters are modeled as functions of observed mobile user characteristics and a small number of unobservable user-specific factors. In literature, factor analytic structures are combined with Probit or Logit models and the primary focus of these models is to understand inter-brand competition by pictorially depicting locations of competing brands in a perceptual map (Chintagunta 1994; Elrod and Keane 1995). Our approach offers a methodological contribution to MDCEV models by allowing correlations in both baseline utilities and satiation levels of various mobile app categories in a parsimonious manner using factor analytic approaches.

Empirical Background and Data Description

Mobile App and Web Time-Use Panel Data

We provide a brief overview of the empirical background for our data. We have gathered large-scale panel data comprising mobile app and web time-use histories provided by Nielsen KoreanClick, a market research company that specializes in consumers' Internet and mobile usage. Audience measurement gauges how many users are in an audience and how long they remain, usually in relation to television viewership (e.g., Nielsen ratings), but also in relation to increasingly traffic on websites and mobile apps. Nielsen KoreanClick maintains a panel of mobile app users with Android operating system-based devices, aged 10 to 70, selected based on stratified sampling in Korea. Android is the dominant operating system of mobile devices worldwide accounting for 67.5 percent of the global market. In Korea, nearly 93 percent of smartphones are powered by the Android platform (Yonhap News 2014). After individuals voluntarily agree to be panel members, they download and install a meter application from Nielsen KoreanClick on their mobile devices. The participating users are rewarded with points for installing the meter app and the incentive points can be accumulated and redeemed for gift cards. This app runs in the background and collects data on panel members' use of mobile apps and the mobile web even while disconnected. The meter app regularly transmits encrypted log files to a server via a secure cellular connection or Wi-Fi. In addition, we acquire individual-level information on user demographics such as age (20's or less, 30's, and 40's or over), gender (male and female), monthly income ($3,000 or less, between $3,000 and $5,000, and $5,000 or over), and education (students, high school graduates, and university graduates).

Table 1. Mobile Content Categories: Mobile Apps and Mobile Websites

Category		Count	Example
Communication	Mobile Messengers, Mobile Internet Phone, Email	247	Kakao Talk, Mypeople Messenger, GO SMS Pro, LINE, NateOn, LightSMS, Gmail, Viber, Skype
Entertainment	Book, Cartoon, Adults, Sports, Travel, Humor, Magazine	1,071	Naver Webtoon, LIVE Scores, TIViewer, T store Book, jjComics Viewer, Naver Books, Score Center
Game	Action, Adventure, Board, Puzzle, Racing, Role Playing, Shooting, Simulation, Sports	2,478	Rule The Sky, TinyFarm, Smurfs' Village, Shoot Bubble Deluxe, Hangame, 2012 Baseball Pro, Angry Birds Space, Jewels Star
Map and Navigation	Map, Navigation	337	T map, Google Maps, SeoulBus, Naver Maps, Olle Navi, Subway Navigation,
Lifestyle	Weather, News, Restaurants, Job, Health, Religion, Fashion	1,808	YTN News, MK News, SBS News, Weather, Bible, Newspapers, JobsKorea
Personal Financing	Banking, Stocks, Finance, Real Estate, Ecommerce	407	Smart Trading, M-Stock Smart, KB Star Banking, Auction Mobile, Coupang, Gmarket Mobile
Music and Radio	Radio, Music	304	Music Player, SKY Music, PlayerPro, Mnet, MyMusicOn, FM Radio, Soribada
Photo	Photo Gallery, Camera	264	Gallery, Camera, Cymera, Photo Editor, Instagram
Portal Search	Portal Site, Search Engine	83	Naver, Daum, NATE, Google Search, Junior Naver
Schedule	Scheduler, Memo, Alarm Clock	1,399	Address book, Alarm/Clock, Calendar, Memo, Polaris Office, Polaris Office, Docviewer
Social	Social Networking Service, Board, Blog, Microblog	163	Kakao Story, Facebook, Twitter, Naver Café, Daum Café, Cyworld, Naver Blog, Me2day
Utilities	Productivity, Decoration, Webhard, Widget, Firewall	2,241	Calculator, Voice Recorder, HD Browser, NDrive, Smart App Protector, Battery Widget
Video	Multimedia, Broadcasting, Movie	261	TV, Youtube, MX Video Player, SKY Movie, Afreeca TV, T-DMB, PandoraTV
Mobile Web	Websites	7,944	Naver.com, Daum.net, Nate.com, Google.co.kr, ppomppu.co.kr, dcinside.com, facebook.com
Sum		19,007	

We collected the data between March 5 and April 30 2012 (eight weeks). These data include 1,425 panel members who used mobile apps and the web throughout the sampling period. Moreover, our data incorporate individual-level, weekly information on the type, name, and duration of mobile apps used and mobile websites visited. Nielsen KoreanClick classified the mobile content into 14 categories; communication, game, map and navigation, entertainment, lifestyle, personal financing, music and radio, photo, portal, schedule and memo, social networking, utility, video, and combined mobile web activities.

Notably Google Play is a leading app market based on Android operating system. When publishing a new app or a new version of the existing app in Google Play, app developers self-select one or more appropriate app categories, however, they are not required to go through the verification process. Hence there are some cases in which app categories reported by app developers are incorrect or inconsistent. To address this issue, Nielsen KoreanClick performed a thorough, manual re-classification task to ensure that a certain app is classified to the single, primary app category. Table 1 demonstrates that user in our sample used 11,063 apps and visited 7,944 websites. Apps in the sample include not only top global apps such as Facebook, YouTube, Twitter but also top local apps such as Kakao Talk, Naver, Cyworld.

We have 11,400 (1,425 users 8 weeks) app choice occasions in our data. Column 1 of Table 2 shows that users access the mobile web most frequently (99.7%), followed by communication apps (99.3%), schedule/memo apps (98.7%), and utility apps (97.7%) while access entertainment apps least frequently (38.0%). In addition, Column 2 of Table 2 shows that that smartphone users in our sample spend an average of 1 hour and 47 minutes every day on consuming content in his or her mobile device. Daily time spent in mobile apps surpasses mobile web consumption. To be specific, Column 4 of Table 2 shows that users spend the most time (24.8%) with communication apps, such as mobile messaging, followed by mobile web (17.7%), game apps (12.1%), music and radio apps (8.9%), social network apps (6.1%), and utility apps (6.1%). In later section, we show that our findings of baseline utilities and satiation levels for app categories based on the proposed model are quite different from those reported in Tables 2.

