Elaboration Document

Ubiquitous Product’s Social Learning Process’s Effect on Individual Choices:

Empirical Study of an African App-store

Author: Meisam Hejazi Nia

Advisors: Dr. Norris Bruce, Dr. Brian Ratchford

Date: 08/01/2014

**Abstract**

Fueled by widespread smartphone adoption and increasing mobile internet access the mobile app market has experienced significant growth during the last 3 years. The mobile app market reached 64 billion download in 2012 up from 8.2 billion downloads in 2010 and its total revenue in 2012 was $18 billion up from $5.2 billion in 2010. In this market three types of agents operate: (1) platform owners (2) app developers and (3) consumers. As mobile phones are ubiquitous, consumer’s preference for an app and app category can influence the preference of others in many ways, ranging from psychological benefits of social identification, social learning, and inclusion to the benefit of network externalities, direct network effect. In other word, individual decision to adopt an app may be driven by aggregate process of population awareness, adoption, and learning. For example if a consumer observes that her friends, two other consumers in the bus, and a group of other consumers in a restaurant are talking about specific mobile app, she may be likely to be curious to search or download the mobile app. In addition, observing this talk, the consumer may generalize this information, inferring that this mobile app is going to be hot app shortly, and concluding that it may worth to try it, yet this generalization may come with misperception. For the mobile app platform owner this social learning and misperception process are managerially relevant, as they may suggest different policies including, on one hand, encouraging trial, e.g. suggesting issuance of free version, or giving viral marketing incentives, e.g. bring new customers and get bonus, and on the other hand, correcting misperceptions through making the data available online to facilitate observational learning. For the platform owner that seeks the diffusion of its platform, to earn more revenue, it may also be relevant to understand how app-categories complement or substitute each other, and what the optimal assortment size within a category is to decide on how to change revenue share contract within each category to maximize the revenue from the app store platform.

Thus, given exceptional characteristics of ubiquitous products (e.g. mobile apps), we develop a hierarchical Bayesian non-linear state space structural model[[1]](#footnote-1) to study the effect of aggregate preference’s interdependence on individual consumers adoption choice, the misperception of individual consumers, and the substitution and complementarity of app-categories, on app store platform diffusion. We apply our model to a unique dataset of consumer’s download choice on a newly launched app store in Africa[[2]](#footnote-2), for 350 days. We run counterfactual analysis to find the effect of various app store platform design and incentive decisions on consumer choices. Our results should be of interest to academics and business community.

Keywords: mobile app store platform, interdependence of consumer preferences, consumer misperception, social learning, diffusion model, app category complementarity and substitution, hierarchical mixture Bayesian, state space, structural model, Markov Chain Monte Carlo

**Introduction**

Smartphone are ubiquitous in telecommunication market today, to the point that no matter which mobile operator you select in U.S., e.g. T-Mobile, Verizon, AT&T, or Sprint; you have an option to adopt a long term contract that includes a smartphone handset. These smartphones may have one of the three dominant operating systems (OS), i.e. Google Android, Apple iOS, or Windows. All the three OS companies have opened their platform for third party developers to develop mobile apps, which in turn help the diffusion of the main OS platform. Fuel by this decision of the OS platforms, and flood of mobile app developers in this eco-system, new actors, or intermediaries called app-store stepped in to facilitate the transaction between app developers and app users, taking revenue share for their services from publishers and developers.

Some of these app-store platforms are showing the number of users who have downloaded a mobile app with some perturbation, i.e. rounding it by thousands or millions sometime, and some do not show it at all. For example while Google play shows the range of number of installs of an app, Apple, Microsoft, and Amazon do not show this information at all. Showing the number of installs of an app, may help to solve consumer’s misperception about the quality of a mobile app, through presenting the other user’s total download actions, so it may be interesting to know why none of the app stores have solved this misperception completely.

Observing other’s user’s action does not have to be online. In other word, as users carry their smartphones in most of their social life events[[3]](#footnote-3), they have become ubiquitous, and this ubiquity can facilitate the diffusion of mobile apps. More precisely, social learning process through observing others’ choice of mobile app may play a role in mobile app’s diffusion. This social learning process may even be central for mobile apps, as they are classified as disruptive innovations[[4]](#footnote-4). For example, one may observe couple of friends, a consumer sitting at the bus stop, another consumer in the restaurant, and the other consumers in class room who use or talk about a specific mobile app, and she may create a mental note for the app category, or app name to check it later, or alternatively she may check it on her cellphone right away. Even this consumer may be affected by the number of her friends who use this app, modifying her preference for specific mobile app or mobile app category. In this scenario, a consumer not only moves from awareness phase to interest phase in purchase funnel, but also she modifies her preferences, inserting mobile app or mobile app category among her consideration set.

Interdependence of consumer preference has long been studied in marketing in the context of geographical or social neighborhood by Yang and Allenby (2003), Stephen and Toubia (2010), Lehmmes and Croux (2006), Nair et al. (2010), etc., yet about mobile apps, one may not only learn from her peers, but also for a random person she observes in a party, restaurant, cinema, or even class, talking or using a mobile app, or perhaps press. This process suggests that the interdependence between customers’ preferences may be an aggregate diffusion process, which affects individual choices.

Mentioning perception of individual about the aggregate process coins the concept of consumer misperception. In other word, bounded rationality theory of Simon (1972) suggests that consumers do not fully solve the problems, rather they use heuristics. Parallel with bounded rationality theory in decision science, marketing scholars have long been modeling consumers’ misperception in referring to consumer learning theory, by assuming normal distribution of individual’s misperception, and an updating process to correct this prior misperception based on the data signals that consumer receives (Erdem and Kean 1996). These studies may suggest that although there is interdependence of preference of individuals, there is an error involved in this process, called misperception. Knowing the variance of this error and its effect may be relevant for the app store platform, as the owner can share information thoroughly or partially to affect consumers’ misperception. In other word, if an app store platform knows that consumers’ perception about the level of diffusion of each app is biased, and this bias is detrimental to its profit, then it may decide to correct this bias through generating an informative marketing signal, to facilitate social learning.

In addition to deciding about marketing signal presentation, an app store platform may be interested to know what promotion policy to follow for different mobile app categories. To know this optimal promotion policy, an app store platform may not only be interested to extract information from those who have already adopted an app in a given app category, but also it may be interested to extract information about the number of potential users who are aware of the app and app category, but they have not taken an action yet. An app store platform owner may also be interested to extract information about whether chatter between consumers or consumer’s joy of testing new app drives the diffusion of a given mobile app, and app category. Answering these questions may not solely be possible by looking at the sales data, as that only contains information about those who have already adopted an app, so it may need individual level choices. Individual choice data allows not only capturing this intricacy, but also it allows policy simulation.

