Idea Document

Complementarity and Substitution of App-Category in App-Store:

Empirical Study of Nigerian App-store

Author: Meisam Hejazi Nia

Supervisor: Dr. Norris Bruce

Date: 05/13/2014

**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. For the platform owner that seeks the diffusion of its platform, it is important to understand which app-category can drive the diffusion of its platform, and what app-categories are substitutes. In this study we develop a structural model that seeks to find the level of substitution and complementarity between different app-store app-categories. We test our model using 9 month unique panel data of consumer purchases. We use state space model and hierarchical Bayesian to model consumer’s choice. Our counterfactual analysis of the incentive of the firm for different product categories has managerial implication for app-store platform owners.

**Model**

As currently we do not have data for the time of each purchase, we use more structure to identify our model. 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 such as Ratchford (1982) and Kok and Xu (2011) suggest that consumer decision may follow nested logit structure. Parallel with these studies search theories in computer science suggest that there are two solutions in searching solution tree for a problem: depth first, and breath first. On the other hand, greedy algorithm in computer science literature, and optimal control theory in management science literature, and search theory in economics literature, suggest that if consumer faces an uncertain life time, and uncertain consumption, she may select a choice that gives her highest utility among all the available choices. Therefore, we model consumer’s decision in two steps. In the first step, consumer selects the category from which she will purchase, and in the second one, she will select among all the available categories which product to purchase. This 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. Therefore, 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 |
|  | Goodwill of app category j at time t-1 |
|  | Mean of price of mobile app’s in app category j at time t |
|  | Variance of price of mobile app’s in app category j at time t |
|  | 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 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 mobile app’s available in app category j at time t. |
|  | Parameters to estimate |

Therefore, we can write the probability that consumer i selects product category j at time t that she is created in the following form:

 (2)

Therefore, we can write the expected total diffusion of category j at time t in the following form:

 (3)

Where is a dummy variable that indicate whether consumer i purchases product category j either at the time of its creation on the app store platform or later. At this stage as a result  can be considered known, or data if we condition on . Therefore, we define the observation equation for latent diffusion of app category users in the following form:

 (4)

In contrast to prior studies such as Carare (2011) do not account for consumers’ decision within the category, we start our model with product category. In this sense, we assume that each product category has a model of bass diffusion, or formally

 (5)

|  |  |
| --- | --- |
| Variable | Definition |
|  | cumulative number of users of each app category j at day t on the platform |
|  | market size of the app category j |
|  | error term of diffusion of category j at time t |
|  | Variance of price of mobile app’s in app 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 |

Our incorporation of the sum of number of users of different product categories (sigma term), allows us to account for complementarity and substitution of different product categories. Needless to say we account for heterogeneity in these parameters, and we later explain this heterogeneity in terms of characteristics of each mobile-app category (i.e. type of service and demographic of users).

Consumer search apps in the product category, given that she has selected the product category, and based on her utility she decide whether to purchase and which app to purchase given they have decided to purchase, therefore, we define consumer i’s utility from purchasing mobile app a, at time t, according formally as:

 (6)

|  |  |
| --- | --- |
| Variable | Definition |
|  | Expected cumulative number of mobile app a users at time t-1 |
|  |  |
|  | Price of mobile app a 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) |
|  | unobserved demand shock (market time dummy) |
|  | the parameters to estimate |

We allowed heterogeneity in these parameters, and we later on explain this based on individual characteristics, such as download counts and tenure, and region (IP) of the downloader.

The logit form of the error term of the utility, allows us to define purchase probability of consumer in the following form:

 (7)

where  is the number of mobile apps in category j.

As we do not know which of the mobile apps has been selected as a first mobile app of given consumer i, again we calculate expected probability that consumer i, selects app a, at time t, in the following form:

 (8)

Where is a dummy variable that indicates whether consumer i purchases mobile app a either at the time of its creation on the app store platform or later. At this stage as a result  can be considered known or data, if we condition on . Therefore, we define the observation equation for latent diffusion of mobile app users in the following form:

 (9)

The next portion of our model accounts for evolution of preference for each mobile app. To model this evolution of preference for the app, we use the following Bass structure as well:

 (10)

|  |  |
| --- | --- |
| Variable | Definition |
|  | Cumulative number of mobile app a users at day t |
|  | The market size of mobile app-a in category j at time t, which is Proportional to the market size of the category |
|  | External market force of diffusion |
|  | Internal market force of diffusion |
|  | Parameter to estimate |

Needless to say allow for heterogeneity in these parameters, and further on we explain this heterogeneity in terms of observed characteristics of mobile app (i.e. file size as a signal of quality and sophistication of the mobile app).

This closes our model specification.

**Extension**

We later on can add the supply side to the model to account for app-developer/publisher’s decision to select the product category to develop mobile app.

**Data**

We have the 9 month daily purchase of consumer on the whole platform. For each user we have created time, and number of downloads. This allows us to distinguish between heavy and light users, and the experience of user with the platform owner. For each purchase we know the price, the number of downloads, source of download, whether the mobile app is featured and the device name. This allows us to explain heterogeneity in each user purchase decision. For each app we also know the category, language and file size which allow us to explain heterogeneity.

**Estimation**

We use Metrapolist-Hasting, Hierarchical Bayesian, and Kalman filter forward filtering backward smoothing for our estimation.

**Conclusion**

We model consumer’s decision to purchase from mobile app category. We used the panel data to estimate our model. Our model explains the level of substitution and complementarity of the mobile app categories. We run counterfactual to understand how incentivizing one category can impact the revenue of the platform. We believe our study has various managerial implications.