Table 2. Choice and Time Use According to App Categories

Categories	Choice	Time Use		
		Hour	Percent	Percent excluding outside option
Outside option (other activities)	100.0%	22.23	92.6%	-
Communication	99.3%	0.44	1.8%	24.8%
Game	65.0%	0.21	0.9%	12.1%
Map/Navigation	69.6%	0.03	0.1%	1.5%
Entertainment	38.0%	0.03	0.1%	1.8%
Lifestyle	68.5%	0.03	0.1%	1.7%
Personal Financing	61.4%	0.04	0.2%	2.2%
Music/Radio	67.5%	0.16	0.7%	8.9%
Photo	90.3%	0.04	0.2%	2.4%
Portal	71.3%	0.07	0.3%	4.2%
Schedule/Memo	98.7%	0.1	0.4%	5.8%
Social Network	69.6%	0.11	0.4%	6.1%
Utility	97.7%	0.11	0.4%	6.1%
Video	77.7%	0.08	0.4%	4.8%
Web	99.7%	0.31	1.3%	17.7%
Total		24	100.00%	100.00%

Model-Free Evidence of Dependence between App Choice and Time Use Decisions

As mobile users increasingly access mobile web and various kinds of apps, their choices become interdependent. Table 3 demonstrates that approximately 98.5% of users in our data access more than four categories of mobile content in a given week, and what is more notable is that all users use at least two categories of mobile content during a given week. These descriptive findings of the joint use of multiple categories of mobile content lends support to the validity of our econometric model in which we incorporate the multiple-discrete choice into the continuous time-use decision.

Table 4. Correlation Matrix of App Choice Incidence

	Communication	Game	Map/Navigation	Entertainment	Lifestyle	Personal Financing	Music & Radio	Photo	Portal	Schedule/Memo	Social Network	Utility	Video	Web
Communication	1.00													
Game	0.08	1.00												
Map/Navigation	0.12	0.04	1.00											
Entertainment	0.06	0.15	0.11	1.00										
Lifestyle	0.11	0.09	0.20	0.15	1.00									
Personal Financing	0.09	0.04	0.17	0.06	0.15	1.00								
Music & Radio	-0.01	0.07	0.14	0.16	0.13	0.05	1.00							
Photo	0.23	0.10	0.16	0.09	0.18	0.11	0.13	1.00						
Portal	0.11	0.10	0.14	0.12	0.17	0.14	0.11	0.18	1.00					
Schedule/Memo	0.54	0.10	0.14	0.07	0.13	0.13	0.05	0.26	0.16	1.00				
Social Network	0.11	0.08	0.11	0.12	0.14	0.11	0.13	0.22	0.13	0.11	1.00			
Utility	0.41	0.13	0.17	0.10	0.16	0.15	0.08	0.26	0.16	0.41	0.11	1.00		
Video	0.14	0.14	0.14	0.13	0.16	0.09	0.14	0.22	0.16	0.16	0.12	0.18	1.00	
Web	0.00	-0.03	-0.03	-0.05	-0.03	-0.03	-0.03	-0.01	-0.03	0.00	-0.03	-0.01	-0.02	1.00

Table 5. Correlation Matrix of App Use Time

	Communication	Game	Map/Navigation	Entertainment	Lifestyle	Personal Financing	Music & Radio	Photo	Portal	Schedule/Memo	Social Network	Utility	Video	Web
Communication	1.00													
Game	0.00	1.00												
Map/Navigation	0.03	0.02	1.00											
Entertainment	0.03	0.03	0.03	1.00										
Lifestyle	0.05	0.04	0.04	-0.01	1.00									
Personal Financing	0.01	-0.01	0.02	-0.02	0.08	1.00								
Music & Radio	0.10	0.03	0.00	0.03	0.07	-0.01	1.00							
Photo	0.35	-0.01	0.00	-0.02	0.01	-0.02	0.05	1.00						
Portal	0.09	0.04	0.02	0.06	0.08	0.07	0.06	0.06	1.00					
Schedule/Memo	0.15	-0.01	0.08	0.02	0.08	0.06	0.03	0.04	0.09	1.00				
Social Network	0.24	0.00	0.00	0.04	0.02	-0.02	0.08	0.25	0.11	0.03	1.00			
Utility	0.08	0.04	0.04	0.01	0.08	0.04	0.01	0.03	0.06	0.09	0.02	1.00		
Video	0.04	0.02	0.03	0.04	0.07	0.02	0.08	0.02	0.09	0.06	0.05	0.02	1.00	
Web	0.03	-0.04	0.00	0.00	0.06	0.04	0.01	0.08	0.08	0.01	0.05	0.03	0.09	1.00

It is more challenging to make inferences on use time decision across multiple app categories using simple methods. We do not observe use time for app categories that are not selected. As Table 3 indicates, all categories of mobile content are used in only 6.4% of our weekly data. Therefore, to use simple metrics like a correlation matrix, we should aggregate the observed weekly data up to monthly or quarterly level or discard observations with zero-time use observations. To impute such incidence with non-app use, one can use a correlation matrix of app choice incidence using dummy variable which takes the value of one if the app is used and zero otherwise. Similarly, one can use a correlation matrix of app time-use quantity by imputing no incident with zero value. These approaches are inferior to the proposed approach because they do not fully use the available information. Tables 4 and 5 show the correlation matrix of app choice incidence variable and the correlation matrix of app time-use quantity variable, respectively. The observed relationships among app use incidents are mostly positive and marginal. We observe substantial positive relationships among communication, utility, and schedule/memo. However, we cannot find any substantial negative correlation. In the section of Correlation across Mobile Web and App Categories, we show that our findings based on the proposed model are drastically different from those reported in Tables 4 and 5, indicating that the simple methods can lead to misleading results in evaluation of correlation across mobile content categories. We discuss our modeling approach in detail below.

Econometric Model

In this section, we present our proposed model to estimate the baseline utility and satiation levels of different categories of mobile web and apps while allowing for user heterogeneity and cross-app use interdependence even when the number of app categories is large. We then discuss how we identify our demand system.

Proposed Model

a. Consumer Utility Function

We observe which app categories are chosen and how much time is spent on each selected app category in our data. Accordingly, we describe a mobile user's behavior by virtue of a multiple discrete/continuous process. Compared to discrete choice models (i.e. Logit or Probit models) and continuous dependent variable models, multiple discrete/continuous models can handle more effectively both discrete and continuous natures of observed data within a single utility-based framework. Because of this advantage, multiple discrete/continuous models have successfully been applied in several academic fields, including marketing, transportation, and economics (Kim et al 2002; Hendel 1999; Bhat 2008). In this paper, we extend Bhat's (2008) MDCEV framework. We specify the latent utility of mobile content usage as follows:

$$U_{ht} = \tfrac{1}{\alpha_0} \cdot \exp(\varepsilon_{h0t}) \cdot q_{h0t}^{\alpha_0} + \sum_{j=1}^{J} \tfrac{1}{\alpha_{hj}} \cdot \mu_{hjt} \cdot \{(q_{hjt}+1)^{\alpha_{hj}} - 1\}, \tag{1}$$

where $h = 1,\ldots,H$ denotes individual mobile users, $j = 0$ and $j = 1,\ldots,J$ denote an outside option (activities other than mobile content use) and mobile content use categories, respectively, and $t = 1,\ldots,T$ denotes time period (weeks). q_{hjt} $(j = 0,\ldots,J)$ is time allocated to alternative *j* by user *h* in time period *t*. This specification is referred to as "alpha-profile" in the literature (see Bhat (2008) for further details on model specification and other possible specifications). ε_{h0t} is an user-, alternative-, and time-specific random term associated with the outside option. μ_{hjt} represents the "baseline marginal utility" (the marginal utility when none is consumed) of alternative *j* in time *t* by user *h*. When a user decides which category to use first, categories with large value of μ_{hjt} have higher probabilities of being selected compared to those with small μ_{hjt}. Also, it can be interpreted as a measure of "perceived quality" because higher values of μ_{hjt} mean that the alternative confers higher levels of utility from any level of consumption, all else the same. Moreover, α_0 and α_{hj} are referred to as a satiation parameter in that it determines how the marginal utility of alternative changes as its consumption quantity increases. As α_{hj} decreases, the utility function in alternative *j* shows more concave patterns, and higher satiation occurs at a lower value of q_{hjt}. Due to this diminishing marginal utility, multiple alternatives can become comparable to each other and they will be chosen together rather than only one option is selected.