Policy simulation will allow the app store platform to study which incentive policy to use (e.g. reducing revenue share on some relevant apps, suggesting trial through free version of an app, or using viral marketing strategies by giving bonus to those who bring new consumers) to drive the diffusion of an app. Another decision that may an app store platform be interested to make informatively is which app category drives the diffusion of the other app categories in an app store. Knowing this information may help app store platform to decide on which app category to promote first, as a loss leader. For example, the app store owner may find app category of Book and References to be complement to the app category of Trivia, so she may decide to give an incentive on one to drive the diffusion of the other. On the flip side, an app store platform may find that two app categories are competitors, so it may decide to hold on the incentive that it had planned to give on the other category.

The fourth question that may be of interest of an app store platform is the optimal assortment size. Deciding how many apps are allowed to be offered in each app category is a governance decision of an app store platform. If an app store platform finds that the more the number of the apps in the category, the less likely the consumer is to consider the category due to the search cost, an app store platform may decide to either increase its revenue share on the established app categories to dis-incentivize the development of new app over the app store, or simply to break the category into sub categories.

Thus given unique features of the diffusion and governance of a mobile app-store platform, we consider the following substantive questions: (1) Does the offline diffusion of mobile app and mobile app categories drive the individual choice of consumers? (2) How can an app-store use the misperception of consumers about aggregate diffusion of an app and an app category to maximize the diffusion of those apps? (3) What type of policy should an app-store platform use (e.g. encourage trial, or encourage sharing) to increase the diffusion of a given mobile app category? (4) Which mobile app categories are substitute, and which are complements? (5) What is the optimal number of mobile apps in each app category?

To address these questions about mobile app and mobile app category diffusion, we adopt an individual choice model of mobile app category and mobile app that uses Bass Diffusion Model (BDM) to model interdependence of consumer’s awareness and preferences. We capture diffusion of mobile app and mobile app category in a latent measures that evolves with parameters of innovation (p), and imitation (q), cumulative adoption(n), and app and app category market potential (M) . This modeling approach allows us to capture not only direct network effect of an app and an app category in imitating parameter, but also complementarity and substitution of different apps and app categories over the app store in innovation parameter. Moreover, we allowed for misperception of consumers in our model. This structural modeling approach allows us to run counterfactual analysis for different policies that an app store platform can follow.

Our study of interdependence of consumer choices is mostly related to studies by Yang and Allenby (2003), Stephen and Toubia (2010), Lehmmes and Croux (2006), Nair et al. (2010), who model the peer effect either due to geographical (distance) or social network neighborhood. In addition our study is related to study by Sriram et al. (2007) and Liu et al. (2010) on the complementarity and substitution of digital goods. Although these studies have contributed greatly to our understanding of the phenomenon, none has considered the impact of observation of ubiquitous product on the individual consumer decision. Given ubiquity of smart phones and mobile apps, it may be important to consider the effect of information about aggregate diffusion of an app and consumer misperception stemmed from induction of consumers after observing app usage in various neighborhoods (e.g. web, social network, society, geographical, press) on the individual choice of an app. In addition, in considering the complementarity and substitution of ubiquitous products, it may be relevant to think more in terms of complementarity and substitution of products in the mind of the consumers rather than choice observations. For example, some consumer’s may find two mobile app’s to be complement, but they do not download them together right away, but they keep it at the back of their mind to download one later.

To help address some of these limitations and research questions, we extended individual logit choice model to model the interdependence of consumers’ choices. To do so, we incorporated Bass Diffusion Model (BDM) for the evolution of goodwill of each app and app category and their substitution and complementarity. The resulting choice model is dynamic, nonlinear, and heterogeneous. In contrast to many studies, in order to separate state dependence from heterogeneity, we use normal mixture model of heterogeneity, as a substitute for normal model of heterogeneity. Moreover, we extract the aggregate diffusion of awareness, goodwill, or preference of an app or an app category from individual choice data; thus, for estimation, we re-cast BDM differential equations into non-linear discrete-time, state-space forms; and then estimate them using MCMC approach to joint estimation of choice model and Extended Kalman Filter.

This paper, thus, contributes to the emerging literature on ubiquitous products. First, it introduces a dynamic choice model to investigate several factors that could affect choice of an app and an app category on the app store platform. Although other studies in marketing have considered dynamic choice model of digital goods, none have the effect of dynamic aggregate goodwill of ubiquitous product (e.g. mobile apps) and the dynamic of interdependence of consumer choice preferences. Second, the paper examines the influence of consumer misperception on the choice of an app and an app category. Misperception is an important element, as it may suggest censoring of information on app store platforms. Third, we extract the diffusion of an app and app store on the platform from the individual choice data. To the best of our knowledge, this is the first time that one extracts this information from choice data for ubiquitous product empirically, yet it may have an advantage over extraction of this information from sales data, as the real diffusion is a latent process in individual minds that may either result in purchase or no purchase based on other factors as well. Fourth, although many studies have normative suggestions about the effect of assortment size based on search theories, yet to the best of our knowledge we are the first paper to empirically test the predictions of these theories. Fifth, to estimate the resulting model, the paper employs hierarchical Bayesian choice approach to Extended Kalman Filter. These approaches should be of interest to academia and a number of commercial entities interested in the performance of ubiquitous products, and relevant store platform.

The data we used for our estimation were collected by an African telecom operator on downloads of mobile apps’ on the app store of its global partner. The app-store is launched within 330 days prior to our study, but we use around 150 days daily downloads of mobile apps on the app store, the period in which the dominant app in our given categories is launched. We selected 10 different app categories, and within each category we selected app’s that have significantly greater number of downloads given their category during the period of study, total of 11 apps that are dominant in their relevant category. This choice allows us to assume that the apps are in an monopolistic competition market, as within each category we have choice of either the dominant app, or app’s that are not dominant. We investigate geographic source of dependence of awareness and preference of consumers through goodwill of the mobile app, by filtering our data on one geographical neighborhood. Geographic neighborhood is created by physical proximity, through mapping IP address of each consumer to the corresponding city. For this purpose we wrote a crawler and scraper in Perl to scrape the location of each corresponding IP in an IP-location mapping website. The data consists of a random sample of 1,300 consumers, live in a town in Africa, as we conditioned only on one city, and with around 1,880 unique downloads for the course of around 150 days in 2013 and 2014.

**Literature Review**

This study draws upon several streams within the literature that have investigated: *i*) Interdependence of consumer preference; *ii*) Product Category Complementarity; iii) Mobile app store dynamics; and *iv*) new product diffusion models in general. Given the breadth of these areas across multiple disciplines, what follows is only a brief review of these relevant streams.

**Interdependence of consumer preference**

Quantitative models of consumer purchase behavior often do not recognize that consumer choices may be driven by the aggregate diffusion of awareness and preferences. Economic models of choice typically assume that an individual’s latent utility is a function of brand and attribute preferences, not the preferences of others, yet in the case of free experience goods, when there is no pricing mechanism, a new model which accounts for aggregate diffusion of awareness and preferences may explain the choice of consumers better.

