# Designing Freemium: a Model of Consumer Usage, Upgrade, and Referral Dynamics

Clarence Lee, Vineet Kumar and Sunil Gupta*

November 2013, Preliminary and Incomplete

**Abstract**

Over the past decade "freemium" (free + premium) has become the dominant business model among internet start-ups for its ability to acquire and monetize a large install-base with limited marketing resources. Freemium is a hybrid strategy where a firm offers both a perpetually free but limited version of their service, and a premium version with enhanced features that require a fee, and where firms regard the free product as a promotional tool. The model leads to several questions interesting to marketers, which we explore in our framework. What is the right referral bonus incentive to offer to customers? When consumers are connected to others, how can we appropriately characterize the value of "free" customers? How does sharing influence customers' likelihood of upgrading to the premium product? 1) How much should be spent to acquire free consumers? (Value of a Free Consumer) 2) What is the appropriate level of incentives to encourage adoption and how much do the referrals account for the value of the consumer? (Value of Referrals) 3) How should firms dynamically design referral incentives to maximize upgrade and usage behavior? (Dynamic Design of Referral Incentives) We develop an empirical microfoundations-based framework to understand dynamics of consumer behavior of plan choice, usage, and referral in the freemium setting and apply it to a novel panel data set from a leading cloud-based storage service. Using Bayesian methodology, we estimate the structural model and perform counterfactual analysis. We find that the value of free consumers is approximately $24 per year, and that the existence of the referral program contributes to 65% of this value – signifying the importance of the referral program. We explore counterfactuals to maximize the average consumer referral rate by changing the referral incentives. Contrary to the belief that more is better, we find the existence of an optimal incentive point for referrals. Thus, we are able to characterize both the individual value of consumers to the firm as well as the network value of customers, providing a mechanism to capture the impact of consumer-to-consumer interactions.

**Keywords:** *Discrete-Continuous Choice Dynamic Structural Models; Freemium; Entrepreneurship*

*Clarence Lee is a doctoral candidate at the Marketing Unit at Harvard Business School (clee@hbs.edu). Vineet Kumar is an Assistant Professor of Business Administration at Harvard Business School (vkumar@hbs.edu). Sunil Gupta is the Edward W. Carter Professor of Business Administration at Harvard Business School (sgupta@hbs.edu).

# 1 Introduction

Over the past decade, several software companies have increasingly turned to the subscription model for revenue generation. Firms often offer a limited but *perpetually free* version of their software in order to rapidly develop a large consumer install-base, with the expectation that users will upgrade to the paid premium version. This business model – referred to by industry as "freemium," a hybrid of free and premium – has been successfully adopted in Silicon Valley, and largely popularized among the newer generation of start-ups. According to the New York Times, freemium is one of the most prevalent business models among Web start-ups because relying on advertising as the sole stream of revenue might not be sufficient or sustainable.[1] To date, over 80% of the top grossing iOS apps have adopted the freemium model, with the largest freemium start-ups having acquired over hundreds of millions in venture funding.[2][3] The success of this business model has been further validated by many of the largest companies across multiple digital sectors, from online social networking sites such as LinkedIn, to music services such as Pandora and Spotify. Even media companies such as the New York Times (and its online pay wall) and mobile payment companies like Paypal utilize freemium. In the offline context, some consumer banks can also be characterized as using this model, with free checking accounts along with premium relationship accounts comprising the differentiated product offerings.

Freemium is often adopted because of its ability to attract a large number of users to its free version, i.e. as a customer acquisition strategy. Start-up technology companies facing capital constraints often choose this strategy over investing in advertising or using a sales force to obtain new customers. When coupled with a powerful consumer-to-consumer referral invite program, its effectiveness in acquisition is often magnified since a free product is easier to recommend. As a result, companies using this strategy see that a large percentage, often as high as 90% or more, of their consumer base are free users who do not contribute directly to the firm's revenues. While attracting a large user base is vital for establishing company value, firms must generate revenue for sustainability. Hence, the challenge requires balancing dual tasks: growing the consumer base by offering a free service and maintaining premium services to incentivize upgrades in order to stay profitable (Needleman and Loten, 2012).

The freemium business model raises several important questions in marketing that we investigate in this study, which uses a data set from a leading online file synchronization, backup and sharing service as its focus. Our research questions in this context include the following:

1. **Referrals:** What is the right referral bonus incentive to offer to consumers? Should the firm limit the number of referrals? Are referral bonuses complements or substitutes to upgrade decision? Should the referral bonus increase or decrease over time? Should the firm set consumer expectations by announcing a change in referral

[1] "Ad Revenue on the Web? No Sure Bet," The New York Times, May 25, 2009. http://www.nytimes.com/2009/05/25/technology/start-ups/25startup.html.

[2] "Freemium apps continue to flourish in 2012." IntoMobile, December 22, 2011. http://www.intomobile.com/2011/12/22/freemium-apps-continue-flourishing-2012/.

[3] TechCrunch Crunchbase for Evernote, Pandora, 37 Signals, Spiceworks, and Dropbox, accessed June 5, 2013. http://www.crunchbase.com/company/.

bonus?

2. **Customer Value:** While free customers do not provide any direct revenue to the company, they have potential to generate revenue by either upgrading in the future, or by referring new customers who may upgrade. How can we then value these "free" customers? Given the interconnected network of consumers, how can we characterize customer network value? Who are the most valuable customers and what is their network value?