According to (Bhat 2008), U_{ht} becomes a proper utility function when $\mu_{hjt} > 0$ and $\alpha_{hj} < 1$ for *j*=1,..., *J*. To ensure that baseline utility is non-negative and satiation parameter is less than one regardless of the values of β_{hj}, λ_{hj}, and ε_{hjt}, we specify the baseline utility parameter μ_{hjt} and the satiation parameter α_{hj} as:

$$\mu_{hjt} = \exp(\beta_{hj} + \varepsilon_{hjt}), \text{for } j = 1,\ldots,J,$$
$$\alpha_{hj} = 1 - \exp(\lambda_{hj}), \text{for } j = 1,\ldots,J. \tag{2}$$

ε_{hjt} represent idiosyncratic elements in utility. Both ε_{h0t} and ε_{hjt} are known to decision makers but unknown to researchers. We assume that these follow Type-I Extreme Value distribution.

b. User Heterogeneity and Correlation Among Mobile Web and App Categories

For user- and alternative-specific β_{hj}, we specify the following factor analytic structure:

$$\beta_h = \overline{\beta} + \Pi_\beta D_h + \Gamma_\beta \psi_h + \Lambda_\beta \upsilon_h, \tag{3}$$

where a $(J \times 1)$ vector $\beta_h = [\beta_{h1}, \beta_{h2}, \ldots, \beta_{hJ}]'$ and $\overline{\beta}$ is a $(J \times 1)$ constant vector. D_h is the $(K \times 1)$ vector of observed demographic variables of user *h* and Π_β is a $(J \times K)$ coefficient matrix. Γ_β is a $(J \times F)$ factor loading matrix and ψ_h is a $(F \times 1)$ vector of orthogonal Gaussian factors ($\psi_h \sim N(0, I_F)$). Λ_β is a $(J \times J)$ diagonal matrix and υ_h is a $(J \times 1)$ vector of independent unit-variance Gaussian random variables ($\upsilon_h \sim N(0, I_J)$). Our specification decomposes individual-level heterogeneity in mobile content choice utility into three parts. The first is what the observed demographic variables explain. The second part is explained by parsimonious factors. In factor analysis literature, ψ_h is referred to as a "common factor". *F* elements in ψ_h influence all *J* elements in β_h. We can understand main characteristics of the factors by interpreting the factor loading matrix Γ_β. Note that, along with $\Pi_\beta D_h$, this common factors generate correlations in β_h. The remaining variation in β_h is explained by the last term in equation (3), $\Lambda_\beta \upsilon_h$, which is referred to as a "specific factor." Unlike ψ_h, *j*-th element in υ_h influences β_{hj} only.

The factor analytic structure is of interest in our empirical setting for a number of reasons. First, a factor model allows us to estimate category similarity in unobserved attributes (Elrod and Keane 1995). We can potentially interpret the factors as inherent mobile users' traits. Moreover, the user-specific factor estimates can be used for targeting purposes. Second, the factor model introduces correlations in the latent baseline utilities across app categories with relatively few parameters. This characteristic is particularly useful in our context to alleviate the concern for dimensionality issues inherent in estimating unobserved heterogeneity and their interdependence when the number of app categories (i.e., alternatives) is large. We can reduce the number of parameters required to estimate a full covariance matrix while remaining highly flexible and minimizing loss of information. Because of these major advantages, researchers have applied factor analytic structure in random coefficient Logit or Probit models (Elrod and Keane 1995; Singh et al. 2005; Hansen et al 2006). A key methodological contribution of our proposed model is that it extends factor analytic structure to multiple discrete/continuous models.

From equation (3), we can derive the following covariance matrix of β_h:

$$Cov(\beta_h) = \Pi_\beta \Omega_D \Pi'_\beta + \Gamma_\beta \Gamma'_\beta + \Lambda_\beta \Lambda'_\beta, \tag{4}$$

where Ω_D is a covariance matrix of D_h. The variance decomposition of equation (4) allows us to quantify the relative contribution of each part. The proportion of variation in β_{hj} explained by observed demographic variable is (*j*-th diagonal element of $\Pi_\beta \Omega_D \Pi'_\beta$)/(*j*-th diagonal element of $Cov(\beta_h)$).

Further, the value of user- and alternative-specific satiation parameter α_{hj} is determined by λ_{hj}. Similar to equation (3), we specify the following factor analytic model structure to λ_{hj}:

$$\lambda_h = \overline{\lambda} + \Pi_\lambda D_h + \Gamma_\lambda \varphi_h + \Lambda_\lambda \nu_h, \tag{5}$$

where $(J \times 1)$ vector is $\lambda_h = [\lambda_{h1}, \lambda_{h2}, \ldots, \lambda_{hJ}]'$, and $\overline{\lambda}$ is a $(J \times 1)$ constant vector. Π_λ is a $(J \times K)$ coefficient matrix. Γ_λ is a $(J \times F)$ factor loading matrix and φ_h is a $(F \times 1)$ vector of orthogonal

Gaussian factors ($\varphi_h \sim N(0, I_F)$). Λ_λ is a $(J \times J)$ diagonal matrix and ν_h is a $(J \times 1)$ vector of independent Gaussian random variables ($\nu_h \sim N(0, I_J)$). The covariance matrix of λ_h can be decomposed as follows:

$$Cov(\lambda_h) = \Pi_\lambda \, \Omega_D \, \Pi'_\lambda + \Gamma_\lambda \, \Gamma'_\lambda + \Lambda_\lambda \Lambda'_\lambda \,. \tag{6}$$