Preferences and choice behavior are influenced by a consumer’s own tastes and awareness, and also the taste and awareness of others. People who live in the same geographical area often become aware of an innovation sooner and their preference are more coherent, resulting in the choices that are interdependent. Interdependence may be driven by social concern, by endorsements from respected individual that increase a brand’s credibility, or by learning the preference of others through diffusion of information. Many studies tried to address this issue including Yang and Allenby (2003), which uses cross sectional data to model preference dependanc. Stephen and Toubia (2010) studies value implications of social network between sellers in a large online social commerce marketplace. Lehmmes and Croux (2006) apply bagging and boosting classification techniques to a customer database of an anonymous U.S. wireless telecommunication company, and find that both methods significantly improve accuracy in predicting churn.

Bell and Song (2007) examine customer trials at Netgrocer.com and drawing on studies in marketing and economics conjecture that exposure spatially to proximate others (through direct social interaction or observation), can influence decisions of those who have yet to try, find that trials arise from utility-maximizing behavior and they find that neighborhood effect is significantly positive and economically meaningful. Hartmann et al. provides a selective review of how insights from other discipline can improve and inform modeling choices of social interactions. Aral and Walker (2011) present a method that uses in vivo randomized experimentation to identify influence and susceptibility in networks while avoiding the biases inherent in the traditional estimates of social contagion to identify social influence in the networks.

Nair et al. (2010) quantify the impact of social interaction and peer effects in the context of physicians’ prescription choices, using detailed individual-level prescription data, along with self-reported social network information, and find that physicians prescription behavior is significantly influenced by the behavior of research-active specialists, or “opinion leaders,” in the physician’s reference group. Bradlow et al. (2005) reviews literature on generalizing the notion of a map to include demographic and psychometric representations, to argue for spatial models to capture a variety of effect that impact firm or consumer decision behavior. Hartmann (2010) develops a model for demand estimation in the context of social interactions. It models the decision of a group of customers as an equilibrium outcome of an empirical discrete game, in which group members must be satisfied with chosen outcome. Yang et al. (2005) estimate the preference independency in television program viewership between spouses using simultaneous equation model, and show that the models that ignore interdependency of preference yield biased estimates of consumers sensitivity to observed attribute preferences.

Narayan et al. (2011) studies how multi-attribute product choices are affected by peer influence, and propose two-stage conjoint-based approach to examine three behavior mechanisms of peer influence, finding that when faced with information on peer choices, consumers update their attribute preferences in Bayesian manner. Kurt et al. (2011) investigate the interactive influence of the presence of an accompanying friend and consumer’s agency communion orientation on the consumer’s spending behavior, finding that shopping with friend can be expensive to agency-oriented consumer (e.g. males) but not for communion-oriented consumer (e.g. females). Chung and Rao (2012) present a general consumer preference model for experience products that allows for situations in which it is not easy to consider some qualitative attributes of a product or when there are too many attributes relative to the available amount of preference data, by capturing the effects of unobserved product attributes with the residuals of reference consumer for the same product. Choi et al. (2010) develop a Bayesian spatio-temporal model to study two imitation effects in the evolution of demand at an Internet retailer, allowing imitation behavior to be reflected both in geographic proximity and in demographic proximity, finding that the proximity effect is especially strong in the early phases of demand evolution, whears the similarity effect becomes more important with time.

**App Category Complementarity**

Sriram et al. (2007) presents a framework of durable good purchasing behavior in related technology product categories, and using a household level information on category-level firmst-time adoption decisions in three categories—personal computer, digital cameras, and printers--- it finds strong complementary relationship between the three categories. Liu et al. (2010) studies complementarities and the demand for home broadband internet services, with an eye on the demand side explanation rather than supply side for large share of phone companies in the DSL broadband market, and it finds evidence of strong complementarities between the consumption of broadband and its related categories. Yang and Ching (2013) develop a structural consumer life-cycle model to investigate consumer’s adoption and usage decision of ATM cards, and using data on ATM cards adoption in Itly, it finds that one could significantly overestimate adoption costs for the elderly when ignoring their shorter life span.

Ma et al. (2012) models household’s contemporaneous brand choice outcome in complementary product categories, explicitly capturing cross-category dependencies in brand choice outcome of household, and using scanner panel data on cake mix and frosting categories, it finds that complementarity accounts for the vast majority of the estimated cross category effects in demand. Karaca-Mandic (2011) estimates complementarities between DVD player adoption and availability of content on DVD, and using household level quasi-panel data on adoption decision of households, onthly data on availability of movies for rental on DVD at local regions and on the number of movies released on DVD nationally, it finds significant complementarity.

Lee et al. (2013) propose a direct utility model with a latent decision sequence for measuring asymmetric complementarity effects to capture differential response to cross-category purchases and inventories, and using scanner panel data on milk and cereal purchases, it investigates cross-price elasticities and spillover effects of merchandising variables. Niraj et al. (2008) models more than one purchase decision of the household simultaneously, through two-stage bivariate model of incidence and quantity outcome in multiple categories, and it finds that cross-category promotional spillovers are asymmetric between the two product categories of bacon and eggs. Wedel and Zhang (2004) develop a sales model that can be used to assess cross-category price effects that are specific to individual items (stockkeeping units) from store-level sales data.

Song and Chintagunta (2006) study the cross-category effects of marketing activities using aggregate store-level scanner data on four product categories--- liquid laundary detergents, powdered laundry detergents, and liquid fabric softeners, and sheet fabric softeners—, and it finds evidence for a complementary relationship between liquid softeners and both forms of detergents. It also finds that the magnitude of cross-category elasticities are brand specific, i.e., different brands in a category have a different price impact on the demand for a brand in another category.

Oestreichter-Singer and Sundrararajan (2008) tests the conjectures that the explicitly visibility of “product networks” can alter demand spillover across their constituent items on the data about copurchase networks and demand levels associated with more than 250,00 interconnected books offered on Amazon.com, finding that on average the explicit visibility of a copurchase relationship can lead to up to an average threefold amplification of the influence that complementary products have on each others’ demand level. Bezawada et al. (2009) investigate the effects of aisle and display placement on the sales affinities between categories, by developing a spatial model of brand sales that allows for asymmetric store-specific affinity effects between two or more categories.