Model Estimation and Identification

We derive our demand system by applying the Kuhn-Tucker method to the latent utility of mobile content use specified in equation (1). By solving the Kuhn-Tucker conditions for constrained utility maximization, we obtain demand functions wherein a mixture of corner solutions and interior solutions are a product of the underlying utility structure. The Lagrangian for the constrained utility maximization problem is:

$$L_{ht} = \tfrac{1}{\alpha_0} \cdot \exp(\varepsilon_{h0t}) \cdot q_{h0t}^{\alpha_0} + \sum\nolimits_{j=1}^{J} \tfrac{1}{\alpha_{hj}} \cdot \mu_{hjt} \cdot \{(q_{hjt}+1)^{\alpha_{hj}} - 1\} + \delta_{ht}(Q - \sum\nolimits_{j=0}^{J} q_{hjt}), \tag{7}$$

where δ_{ht} is the Lagrange multiplier, and Q denotes total amount of time given to each mobile user (i.e. 24 hours per day). The Kuhn-Tucker first order conditions can be derived as the following:

$$\begin{aligned} &\mu_{hjt} \cdot (q_{hjt}+1)^{\alpha_{hj}-1} - \delta_{ht} = 0, \quad \text{if } \; q_{hjt} > 0, \\ &\mu_{hjt} \cdot (q_{hjt}+1)^{\alpha_{hj}-1} - \delta_{ht} < 0, \quad \text{if } \; q_{hjt} = 0, \end{aligned} \tag{8}$$

and $\exp(\varepsilon_{h0t}) \cdot q_{h0t}{}^{\alpha_0 - 1} - \delta_{ht} = 0$ since the outside option is always consumed (i.e. time used for activities other than mobile app and web uses is non-zero). We use the expression for δ_{ht} from the first-order condition for the outside option to eliminate the Lagrange multiplier form (7) and then take log in both sides, so the Kuhn-Tucker conditions for the interior and corner solutions can be written, respectively, as:

$$\begin{aligned} &\beta_{hj} + (\alpha_{hj}-1)\cdot\ln(q_{hjt}+1) + \varepsilon_{hjt} = (\alpha_0 - 1)\cdot\ln(q_{h0t}) + \varepsilon_{h0t}, \quad \text{if } \; q_{hjt} > 0, \\ &\beta_{hj} + (\alpha_{hj}-1)\cdot\ln(q_{hjt}+1) + \varepsilon_{hjt} < (\alpha_0 - 1)\cdot\ln(q_{h0t}) + \varepsilon_{h0t}, \quad \text{if } \; q_{hjt} = 0. \end{aligned} \tag{9}$$

From this Kuhn-Tucker first order condition, we derive the following probability that any *M* of the *J* alternatives is chosen (see Bhat 2008):

$$\begin{aligned} &P_{ht}(q_{h0t}, q_{h1t}, \ldots, q_{hMt}, 0, 0, \ldots, 0) \\ &= \int_{\Theta_h} M! \cdot \left(\prod\nolimits_{j=0}^{M} \frac{1-\alpha_{hj}}{q_{hjt}+\tau_j} \right) \cdot \left(\sum\nolimits_{j=0}^{M} \frac{q_{hjt}+\tau_j}{1-\alpha_{hj}} \right) \cdot \frac{\prod_{i=0}^{M} e^{V_{hit}}}{\left(\sum_{j=0}^{J} e^{V_{hjt}} \right)^{M+1}} \, dF(\Theta_h), \end{aligned} \tag{10}$$

where $V_{h0t} = (\alpha_0 - 1)\cdot\ln(q_{h0t})$, $V_{hjt} = \beta_{hj} + (\alpha_{hj}-1)\cdot\ln(q_{hjt}+1)$ for *j*=1,…,*J*, $\tau_0 = 0$, and $\tau_j = 1$ for *j*=1,…,*J*. Θ_h is a vector of common factors and specific factors in α_{hj} and β_{hj} (ψ_h, υ_h, φ_h, and ν_h) and $F(\Theta_h)$ is a joint distribution function of Θ_h. We use Monte Carlo simulation methods to calculate the probability (10) and estimate the model parameters by maximizing a likelihood function derived from equation (10) (see Keane 1993 for detail). To compute the integrals in the likelihood function, we generate random normal draws for ψ_h, υ_h, φ_h, and ν_h from their distributions, and then take averages of computed integrands. In our application, we use 1,000 random draws. The resulting estimator becomes a simulated maximum likelihood (SML) estimator. This procedure is the same as maximum likelihood except that simulated probabilities are used instead of the exact probabilities. The properties of SML (i.e. its consistency, efficiency, and asymptotic normality) have been derived by, for example, Keane (1993).

Bhat (2008) discusses the identification issues of general MDCEV models and shows the empirical identifiability of "alpha-profile" specification, which is adopted in the proposed model. A distinctive feature of our model is that it incorporates Gaussian factor analytic structure into a MDCEV model. Note that our Gaussian factor specification is the most general and widely used specification (Elrod and Keane 1995; Singh et al. 2005) but other distributional assumptions can be used. For the identification of model parameters, appropriate restrictions on factor loading matrices are required. Following Singh et al (2005) and Hansen et al (2006), we impose a standard triangular restriction on a loading matrix. This approach imposes the minimum restriction for the parameter identification. To be more specific, with two factors, one of the elements of the second column of factor loading matrix is restricted to be equal to zero. With three factors, a 3×3 submatrix of factor loading matrix is lower triangular. We estimate one-, two-, and three-factor versions of the model specified in equations (3) and (5). The log-likelihood values for the one-, two-, and three-factor models are 97,128, 97,926, and 98,040, respectively. To determine the number of factors, we use the Bayesian Information Criterion (BIC). The BIC values for the one-, two-, and three-factor models are -191,631, -192,984, and -192,979, respectively, and the two-factor model is chosen. The estimation results of one- and three-factor models are available upon request from the authors. Accordingly, we report and discuss the results for the two-factor model below.

Results

Baseline Utility and Satiation Levels for Mobile Web and App Categories

Table 6 and Table 7 show the estimates for baseline utility parameters ($\overline{\beta}$ and Π_{β}) and satiation parameters ($\overline{\lambda}$ and Π_{λ}), respectively. Results in Table 6 show that mobile users' baseline utility for communication apps is the highest and their baseline utility for personal financing apps is the lowest among the mobile web and app categories. The highest baseline utility for communication apps can be attributed to the high penetration and wide use of mobile messaging apps. For example, 93% of smartphone users in South Korea use a mobile messaging app, Kakao Talk (Frier 2013). The lowest baseline utility for personal financing apps suggests that they remain a niche market. Note that as $\overline{\beta}$ increases the baseline utility increases. The sign for $\overline{\beta}$ estimates for all types of mobile content is negative due to the relatively higher utility of outside options (e.g., mobile users spend more than 22 hours daily engaging in activities other than mobile content use), which is normalized to zero for model identification ($\beta_{h0} = 0$). We can also interpret the degree of baseline utility in terms of $\overline{\mu}$ because the sign of the estimates is reversed but the relative magnitude of estimates remains the same after exponential transformation in equation (2). Furthermore, among several demographic variables, age, gender and education explain substantial heterogeneity in baseline utilities across mobile users. For example, older users show a higher intrinsic preference for portal, schedule/memo, and video apps and a lower intrinsic preference for the remaining apps and the mobile web. In addition, women exhibit a higher intrinsic preference for communication and photo apps and a lower intrinsic preference for entertainment and personal financing apps as compared to men. Lastly, users with high education levels show significantly lower baseline utilities in most categories, except for personal financing apps.

Results in Table 7 show that the satiation level is the highest in the personal financing app category and the lowest in the game app category. Note that as $\overline{\lambda}$ increases the satiation effect also increases. The highest satiation for personal financing apps implies that users access these apps quickly (e.g., checking balance information or making a deposit by simply taking a photo of a check). In contrast, the lowest satiation for game apps suggests that users tend to continue playing games without growing tired of them. Further, we find that there exists substantial user heterogeneity in terms of satiation levels. For example, as age increases satiation with entertainment and music/radio apps increases, while satiation with personal financing and schedule/memo apps simultaneously decreases. In addition, women show significantly lower satiation levels regarding photo and social networking apps than men. Moreover, users with high education levels show significantly higher satiation levels regarding communication, entertainment, and social networking apps. We can also interpret the degree of satiation in terms of $\overline{\alpha}$. However, unlike $\overline{\mu}$, $\overline{\alpha}$ decrease as the $\overline{\lambda}$ increases after exponential transformation in equation (2). Hence, lower $\overline{\alpha}$ represents higher satiation levels.

Table 6. Estimates for Baseline Utility Parameters

	Constant ($\overline{\beta}$)	Demographic Variables (Π_β)						
		Age 30's	Age 40's & over	Female	Income Mid-class	Income Upper-class	Education High-School Graduates	Education University Graduates
Communication	**-1.30**	**-0.18**	-0.03	**0.34**	0.02	**0.21**	-0.05	**-0.13**
	(0.05)	(0.05)	(0.06)	(0.05)	(0.04)	(0.05)	(0.07)	(0.06)
Game	**-2.94**	0.04	**-0.13**	-0.04	-0.04	**-0.22**	0.09	-0.01
	(0.05)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)	(0.08)	(0.07)
Map/Navigation	**-2.76**	**-0.37**	**-0.26**	0.00	-0.02	**0.12**	-0.07	0.06
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)
Entertainment	**-3.15**	**-0.35**	**-0.45**	**-0.23**	-0.04	-0.09	**-0.21**	**-0.39**
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)
Lifestyle	**-2.85**	0.03	-0.06	**0.17**	0.04	0.02	**-0.36**	**-0.21**
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)	(0.07)	(0.05)
Personal Financing	**-3.34**	**-0.12**	**-0.19**	**-0.10**	0.05	**0.15**	**0.33**	**0.39**
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.07)	(0.06)
Music/Radio	**-2.47**	**-0.41**	**-0.43**	0.02	0.01	0.02	**-0.22**	**-0.38**
	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)	(0.07)	(0.05)
Photo	**-2.04**	**-0.16**	**-0.27**	**0.33**	-0.05	-0.05	**-0.17**	**-0.12**
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.07)	(0.05)
Portal	**-2.94**	**0.15**	**0.18**	**0.12**	-0.04	0.03	-0.06	**-0.11**
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)	(0.07)	(0.05)
Schedule/Memo	**-1.81**	-0.05	**0.16**	**0.15**	-0.02	0.04	0.06	0.02
	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.07)	(0.06)
Social Network	-2.54	**-0.40**	**-0.53**	**0.10**	-0.07	0.04	-0.11	-0.10
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.07)	(0.05)
Utility	**-1.60**	**-0.14**	-0.09	0.00	-0.06	-0.09	-0.10	-0.10
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.07)	(0.06)
Video	**-2.76**	**0.12**	**0.10**	-0.04	-0.01	0.02	0.00	-0.07
	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.06)	(0.05)
Web	**-1.77**	-0.02	-0.02	0.01	-0.01	0.03	**-0.17**	**-0.11**
	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.07)	(0.05)

Note: Standard errors in parentheses. Bold: significant at the .05 level. Following Bhat (2008), α_0 is bounded between 0 and 1. The estimated value of α_0 is 0.00 with a standard error of 0.03. To code discrete demographic variables, we use the following as a reference level: we use age 20's or less for age and male for gender, respectively. For monthly income, we create two dummy variables – mid-class ($3,000–$5,000) and upper-class ($5,000 or over), and we use lower-class ($3,000 or less) as a reference level. Similarly, for education, we create two dummy variables – high school graduates and university graduates, and we use students (elementary, middle-and-high school students, university students) as a reference level.

In Figure 1, we map mobile web and app categories according to baseline utility and satiation levels – $\overline{\mu}$ and $\overline{\alpha}$ .The four quadrants of the scatterplot provide some interesting insights into how different the mobile content is in terms of its choice and time use. For example, the top-right, quadrant *I*, represents mobile content that is used widely as well as extensively, including communication apps, utility apps, and mobile web. Furthermore, we can distinguish among the apps with the similar level of baseline utility: low satiation apps from high satiation counterparts. For example, baseline utility levels are similar to each other for apps in quadrant *III* (photo, map and navigation, life style, personal financing, and portal search apps) and quadrant *IV* (social, music, video, entertainment, and game apps), however, satiation levels are greater in the former than in the latter. One notable observation that can be made here is that hedonic app categories grouped in quadrant *IV* tend to be used more extensively than utilitarian counterparts grouped in quadrant *III* when their baseline utility levels are relatively low.

Table 7. Estimates for Satiation Parameters

	Constant ($\bar{\lambda}$)	Demographic Variables (Π_λ)						
		Age 30's	Age 40's & over	Female	Income Mid-class	Income Upper-class	Education High-School Graduates	Education University Graduates
Communication	**1.74**	**0.31**	**0.31**	**-0.22**	**0.05**	0.03	**0.31**	**0.23**
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Game	**1.73**	**0.09**	**0.19**	-0.03	**0.11**	0.04	0.05	-0.04
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)
Map/Navigation	**3.80**	-0.03	**-0.27**	**0.27**	-0.02	0.04	-0.05	-0.07
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)
Entertainment	**2.75**	**0.53**	**0.68**	**0.08**	-0.05	**0.14**	**0.23**	**0.25**
	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.06)	(0.05)
Lifestyle	**3.82**	**0.19**	**-0.22**	-0.02	0.04	**0.30**	-0.02	**-0.13**
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)
Personal Financing	**3.84**	**-0.28**	**-0.30**	**0.19**	**-0.07**	**-0.26**	**-0.21**	**-0.14**
	(0.04)	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)
Music/Radio	**2.28**	**0.44**	**0.51**	**-0.10**	**-0.23**	**-0.12**	-0.05	0.07
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)
Photo	**3.67**	0.05	-0.03	**-0.51**	**0.11**	**0.17**	**0.31**	**0.33**
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)
Portal	**3.59**	**0.33**	0.06	-0.03	-0.01	**-0.08**	**0.28**	**-0.12**
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)
Schedule/Memo	**3.32**	**-0.13**	**-0.32**	**0.15**	0.00	**-0.16**	0.04	-0.01
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)
Social Network	**2.47**	**0.30**	**0.19**	**-0.48**	-0.04	**0.14**	**0.24**	**0.25**
	(0.03)	(0.04)	(0.04)	(0.02)	(0.03)	(0.03)	(0.05)	(0.04)
Utility	**3.13**	**0.25**	**0.19**	**0.18**	0.04	**0.18**	**-0.15**	**-0.08**
	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.05)	(0.04)
Video	**2.96**	**0.08**	**0.14**	**0.06**	**-0.08**	**-0.12**	**-0.12**	(0.03)
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)
Web	**2.23**	**-0.10**	**0.26**	-0.04	0.00	-0.03	**0.25**	**0.07**
	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.05)	(0.04)

Note: Standard errors in parentheses. Bold: significant at the .05 level. To code discrete demographic variables, we use the following as a reference level: we use age 20's or less for age and male for gender, respectively. For monthly income, we create two dummy variables – mid-class ($3,000–$5,000) and upper-class ($5,000 or over), and we use lower-class ($3,000 or less) as a reference level. Similarly, for education, we create two dummy variables – high school graduates and university graduates, and we use students (elementary, middle-and-high school students, university students) as a reference level.