**Mobile app store dynamics**

Ghose and Han (2014) build a structural econometric model to quantify the vibrand platform competition between mobile (smartphone and tablet) apps on the Apple iOS and Google Android platforms and estimate consumer preference toward different mobile app characteristics, finding positive effect of in-app purchase option wherein user can complete transactions within the app, and negative effect of in-app advertisement option. Carare (2012) uses daily data on the ranking by sales of the top 100 apps sold through Apple’s App Store, and it finds that lagged bestseller rank information has causal effect on mobile apps’ demand. Garg and Telang (2013) present an innovative method to use public data to infer the rank-demand relationship for the paid apps on Apple’s iTune App store, and it finds that the top-ranked paid app for iPhone generates 150 times more downloads compare to the paid app ranked at 200. Liu et al. (2012) examines the effect of the freemium strategy on Google play, and by analyzing a large panel dataset, it finds that the freemium strategy is positively associated with increased sales volume and revenue of the paid apps.

Ghose et al. (2011) explores how internet browsing behavior varies between mobile phones and personal computers, and arguing that smaller screen sizes on mobile phone increase the cost to the user of browsing for information, and wider range of offline locations for mobile internet usage that causes importance of local activity, it finds significant ranking effect and more geographic match relevance on mobile internet. Ghose and Han (2011b) develops and estimates a dynamic structural model of user behavior and learning with regard to content generation and usage activities in mobile multimedia environments, in which user learn about two different categories of web content: social network and community and mobile portal, and it allows for both direct experience and indirect experience from neighbors, finding that learning based on direct experience is more accurate (has less variability) than learning based on indirect experience. Ghose and Han (2011a) quantify how user mobile internet usage relates to unique characteristics of the mobile internet, focusing on examining how the mobile-phone-based content generation behavior of users relates to content usage behavior, analyzing whether there is positive or negative interdependence. Kim et al. (2008) develop a structural model to examine user demand for voice service and SMS, and it measures the own- and cross-price elasticity of these services, finding cross price elasticity significant due to complementarity and substitution.

**New product diffusion models**

Chaterjee and Eliashberg (1990) develop a model of innovation diffusion using micromodeling approach that explicitly consider the determinants of adoption at the individual level in a decision analytic framework, incorporating heterogeneity in the population with respect to initial perceptions, preference characteristics, and responsiveness to information. More recently, Young (2009) examines three broad classes of diffusion models—contagion, social influence, and social learning—and shows how to incorporate heterogeneity into each at a high level of generality without losing analytical tractability. Lelarge (2012) studies a model of diffusion where the individuals behavior is the result of strategic choice, using simple coordination game with binary choice, and it describes the condition of a new action to become widespread in a random network.

Iyengar and Van den Bulte (2011) studies how opinion leadership and social contagion within social networks affect the adoption of new product, and find evidence of contagion operating over network ties, even after controlling for marketing efforts and arbitrary system wide changes. Moon and Russel (2008) develops a product recommendation mechanism based on principal of customer preference similarity stemming from prior purchase behavior is a key element in predicting customer product purchase. Song and Chitagunta (2003) develops an empirical model for the adoption process of a new durable product that accounts for consumer heterogeneity as well as consumers’ forward-looking behavior, and it decomposes the effect of the entry of Sony into primary demand expansion and switching from other brands, finding that there exist several segments in the market with different preferences for the brand and different price sensitivities leading to differences in adoption timing and brand choice across segments.

Dover et al. (2012) shows how networks modify the diffusion curve by affecting its symmetry, and it demonstrates that network’s degree distribution has a significant impact on the contagion properties of the subsequent adoption process, proposing an agent based method for uncovering the degree distribution of the adopters network underlying the dissemination process. Trusov et al. (2013) proposes an approach for identifying systematic conditions that are stable across diffusions, transferable to new product introduction within a given network, using Facebook applications data, , and it shows that incorporation of such systematic condition improves prelaunch forecast. For more complete review of models please review Peres et al. (2010).

This study is different from above mentioned studies in that it uses a unique long time series of a panel of consumers over a course of 11 month over free mobile apps, which allows us to identify to identify the diffusion of awareness and preferences of consumers in different geographical locations. In addition multiple app categories over a long time allow us to identify the relationship between different app categories. In addition, to the best of our knowledge we are the first to model the evolution of social preference in consumer choice.

**Model**

We observe choice data for a set of individuals. In contrast to the prior studies such as Carare (2011), which do not account for consumers’ decision within the app category, we start our model with a choice of app category. More specifically, we assume consumers go through two stages of decision making: First consumer selects an app category that she is interested in among the available app categories, where  denotes the outside good option, and then given the app category in the second stage consumer selects among available apps within the app category, where  denotes outside good option or any app other than the apps under this study. In this specification we assume that if a consumer selects an app category, she will definitely download one of the apps within that category. This modeling decision may be justified by referring to sunk opportunity cost of consumer for decision making, and tiny cost of download. More formally we can define the probability of purchase of an app in the category as:



Consumer information processing theory and household production model, assume that consumer considers perceived risk of a purchase decision, and she is more prone to select a product that she has more familiarity with. These studies including Ratchford (1982) and Kok and Xu (2011) suggest that consumer decision can be modeled using nested logit structure. Our different formulation of app choice probability by conditioning on category, allows us to have a more structural model, which in turn facilitates studying search behavior of consumers over different app categories on the app store. Our two stage modeling approach is close to the specification of Kok and Xu (2011) on consumer’s decision among product types in the first stage and among the brands for a given product type in the second stage. Figure below shows consumer’s decision to download an app. Another interpretation of our model is that first consumer faces a universal set which she uses a screening rule to filter. We assume the screening rule in our context is selecting the app category. Other screening rules could be to select one of the attributes as Gilbrid and Allenby (2004) suggest.

We specify the utility of consumer purchase among app categories in the app store in the following form:

 (1)

|  |  |
| --- | --- |
| Variable | Definition |
|  | Utility of consumer i, from purchasing in category j at time t |
|  | Consumer i’s perception of goodwill of app category j at time t |
|  | State variable that summarizes the history number of app’s that consumer i has downloaded in category j until now (to account for state dependency) |
|  | Variance of price of mobile app’s in app category j at time t. According to search theory, it indicates more uncertainty, so it should have negative effect. |
|  | Mean of file size of mobile app’s in app category j at time t (first for the limitation of mobile phone memory size, second as the proxy for complexity or quality of mobile app) |
|  | Mean of the tenure of mobile app’s in app category j (proxy for maturity of app’s in the category) |
|  | Total number of free mobile app’s available in app category j at time t, while search theory suggests lower is better, product differentiation suggests more is better. |
|  | Total number of paid mobile app’s available in app category j at time t. |
|  | Parameters to estimate |

We accounted for state dependence in our model by incorporating the history of consumer downloads of apps in categories j from inception until time t (). If consumer downloads an app from app category j in previous occasion, t-1, then  , otherwise if the consumer selects outside option, then  remains unchanged: . This specification induces a first-order Markov process on the choices. Furthermore, our incorporation of consumer i's perception of goodwill of app category j at time t () is an alternative approach to the approach used by Yang and Allenby (2003) to incorporate interdependence of awareness and preference (perhaps due to the social learning about ubiquitous good) of consumers.