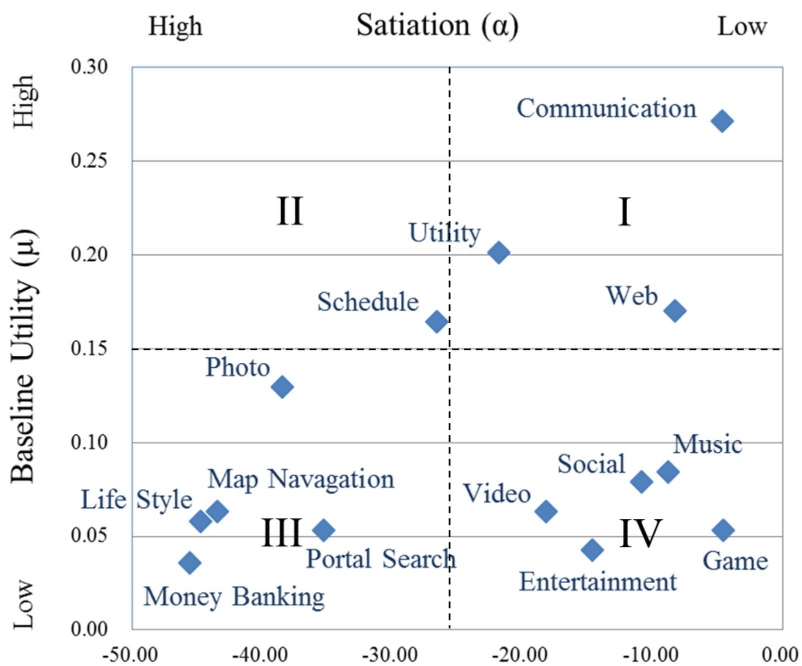

Figure 1. Mapping App Categories According to Baseline Utility and Satiation

Variance Decomposition for Mobile Web and App Categories

Tables 8 and 9 report the estimation results for common factors (Γ_β and Γ_λ) and specific factors (Λ_β and Λ_λ). In addition, in the last three columns in Tables 8 and 9, we decompose the total variations in baseline utilities and satiations for mobile web and app categories into variations attributed to the demographic variables, common factors, and specific factors. Interpretation of the factors is based on estimated factor loading matrices and variance decompositions. For the baseline utility, the first factor loads strongly with positive signs for major utilitarian app categories (communication, map/navigation, lifestyle, personal financing) but with negative signs for major hedonic app categories (game, music/radio, video). A large positive factor score implies high utility for utilitarian apps but low utility for hedonic apps. In contrast, the second factor loads strongly with positive signs for hedonic and social networking app categories (game, entertainment, portal, social network, video) but with negative signs for some utilitarian apps (communication, schedule/memo, web). The variance decomposition shows that, on average, 63% of variation in baseline utility can be explained by demographic variables and 33% by common factors. Specific factors account for the remaining 5%. In marketing, empirical researchers have documented that demographic variables are poor predictors of consumer brand preferences that are estimated on scanner panel data (Singh et al. 2005). However, our result indicates that the demographic variables are powerful predictors of overall variation in app preference captured by the baseline utility. This implies that companies can derive strategic advantage from their knowledge on user demographics.

Table 8. Baseline Utility Factor Estimates and Variance Decomposition

	Estimates			Variance Decomposition		
	Common Factor		Specific Factor	Demographic	Common Factor	Specific Factor
Communication	**0.13** (0.02)	**-0.05** (0.02)	0.02 (0.02)	0.69	0.29	0.01
Game	**-0.08** (0.02)	**0.10** (0.02)	**0.10** (0.02)	0.37	0.38	0.26
Map/Navigation	**0.08** (0.02)	-0.01 (0.02)	**0.06** (0.02)	0.75	0.15	0.10
Entertainment	-0.01 (0.02)	**0.12** (0.02)	**0.05** (0.02)	0.86	0.12	0.02
Lifestyle	**0.07** (0.02)	0.03 (0.02)	0.03 (0.02)	0.76	0.21	0.03
Personal Financing	**0.11** (0.02)	0.00 -	0.00 (0.02)	0.64	0.34	0.02
Music/Radio	**-0.12** (0.02)	0.02 (0.02)	0.02 (0.02)	0.86	0.13	0.01
Photo	0.02 (0.02)	-0.01 (0.02)	0.02 (0.01)	0.99	0.01	0.01
Portal	**0.09** (0.02)	**0.18** (0.02)	0.00 (0.02)	0.15	0.83	0.02
Schedule/Memo	**0.10** (0.02)	**-0.05** (0.02)	0.02 (0.01)	0.49	0.49	0.02
Social Network	0.03 (0.02)	**0.06** (0.02)	**0.04** (0.02)	0.92	0.06	0.02
Utility	0.02 (0.02)	-0.02 (0.02)	0.02 (0.01)	0.90	0.05	0.05
Video	**-0.06** (0.02)	**0.06** (0.02)	**0.03** (0.01)	0.29	0.63	0.07
Web	**-0.05** (0.02)	**-0.18** (0.02)	**0.05** (0.01)	0.08	0.86	0.06

Note: Standard errors in parentheses. Bold: significant at the .05 level.

Table 9. Satiation Factor Estimates and Variance Decomposition

	Estimates			Variance Decomposition		
	Common Factor		Specific Factor	Demographic	Common Factor	Specific Factor
Communication	**0.36** (0.01)	**0.11** (0.01)	**0.02** (0.01)	0.33	0.66	0.01
Game	**-0.05** (0.01)	**0.18** (0.01)	**0.53** (0.01)	0.03	0.11	0.86
Map/Navigation	-0.01 (0.01)	**0.35** (0.01)	**0.44** (0.01)	0.11	0.34	0.55
Entertainment	0.04 (0.02)	**0.08** (0.02)	**0.90** (0.02)	0.14	0.01	0.86
Lifestyle	**0.05** (0.02)	**0.76** (0.02)	**0.46** (0.02)	0.05	0.70	0.25
Personal Financing	**0.18** (0.02)	0.00 -	**0.76** (0.01)	0.09	0.05	0.86
Music/Radio	**0.13** (0.01)	**0.31** (0.01)	**0.84** (0.01)	0.08	0.13	0.80
Photo	**0.46** (0.01)	**0.10** (0.01)	**0.09** (0.01)	0.28	0.69	0.03
Portal	**0.35** (0.01)	**0.83** (0.01)	**1.43** (0.01)	0.01	0.28	0.71
Schedule/Memo	**0.23** (0.01)	**0.28** (0.01)	**0.45** (0.01)	0.09	0.35	0.55
Social Network	**0.52** (0.01)	**0.12** (0.01)	**0.43** (0.02)	0.20	0.48	0.32
Utility	**0.13** (0.01)	**0.45** (0.01)	**0.70** (0.01)	0.03	0.30	0.68
Video	**-0.22** (0.01)	**0.47** (0.01)	**0.63** (0.02)	0.01	0.40	0.59
Web	**-0.16** (0.01)	**0.10** (0.01)	**0.85** (0.01)	0.04	0.05	0.91

Note: Standard errors in parentheses. Bold: significant at the .05 level.

For the satiation parameters, we find that the first factor loads with positive signs for communication, photo, portal, and social network apps, but with negative signs for game, video, and the mobile web. A large positive factor score implies high satiation for communication, photo, portal, and social network apps and low satiations for game, video, and the mobile web. The second factor loads strongly with positive signs for all categories, capturing uniformly lower (higher) levels of usage with a large positive (negative) factor score. The variance decomposition shows that only 11% of variation in satiation parameters is explained by demographic variables. A further 57% of variation is accounted for by specific factors and 33% by common factors. Unlike in baseline utility, demographic variables are poor predictors of heterogeneity in app satiation. Large variation is explained by specific factors, indicating that satiation is user-specific trait. It is difficult to predict user's app satiations based on her demographic information. Instead, one has to use individual level history data for reliable prediction.

Correlation across Mobile Web and App Categories

Tables 10 and 11 show the correlation matrices for baseline utility and satiation parameters across mobile web and app categories. Table 10 demonstrates correlation in the baseline utility. We find that people who use social apps (e.g., Facebook) also frequently use photo apps (e.g., Instagram) and people who use entertainment apps (e.g., Naver Cartoon) frequently use music/radio apps (e.g., Music Player) as well. These results indicate that social apps and photo apps are economic complements and entertainment apps and music/radio apps are also frequently used together as complements. In contrast, people who use mobile portal apps (e.g., Google app) less frequently use the mobile web (e.g., Google on browser), indicating that portal apps and the mobile web are economic substitutes. This result implies that mobile portal apps offer similar functionality and usability to the mobile web. For example, users can access information quickly and easily regardless of whether they use a mobile search app or visit a mobile search website. Because people have a limited amount of time to spend on the mobile web and app categories, substitution occurs in our context. Table 11 shows correlation in satiation levels. People who spend a great deal of time on social apps (e.g., Facebook) also spend a lot of time on communication apps (e.g., WhatsApp) and photo apps (e.g., Instagram), suggesting that social, communication, and photo apps are economic complements to each other. For example, some people take a picture or video, use Instagram to choose a filter to transform its look and feel, and post it to Facebook or share it on WhatsApp.

Table 10. Correlation Matrix for Baseline Utility Parameters

	Communication	Game	Map/Navigation	Entertainment	Lifestyle	Personal Financing	Music /Radio	Photo	Portal	Schedule/Memo	Social Network	Utility	Video	Web
Communication	1.00													
Game	-0.47	1.00												
Map/Navigation	0.60	-0.34	1.00											
Entertainment	0.00	0.40	0.41	1.00										
Lifestyle	0.59	-0.01	0.44	0.33	1.00									
Personal Financing	0.09	-0.37	0.24	-0.38	-0.14	1.00								
Music/Radio	0.21	0.28	0.42	0.81	0.33	-0.62	1.00							
Photo	0.66	0.09	0.45	0.35	0.77	-0.32	0.61	1.00						
Portal	0.20	0.20	0.00	0.09	0.36	0.14	-0.20	0.06	1.00					
Schedule/Memo	0.78	-0.63	0.34	-0.40	0.19	0.31	-0.30	0.16	0.16	1.00				
Social Network	0.48	0.17	0.68	0.74	0.64	-0.17	0.78	0.80	0.12	-0.08	1.00			
Utility	0.41	0.03	0.60	0.70	0.45	-0.40	0.75	0.61	-0.19	0.11	0.71	1.00		
Video	-0.63	0.48	-0.58	0.06	-0.40	-0.31	-0.05	-0.45	0.32	-0.59	-0.31	-0.48	1.00	
Web	0.15	-0.39	0.10	-0.13	-0.06	-0.24	0.23	0.16	-0.84	0.11	-0.02	0.31	-0.44	1.00

Table 11. Correlation Matrix for Satiation Parameters

	Communication	Game	Map/Navigation	Entertainment	Lifestyle	Personal Financing	Music/Radio	Photo	Portal	Schedule/Memo	Social Network	Utility	Video	Web
Communication	1.00													
Game	0.07	1.00												
Map/Navigation	-0.02	0.15	1.00											
Entertainment	0.23	0.06	-0.01	1.00										
Lifestyle	0.23	0.25	0.52	0.07	1.00									
Personal Financing	0.02	-0.05	0.06	-0.09	0.00	1.00								
Music/Radio	0.31	0.11	0.13	0.12	0.28	-0.04	1.00							
Photo	0.92	0.02	-0.05	0.14	0.20	0.05	0.25	1.00						
Portal	0.30	0.14	0.28	0.06	0.43	0.04	0.20	0.26	1.00					
Schedule/Memo	0.27	0.08	0.34	-0.03	0.41	0.17	0.14	0.27	0.30	1.00				
Social Network	0.80	0.02	-0.03	0.15	0.19	0.03	0.24	0.80	0.23	0.22	1.00			
Utility	0.26	0.16	0.32	0.08	0.46	0.01	0.21	0.20	0.29	0.28	0.19	1.00		
Video	-0.06	0.21	0.34	0.05	0.46	-0.06	0.17	-0.12	0.22	0.16	-0.09	0.27	1.00	
Web	-0.07	0.08	0.02	0.03	0.05	-0.06	0.03	-0.11	0.01	-0.05	-0.09	0.02	0.11	1.00

Conclusions

This paper contributes to an emerging stream of literature on the economics of mobile internet and mobile marketing as the first study to quantify the baseline utility and satiation levels of mobile app categories. Moreover, based on novel approaches and sound analytics we examined use interdependence and dynamics among diverse categories of mobile web and apps. Furthermore, using large-scale panel data on mobile app and web time-use histories, we developed a unique multiple discrete/continuous model of app selection and time-use decisions. The vectors of individual-level baseline utility and satiation parameters were modeled as functions of observed mobile user characteristics and a small number of unobservable user-specific factors. These observed and unobserved user-specific components capture the dependence on mobile web and app category selection and time-use decisions. Our approach makes several methodological contributions to empirical frameworks involving multiple discrete/continuous extreme value choices (Bhat 2005) by reducing the number of parameters required to estimate a full covariance matrix using a factor analytic structure. The frameworks, mechanisms, and processes articulated in this study offer high flexibility and efficiency in estimating parameters even when the number of alternatives is large, a phenomenon that is increasingly prevalent in big data analysis. Consequently, our analytical paradigm and comprehensive computational procedures can supply illuminating ideas for the advancement of big data analytics in general and mobile app analytics.