Assuming that random utility term, , is type 1 extreme value distribution, consumer’s i's probability of selecting app category j at time t is given by a multinomial logit model as follows:

 (2)

Here, we assume that the mean utility of outside good is zero, .

As Erdem and Keane (1996) assumes that misperception error of consumer can be modeled with normal distribution, and many structural modelers follow the same approach e.g. Zhao et al. (2013), we also model misperception of consumers about goodwill of an app, assuming normal misperception error. More formally, we assume that individuals misperception of the goodwill of a mobile app category can be modeled as:

(3)

Where  and  are parameters to estimate, and  is a common factor across individuals. In order to identify the model we set .

Bass diffusion model (1969) suggests a specific hazard rate functional form for the diffusion of a product. Here we may be able to consider that goodwill of the app also follows the same form, as mobile apps diffuse as a result of the category need adoption. In addition, given the model of Naik, Prasad and Sethi (2008) and Simon and Horskey (1983), which incorporate competition to find the effect of complementarity or substitution effect, we also adopt the following form for the diffusion of each app category:

 (4)

|  |  |
| --- | --- |
| Variable | Definition |
|  | cumulative goodwill of each app category j at day t |
|  | market size of the app category j |
|  | error term of diffusion of category j at time t |
|  | Coefficient of substitution or complementarity of app category k and j |
|  | External market diffusion factor for app category j |
|  | Internal market diffusion factor for app category j |
|  | Variance parameter to estimate |

In summary equations (1-3) are our observation equation and equation (4) is our state equation of app category state space choice model.

We account for heterogeneity in individual and category parameters, through a random effect model. In order to avoid misspecification of the distribution of consumer heterogeneity in preferences, and allow for multimodal densities of preferences across groups (or “segments”) of consumers, we adopt the method proposed by Dube et al. (2010) to use a flexible semi-parametric heterogeneity specification consisting of a mixture of multivariate normal distributions, to allow for thick tail, skewness, and multimodality. We denote vector of non-state consumer-level parameters by  and vector of non-state category level parameters by . We accommodate consumer and category heterogeneity by assuming that  and are drawn from two separate distribution common across consumers and common across categories. We employ mixture of normal as the first stage prior, to specify an informative prior, avoiding the problem of over fitting. The first stage consists of a mixture of  multivariate normal distributions respectively, and the second stage consists of priors on the parameters of the mixture of normal:

 (5)

 (6)

 (7)

Where represents the hyper-parameters of the priors on the mixing probabilities and the parameters governing each mixture component. To obtain a truly non-parametric estimate using the mixture of normal model requires that the number of mixture components  increase with the sample size. We adopt the approach proposed by Dube et al. (2010) to fit models with successively large numbers of components and to gauge the adequacy of the number of components by examining the fitted density associated with each number of components.

In the second stage of decision making, consumer searches apps of their favor given the app category, so we specify consumer i’s utility from purchasing mobile app a in category j, at time t, according formally as:

 (8)

|  |  |
| --- | --- |
| Variable | Definition |
|  | Utility of consumer i, from purchasing app a in category j at time t |
|  | Consumer i’s perception of goodwill of an app a in app category j at time t |
|  | Dummy variable indicating whether app a is featured at time t |
|  | Tenure of mobile app a at time t(number of days from creation) |
|  | State variable that summarizes the history of consumer i download of app a in category j until now (to account for state dependency) |
|  | File size of mobile app a in app category j at time t (first for the limitation of mobile phone memory size, second as the proxy for complexity or quality of mobile app) |
|  | Parameters to estimate |

We accounted for state dependence in our model by incorporating the history of consumer i's downloads of apps a in category j from inception until time t (). If consumer downloads app a in category j in previous occasion, t-1, then  , otherwise if the consumer selects outside option, then  remains unchanged: . This specification induces a first-order Markov process on the choices. Like app category portion of our model, our incorporation of consumer i's perception of goodwill of app a in category j at time t is an alternative approach to the approach used by Yang and Allenby (2003) to incorporate interdependence of awareness and preference (perhaps due to the social learning about ubiquitous good) of consumers.

Assuming that random utility term, , is type 1 extreme value distribution, consumer i's probability of selecting app a in category j at time t is given by a multinomial logit model as follows:

 (9)

Here, we assume that the mean utility of outside good (app’s that we do not consider in our study) is zero, .

Like app category model, here we also assume that consumers have misperception about the real goodwill of each mobile app. Formally, we assume that consumer i's misperception of app a’s goodwill in category j, at time t can be modeled as:

(10)

Where  and  are parameters to estimate, and  is a common factor across individuals. In order to identify the model we set .

Similar to app category goodwill diffusion model and to account for complementarity and substitution effect between different app’s within the category we specify diffusion of goodwill of an app in the following form:

 (11)

|  |  |
| --- | --- |
| Variable | Definition |
|  | cumulative goodwill of each app a at day t |
|  | Market size of mobile app a in category j at time t |
|  | error term of diffusion of app a at time t |
|  | Coefficient of substitution or complementarity of app k and a |
|  | External market diffusion factor for app a |
|  | Internal market diffusion factor for app a |
|  | Variance parameter to estimate |

In summary equations (8-10) are our observation equation and equation (11) is our state equation in app state space choice model.

We account for heterogeneity in individual and app level parameters similar to the approach we used for app category. We denote vector of non-state consumer-level parameters by  and vector of non-state category level parameters by . We accommodate consumer and category heterogeneity by assuming that  and are drawn from two separate distribution common across consumers and common across apps. We employ mixture of normal as the first stage prior, to specify an informative prior, avoiding the problem of over fitting. The first stage consists of a mixture of multivariate normal distributions respectively, and the second stage consists of priors on the parameters of the mixture of normal:

 (12)

 (13)

 (14)

Whererepresents the hyper-parameters of the priors on the mixing probabilities and the parameters governing each mixture component. To obtain a truly non-parametric estimate using the mixture of normal model requires that the number of mixture components  increase with the sample size. Just like category level parameters, we adopt the approach proposed by Dube et al. (2010) to fit models with successively large numbers of components and to gauge the adequacy of the number of components by examining the fitted density associated with each number of components.

This closes our model specification. The Direct Acyclic Graph (DAG) of conditional distributions is presented in Appendix A.

**Extension**

To extend this research we plan to add supply side down the road, to understand parameters of choices of app-developers that determine her entrance into specific app-category. Another admissible extension is to model the diffusion of app and app category in groups of users with similar purchase behavior. In other word, at each point in time we can cluster users based on their purchase behavior and study how each mobile app and app category diffuses in each of those clusters. Another extension to this study is to model consumer learning in the category, as one may posit that consumers who purchase from a category may have less misperception about the category of product evolution, as more purchase may be correlated with more active search behavior. Another extension could be to find interaction between diffusion of different locations, identifying opinion leadership across locations. Another extension could be to consider that consumers are forward looking and form step-ahead expectation about the goodwill of a mobile app and the mobile app category.

**Data**

The data were collected by an African telecom operator on individual choice of downloading mobile apps from the app store platform of its global partner. The app-store is launched within around 350 days prior to our study, so we use around 150 daily download choice of consumers over this newly launched app store, from the time that dominant app in each selected category is launched. We selected 10 different app categories, and within each category we selected dominant apps, which have significantly greater number of downloads than other apps within a given category during the period of study, total of 11 dominant apps across all categories in the sample. This choice allows us to assume that the app’s are in an monopolistic competition market, as the choice set of consumers consists of either dominant app or another app within the category. We obscure the identity of some apps for the purpose of confidentiality. Following table shows the list of the categories that we select and their corresponding apps:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| index | Category | Downloads in the category | Number of apps in our sample | Name of the apps  (total downloads) |
| 1 | Dating | 27 | 1 | Adult Friend Finder (25) |
| 2 | eBook | 414 | 2 | Open Heavens 2014 (335)  Okadabooks(50) |
| 3 | Education & Learning | 24 | 1 | Gidimo (22) |
| 4 | Health/Diet/Fitness | 42 | 1 | Doctor(30) |
| 5 | Internet & WAP | 52 | 1 | Operator’s info App(44) |
| 6 | Movie/Trailer | 597 | 1 | Dobox(584) |
| 7 | POI/Guides | 22 | 1 | Nearest Locator(20) |
| 8 | Reference/Dictionaries | 55 | 1 | Fact about Life and Physics (53) |
| 9 | TV/Shows | 135 | 1 | Live Mobile TV(114) |
| 10 | Video & TV | 105 | 1 | YouTube Downloader (80) |

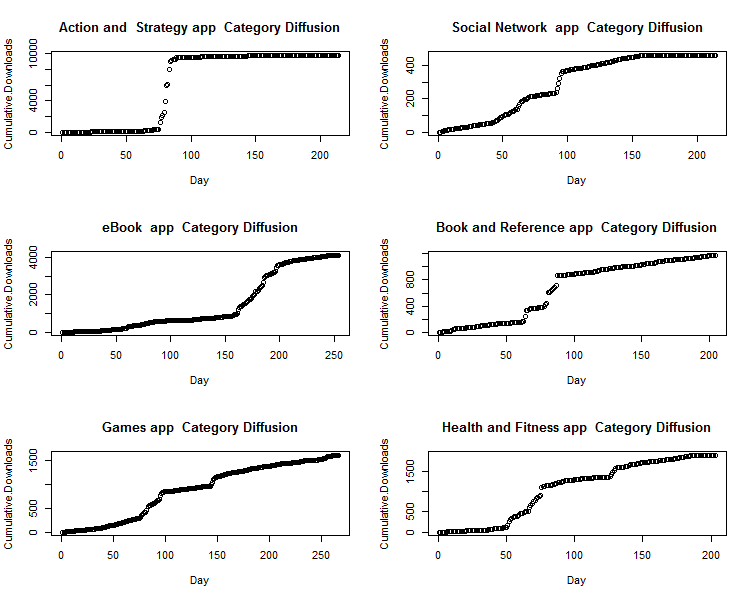
We investigate geographic source of dependence of awareness and preference of consumers through good will of the mobile app. Geographic neighborhood is created by physical proximity, through mapping IP address of each consumer to the corresponding city. For this purpose we wrote a crawler and scraper in Perl to scrape the location of each corresponding IP in an IP-location mapping website[[5]](#footnote-5). We consider only one city in Africa, in this study, so all our basic statistics is also relevant to that city.

Our data also included longitude and latitude of each IP address, yet as mobile phones are nomadic and ubiquitous, we find the choice of city a more reasonable one. Social nature of mobile phone and mobile apps is another reason which may reinforce that our assumption is innocuous. As mobile phone has become inseparable part of humanity, to the point that not only they use it when they are alone in bus, before sleeping, or even in the class, but also they use it in parties, in office, and in the leisure time, and generally in any social events. Usage in social events makes mobile apps visible, so it creates opportunity of learning in many social events.

The random sample consists of around 1,300 consumers, living in the selected city, and with around 1,880 unique daily downloads for the course of around 150 days in 2013 and 2014. Following table shows the basic statistics of data at mean of category level:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Category Data Summary | Mean | Variance | Min | Max |
| Number of available apps in the Category | 35 | 1250 | 12 | 141 |
| Average Tenure of App’s in the category (Days) | 316 | 6,386 | 169 | 498 |
| Number of available free apps in the category | 32 | 908 | 7 | 120 |
| Average day’s that an app is featured in the category | 0.12 | 0.05 | 0.00 | 0.71 |
| Average file size of app’s in the category (MB) | 2 | 4 | 0 | 8 |
| Variance of prices of app’s in the category | 0.51 | 1.09 | 0.00 | 3.75 |

In the following figure we draw diffusion curve of cumulative adoption of each sample of 6 app categories:



For each user we have created time, and number of downloads. This allows us to distinguish between heavy and light users, so we can probably use it to explain heterogeneity of users. For each download we know source of download, whether the mobile app is featured and the device name. For each app category and app we can find the target demographic, so it may be helpful to explain heterogeneity in terms of these data.

**Identification and Estimation**

In order to identify our choice model we use logit model, which has a fix scale. To set the location of the utility we normalize the utility of outside good to zero, as explained before. We identify individual level parameters using large panel of individuals we have in our sample. We identify the goodwill of mobile app category and mobile app through the diffusion structure we impose. We can identify mobile app category, and mobile app parameters through shrinkage and variation of different app categories, and mobile apps in our data, and finally we identify state space parameters through variation through long time series in our panel data.

We use hybrid of Metropolis-Hasting (MH), Gibbs sampling (GS), and Extended Kalman Filter (EKF) forward filtering backward sampling techniques for our estimation. Our algorithm consists of three stages, and we parallelized the algorithm for MH and GS of individual choices, and EKF and GS of app category and mobile apps. Our algorithm consists of “six” stages as presented in Appendix B.

The first stage of the MCMC is a set of I parallel Metropolis Hasting algorithms tuned for each consumer MNL likelihood on Metropolis proposal distributions. We extend the Metropolis Hasting Random Walk (MH-RW) method proposed by Rossi, Allenby and McCulloch (2005), chapter 5 to optimize the proposal density to identify latent individual perception of the diffusion of an app and an app category. We first draw non-state parameters from the proposal density using MH-RW, and then conditional on the non-state parameters we draw the latent state around the mode, using Metropolis Hasting Independent Chain (MH-IC) . We code the tune to be done automatically. This procedure avoids the problem of undefined likelihood for tuning purposes as suggested by Rossi, Allenby and McCulloch (2005).