This paper has direct managerial implications for decisions surrounding the allocation of advertising dollars across mobile web and different app categories. Our empirical results show that both positive and negative correlations exist in the baseline utility and satiation levels of mobile web and app categories. For example, social and communication apps are strong economic complements to each other. Hence, brand messaging through social apps (e.g., Facebook) can reinforce brand advertising messages that users receive from communication apps (e.g., WhatsApp).That is, advertising displayed through Facebook can convey a persuasive marketing message by serving as a memory cue that can trigger users to consider the advertised brand based on existing knowledge stored in their memories from prior advertisements through WhatsApp. Further, we find that users' demographic variables and unobservable factors play important roles in explaining app selection and time-use decisions. These results offer insights related to mobile user segmentation, targeting, and optimal media planning in the mobile app economy, allowing a firm to identify which consumers will be most effectively reached based on consumers' diverse needs, varying demographic characteristics, and idiosyncratic behavioral patterns. Our findings also offer valuable insights into predicting app use time for a particular user and identifying a segment with extensive use of a certain app. Lastly, this paper contributes to the methodology literature in IT big data analytics by proposing a unique model for time-use of various IT artifacts and their interdependence using large-scale, micro-level data.

References

Bhat, C. 2005. "A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decision," *Transportation Research Part B* (39), pp. 679-707.

Bhat, C. 2008. "The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions," *Transportation Research Part B* (42), pp. 274-303.

Bhatnagar, A. and Ghose, S. 2004. "Online Information Search Termination Patterns across Product Categories and Consumer Demographics," *Journal of Retailing* (80:3), pp.221-228

Carare, O. 2012. "The Impact of Bestseller Rank on Demand: Evidence from the App Market," *International Economic Review* (53:3), pp. 717-742.

Chintagunta, P.K. 1993. "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science* (12: 2), pp. 194–208.

Chintagunta, P.K. 1994. "Heterogeneous Logit Model Implications for Brand Positioning," *Journal of Marketing Research* (31: May), pp. 304-311.

Dubin, J.A. and McFadden, D.L. 1984. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica* (52: 2), pp. 345-362.

Einav, L., Levin, J., Popov, I., and Sundaresen, N. 2014. "Growth, Adoption, and Use of Mobile Commerce," *American Economic Review: Papers & Proceedings* (104: 5), pp. 489-494.

Elrod, T. and Keane, M.P. 1995. "A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data," *Journal of Marketing Research* (32: February), pp. 1-16.

Flurry. 2013. There's An App Audience for That, But It's Fragmented. http://www.flurry.com/bid/96368/There-s-An-App-Audience-for-That-But-It-s-Fragmented#.U2Q5k_mSxjM.

Forbes. 2014. Facebook Buying WhatsApp - $19B In Cash And Stock, http://www.forbes.com/sites/benkepes/2014/02/19/breaking-facebook-buying-whatsapp-19b-in-cash-and-stock/.

Frier, S. 2013. "With $200 Million in Revenue, South Korea's Top Messaging App Is All Smiley Faces," Bloomberg, http://www.bloomberg.com/news/2013-12-22/with-200-million-in-revenue-south-korea-s-top-messaging-app-is-all-smiley-faces.html.

Garg, R. and Telang, R. 2013. "Inferring App Demand from Publicly Available Data," *MIS Quarterly* (37: 4), pp.1253-1264.

Ghose, A. and Han, S. 2011. "An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet," *Management Science* (57: 9), pp. 1671-1691.

Ghose, A., Goldfarb, A., and Han, S. 2013. "How is the Mobile Internet Different? Search Costs and Local Activities," Information Systems Research (24: 3), pp. 613-631.

Ghose, A. and Han, S. 2014. "Estimating Demand for Mobile Applications in the New Economy," *Management Science* (60: 6), pp. 1470-1488.

Gordon, M. E. The History of App Pricing, And Why Most Apps Are Free. http://www.flurry.com/bid/99013/The-History-of-App-Pricing-And-Why-Most-Apps-Are-Free#.VAVYBfmSz1B.

Hansen, K., Singh, V., Chintagunta, P. 2006. "Understanding Store-Brand Purchase Behavior across Categories," *Marketing Science* (25: 1), pp. 75–90.

Hendel, I. 1999. "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," *Review of Economic Studies* (66), pp. 423–446.

Hong, S. and Tam, K. 2006. "Understanding the Adoption of Multipurpose Information Appliances: The Case of Mobile Data Services," Information Systems Research (17:2), pp.162-179

Igarashi, T., Takai, J. and Yoshida, T. 2005. "Gender Differences in Social Network Development via Mobile Phone Text Messages: A Longitudinal Study," Journal of Social and Personal Relationships (22:5), 691-713

Keane, M.P. 1993. "Simulated Estimation for Panel Data Models with Limited Dependent Variables," in Rao C.R., Maddala G.S., Vinod D. (eds) *The Handbook of Statistics and Econometrics*. Elsevier Science Publishers, Amsterdam, pp. 545-572.

Khalaf, S. Apps Solidify Leadership Six Years into the Mobile Revolution. http://www.flurry.com/bid/109749/Apps-Solidify-Leadership-Six-Years-into-the-Mobile-Revolution#.VAVXI_mSz1C.

Kim, J., Allenby, G.M., Rossi, P.E. 2002. "Modeling Consumer Demand for Variety," *Marketing Science* (21: 3), 229–250.

Luo, L., B.T. Ratchford, B. Yang. 2013. "Why We Do What We do: A Model of Activity Consumption," *Journal of Marketing Research* (50: 1), pp.24-43

Singh, V., Hansen, K., and Gupta, S. 2005. "Modeling Preferences for Common Attributes in Multicategory Brand Choice," *Journal of Marketing Research* (42: 2), 195–209.

Spissu, E., A.R. Pinjari, C.R. Bhat, R.M. Pendyala, K.W. Axhausen. 2009. "An Analysis of Weekly Out-of-Home Discretionary Activity Participation and Time Use Behavior," *Transportation* (36: 5), pp.483–510.

Venkatesh, V. and Morris, M. 2000. "Why Don't Men Ever Stop To Ask For Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behaviors," *MIS Quarterly* (24:1), 115-139

Xu, J., Forman, C., Kim, J., and van Ittersum, K. 2014. "News Media Channels: Complements or Substitutes? Evidence from Mobile Phone Usage," *Journal of Marketing* (78: 4), pp. 97-112.

Yonhap News. 2014. Android rules S. Korean market in 2013. http://english.yonhapnews.co.kr/business/2014/01/21/30/0501000000AEN20140121001000320F.html.