In the second stage we use standard unconstrained Gibbs Sampler for a mixture of normal to find the hyper parameters of individual level coefficients. The “label-switching” problem for identification in mixture of normal is not relevant in our application, as we are interested only in the joint posterior distribution of quantity which is label-invariant, i.e. mixture of normal density itself. For drawing from multinomial distribution we have the following conditional probabilities:

 (15)

 (16)

Where  are discrete random latent variables with outcome probabilities and. This conditioning allows us to draw from mixture of normal distribution. Similarly we have conditions to draw from mixture of normal distribution for other category, app, and individual level parameters we specified conditional distributions in Appendix B..

Third stage of our estimation involves a Gibbs sampler to shrink individual’s perception of goodwill of mobile app and app category misperception toward true latent measure of goodwill of an app and an app category. Fourth stage of our estimation involves Extended Kalman Filter to draw unobserved state of shrunk goodwill of app and app category, by linearizing Bass diffusion model (BDM), and finally fifth stage of our estimation includes another standard unconstrained Gibbs sampler for mean and variance of state equation of goodwill of an app and an app category.

MCMC that we apply here is a hybrid method with a customized Metropolis chain for the draw of the consumer, category, and app level parameters coupled with a standard Gibbs sampler for a mixture of normal conditional on the draws of consumer, category and app’s parameters. In other word, once the collection of consumer, category, and app parameters are drawn, the MCMC algorithm treats these as “data” and conducts Bayesian inference for a mixture of normal.

**Counterfactual analysis**

We analyze “what if” scenarios to find optimal incentive policy of platform owner for different product categories to increase the diffusion of the platform. This incentive policy could be in the form of lower revenue share for specific category. In addition, for platform owner it may be relevant to know what is the optimal assortment size to generate more revenue in a category, so we simulate different scenarios to find optimal policy. Third paid apps may cannibalize the free apps, or vise versa, so it may be interesting to know what is the optimal combination of free and paid apps that result in higher diffusion level. Another counterfactual could study how removing one app category affect consumer’s choice. Platform owner may be interested to know the effect of specialization or differentiation, and the optimal structure of the app store. To study these issues we use indirect utility of consumer, by numerically integrating the utility of individual consumer across population to get the expected market size.

**Conclusion**

We model consumer’s decision to download from mobile app store, focusing on complementarity and substitution of app categories and social dependency of awareness and preferences of consumers. We used a unique panel data to estimate our model. Our model explains the level of substitution and complementarity of the mobile app categories, and it investigates the level of social dependency of preference and awareness of consumers in a geographical area. We run counterfactual to understand how incentivizing one category can impact the total diffusion of the platform, and to know what is the optimal assortment size and optimal category hierarchy on the app store. Our analysis also allows us to find the impact of paid apps on the diffusion of free apps. We believe our study has various managerial implications.

**Appendix A: Direct Acyclic Graph of Conditional Distributions**













































Mobile app

Mobile app category



**Appendix B: Conditional Distributions for Estimation**

Following we present list of conditional distributions for our estimation:

 (B1)

 (B2)

 (B3)

 (B4)

 (B5)

 (B6)

Where and  are the matrix of covariates.

The prior used are:



We assessed the prior hyperparameters to provide proper but diffuse distributions:



The Dirichlet prior on  is conjugate to the multinomial distribution, and can be interpreted as the size of a prior sample of data for which the classification”observations” is known. The number of observation of each “type” or mixture component is given by the appropriate element of . This prior assumes that each type is equally likely and that there is only a very small amount of information in the prior equal to a sample size of .5. As the number of normal components increases, we do not change how informative the prior is, through scale of elements of the a vector by . As our model become more complicated, we parallelized the computation of many stages as much as possible to use multi-core of server for our estimation.

**References**

Bell, David R., and Sangyoung Song. "Neighborhood effects and trial on the Internet: Evidence from online grocery retailing." Quantitative Marketing and Economics 5.4 (2007): 361-400.

Bezawada, R., Balachander, S., Kannan, P. K., & Shankar, V. (2009). Cross-category effects of aisle and display placements: A spatial modeling approach and insights. *Journal of Marketing*, *73*(3), 99-117.

Bradlow, Eric T., et al. "Spatial models in marketing." *Marketing Letters* 16.3-4 (2005): 267-278.

Carare, Octavian. "THE IMPACT OF BESTSELLER RANK ON DEMAND: EVIDENCE FROM THE APP MARKET\*." *International Economic Review* 53.3 (2012): 717-742.

Chatterjee, R. A., & Eliashberg, J. (1990). The innovation diffusion process in a heterogeneous population: A micromodeling approach. *Management Science*,*36*(9), 1057-1079.

Jaihak Chung, Vithala R. Rao (*2012*) A General Consumer Preference Model for Experience Products: Application to Internet Recommendation Services. Journal of Marketing Research: June 2012, Vol. 49, No. 3, pp. 289-305.

Choi, Jeonghye, Sam K. Hui, and David R. Bell. "Spatiotemporal analysis of imitation behavior across new buyers at an online grocery retailer." *Journal of Marketing Research* 47.1 (2010): 75-89.

Dover, Yaniv, Jacob Goldenberg, and Daniel Shapira. "Network traces on penetration: Uncovering degree distribution from adoption data." *Marketing Science* 31.4 (2012): 689-712.

Dubé, J. P., Hitsch, G. J., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*,*41*(3), 417-445.

Erdem, Tülin, and Michael P. Keane. "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets." *Marketing science* 15.1 (1996): 1-20.

Oestreicher-Singer, Gal, and Arun Sundararajan. "The visible hand? Demand effects of recommendation networks in electronic markets." *Management Science* 58.11 (2012): 1963-1981.

Ghose, Anindya, Avi Goldfarb, and Sang Pil Han. "How is the mobile Internet different? Search costs and local activities." *Information Systems Research*24.3 (2012): 613-631.

Ghose, Anindya, and Sang Pil Han. "Estimating Demand for Mobile Applications in the New Economy." *Management Science* (2014).

Ghose, Anindya, and Sang Pil Han. "An empirical analysis of user content generation and usage behavior on the mobile Internet." *Management Science*57.9 (2011): 1671-1691.

Ghose, Anindya, and Sang Pil Han. "A dynamic structural model of user learning on the mobile Internet." *Available at SSRN 1485049* (2011).

Ghose, Anindya, and Sang Pil Han. "A dynamic structural model of user learning on the mobile Internet." *Available at SSRN 1485049* (2011).

Gilbride, T. J., & Allenby, G. M. (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, *23*(3), 391-406.

Hartmann, Wesley R. "Demand estimation with social interactions and the implications for targeted marketing." *Marketing Science* 29.4 (2010): 585-601.

Hartmann, Wesley R., et al. "Modeling social interactions: Identification, empirical methods and policy implications." *Marketing letters* 19.3-4 (2008): 287-304.

Horsky, Dan, and Leonard S. Simon. "Advertising and the diffusion of new products." *Marketing Science* 2.1 (1983): 1-17.

Iyengar, Raghuram, Christophe Van den Bulte, and Thomas W. Valente. "Opinion leadership and social contagion in new product diffusion." *Marketing Science* 30.2 (2011): 195-212.

Didem Kurt, J. Jeffrey Inman, Jennifer J. Argo (*2011*) The Influence of Friends on Consumer Spending: The Role of Agency– Communion Orientation and Self-Monitoring. Journal of Marketing Research: August 2011, Vol. 48, No. 4, pp. 741-754.

Karaca-Mandic, Pinar. "Role of complementarities in technology adoption: The case of DVD players." *Quantitative Marketing and Economics* 9.2 (2011): 179-210.

Kim, Y., Telang, R., Vogt, W. B., & Krishnan, R. (2010). An empirical analysis of mobile voice service and SMS: a structural model. *Management Science*,*56*(2), 234-252.

Lee, Sanghak, Jaehwan Kim, and Greg M. Allenby. "A direct utility model for asymmetric complements." *Marketing Science* 32.3 (2013): 454-470.

Lelarge, Marc. "Diffusion and cascading behavior in random networks." *Games and Economic Behavior* 75.2 (2012): 752-775.

Aurélie Lemmens and Christophe Croux (2006) Bagging and Boosting Classification Trees to Predict Churn. Journal of Marketing Research: May 2006, Vol. 43, No. 2, pp. 276-286.

Liu, Charles Z., Yoris A. Au, and Hoon Seok Choi. "An Empirical Study of the Freemium Strategy for Mobile Apps: Evidence from the Google Play Market." (2012).

Liu, Hongju, Pradeep K. Chintagunta, and Ting Zhu. "Complementarities and the demand for home broadband internet services." *Marketing Science* 29.4 (2010): 701-720.

Ma, Yu, P. B. Seetharaman, and Chakravarthi Narasimhan. "Modeling dependencies in brand choice outcomes across complementary categories."*Journal of Retailing* 88.1 (2012): 47-62.

Manski, Charles F. "Identification of endogenous social effects: The reflection problem." *The review of economic studies* 60.3 (1993): 531-542.

Garg, Rajiv, and Rahul Telang. "Inferring app demand from publicly available data." *MIS Quarterly* 37.4 (2013): 1253-1264.

Naik, Prasad A., Ashutosh Prasad, and Suresh P. Sethi. "Building brand awareness in dynamic oligopoly markets." *Management Science* 54.1 (2008): 129-138.

Harikesh S. Nair, Puneet Manchanda, Tulikaa Bhatia (*2010*) Asymmetric Social Interactions in Physician Prescription Behavior: The Role of Opinion Leaders. Journal of Marketing Research: October 2010, Vol. 47, No. 5, pp. 883-895.

Narayan, Vishal, Vithala R. Rao, and Carolyne Saunders. "How peer influence affects attribute preferences: A Bayesian updating mechanism." *Marketing Science* 30.2 (2011): 368-384.

Niraj, Rakesh, V. Padmanabhan, and P. B. Seetharaman. "Research Note-A Cross-Category Model of Households' Incidence and Quantity Decisions."*Marketing Science* 27.2 (2008): 225-235.

Peres, Renana, Eitan Muller, and Vijay Mahajan. "Innovation diffusion and new product growth models: A critical review and research directions." *International Journal of Research in Marketing* 27.2 (2010): 91-106.

Rossi, P. E., & Allenby, G. M. R. McCulloch (2005). *Bayesian Statistics and Marketing*.

Yang, Sha, Vishal Narayan, and Henry Assael. "Estimating the interdependence of television program viewership between spouses: A Bayesian simultaneous equation model." *Marketing Science* 25.4 (2006): 336-349.

Moon, Sangkil, and Gary J. Russell. "Predicting product purchase from inferred customer similarity: An autologistic model approach." *Management Science*54.1 (2008): 71-82.

Yang, Sha, and Greg M. Allenby. "Modeling interdependent consumer preferences." Journal of Marketing Research 40.3 (2003): 282-294.

Simon, Herbert A. "Theories of bounded rationality." Decision and organization1 (1972): 161-176.

Song, I., & Chintagunta, P. K. (2003). A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category. *Quantitative Marketing and Economics*, *1*(4), 371-407.

Song, Inseong, and Pradeep K. Chintagunta. "Measuring cross-category price effects with aggregate store data." *Management Science* 52.10 (2006): 1594-1609.

Sriram, S., Pradeep K. Chintagunta, and Manoj K. Agarwal. "Investigating consumer purchase behavior in related technology product categories."*Marketing Science* 29.2 (2010): 291-314.

Andrew T. Stephen, Olivier Toubia (2010) Deriving Value from Social Commerce Networks. Journal of Marketing Research: April 2010, Vol. 47, No. 2, pp. 215-228.

Michael Trusov, William Rand, and Yogesh V. Joshi (*2013*) Improving Prelaunch Diffusion Forecasts: Using Synthetic Networks as Simulated Priors. Journal of Marketing Research: December 2013, Vol. 50, No. 6, pp. 675-690.

Wedel, Michel, and Jie Zhang. "Analyzing brand competition across subcategories." *Journal of Marketing Research* 41.4 (2004): 448-456.

Yang, Botao, and Andrew T. Ching. "Dynamics of consumer adoption of financial innovation: The case of ATM cards." *Management Science* 60.4 (2013): 903-922.

Young, H. Peyton. "Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning." *The American economic review* (2009): 1899-1924.

Zhao, Yi, et al. "Modeling consumer learning from online product reviews."*Marketing Science* 32.1 (2013): 153-169.

1. I would like to thank Dr. Soysal and Dr. Ratchford for sending me to Duke workshop on structural modeling, and teach me related techniques. Also I would like to thank Dr. Bruce for teaching me how to conduct a quality empirical research in marketing, and code Kalman Filter. [↑](#footnote-ref-1)
2. I would like to thank Dr. Ying Xie and Dr. Elisabeth Honka for the revision of the proposal I send out to the company. I also would like to thank Dr. Rao for making us to write a lot during his seminar, allowing us to organize our mind, and positioning our research paper. [↑](#footnote-ref-2)
3. This process has been so extensive that some scholars study how cell phone affects social interaction of individuals http://www.michigandaily.com/news/%E2%80%98u%E2%80%99-researchers-identify-link-between-cell-phones-and-socialization-habits [↑](#footnote-ref-3)
4. http://www.businessweek.com/articles/2014-02-20/uber-leads-taxi-industry-disruption-amid-fight-for-riders-drivers [↑](#footnote-ref-4)
5. http://www.iplocation.net/index.php [↑](#footnote-ref-5)