3. **Shared Product Use:** How does sharing influence customers' likelihood of upgrading to the premium product? Should the firm subsidize shared activities to incentivize consumers to increase their marginal use, leading to positive externalities on other consumers, or should the firm attempt to extract correspondingly more value from sharing activities? How does this relate to the cost of shared activities?

We use a unique panel data set of consumer activities to examine these questions relating to the freemium setting. There are multiple sources of value consumers obtain from the service. First, their files in their accounts are synchronized immediately across all connected devices, including computers, mobile phones and tablets. Second, the files are backed up in the firm's online storage repositories, and accessible from any Internet-connected computer using a web interface. In the course of using the service, consumers add, delete and share files and also refer other consumers to the service; however, the primary revenue generating activity is when consumers upgrade from a free to a premium account, allowing for more storage capacity.

We develop a framework to characterize the dynamic behaviors of consumers in this setting, accounting for their motivations to undertake these various activities in their accounts. Consumers using the free version of the product in each period choose whether to upgrade to an annual or monthly plan that provides them an increased storage quota. In addition to this decision, they also have the choice of referring a friend, and obtain a referral bonus quota if the friend becomes a customer of the firm by subscribing to either the free or premium service. Consumers can also choose to delete files to maintain the limited space in their account, freeing up storage for future use. Thus, in the model consumers obtain a *flow utility* from the amount of storage current used, as well as decision or *action utilities* corresponding to the addition, deletion or sharing of files, and face a potential disutility related to the decision to upgrade their service by paying a price. Note that the benefits of upgrading accrue over time, since the action increases the constraint on storage from the free quota (2 GB) to the premium quota (50 GB). In this setting, inter-temporal dynamics and trade-offs play a very significant role, since the consumer has to predict future usage (and available storage) in determining the tradeoff of current decisions on upgrading, deleting or referring friends weighted against the costs of those decisions. Thus, we model consumers to be forward-looking, in order to trade off the cost of upgrading to a premium plan with the cost of finding and determining older files to delete, when newer content needs to be synchronized over time, as well as the likelihood that they would hit the limit of the free product. Consumers refer friends to the service, and while they receive a referral bonus when the friend joins, the referring consumer is not able to control the timing of joining, and thus forms an expectation over how many of her referred friends might join during each usage period. Our microfoundations-based model thus incorporates discrete and continuous choices for consumer upgrade, usage, and referral behaviors.[4]

The estimation of our dynamic structural model involves several computational challenges, given that the state space has both continuous (amount of usage) and discrete (type of plan, referrals accepted by friends) dimensions. In addition, whereas upgrading and referral behavior are discrete choices, the amount of files to delete to create free storage space is a continuous decision, complicating the modeling and estimation process. We find that our likelihood function is highly irregular and jagged, making it important to use a robust method to obtain the global maximum. We use a conjunction of different approaches to overcome these computational and estimation-related challenges. First, we use a Bayesian methodology, using a modified version of the Imai-Jain-Ching (IJC) algorithm (Imai et al., 2009) that helps deal with the complex, highly irregular likelihood function. Second, we make extensive use of quadrature approaches to computing integrals for the likelihood to improve accuracy and computational time. Finally, to deal with the constrained continuous decision, we use analytical inversion to obtain the exact value of the unobservable shock corresponding to the decision using a stochastic Euler-equation based approach. In contrast, the grid inversion technique used by Timmins (2002) is not only more computationally demanding, but depends significantly on the accuracy of the discretization, and its use in a setting with highly non-monotonic likelihood like ours could be problematic.

We find that consumers on average obtain significant flow utility from having an amount of storage to synchronize and back up their files, which is expected since it is the primary value of the service. They also have a high and convex cost of deleting files, likely from being able to pick appropriate files that are no longer needed in order to maintain sufficient free storage capacity for future usage. Consumers also have a negative utility, or a cost of referring friends to the service, and weigh that against the probability that the referral will be accepted as well as the firm's offered amount of increase in baseline quota from the referral bonus incentive.

With these estimates, we then simulate counterfactuals that help deconstruct the consumer value to consumer personal usage and referral behavior, and in turn, we examine the impact of changing various design policies on consumer value. We find that the lower-bound estimate of the value of a free consumer is approximately \$2. 35% of this value is attributed to the purely personal usage aspect of the service, and 27% of the value can be attributed to purely referral aspect of consumers. A more unexpected finding lies in early evidence of a referral-personal usage synergistic effect, in which the existence of a referral program actually encourages consumers to store more files on average, and thereby increases the probability that a consumer will upgrade at any given period. This aspect is attributed to 38% of the free consumer value.

[4]While we currently account for the value from personal usage and WOM referral, we can establish a lower-bound value of the consumer value, and will extend these results to incorporate the tradeoffs associated with social sharing in this work.

Examining the impact of changing referral incentives is crucial because it can change the speed of product adoption, and therefore help the firm rapidly reach a critical mass of install-base. Even without considering the cost of supporting free consumers, we find that giving away too much referral incentive may actually decrease the overall output of referrals. If a consumer can receive the same amount of bonus space for one referral, which is sufficient for use, then what motivation is there to send out another two or three? Furthermore, even if an optimal referral amount exists, the firm is not limited to statically changing the incentives for perpetuity. If the firm's object is to maximize growth, but can only support enough referral bonus for a limited amount of time, should the firm choose a ramp-up or ramp-down strategy in terms of releasing the bonus incentive? Should the firm announce the bonus ahead of time, encouraging consumers to be anticipate the change? These are important questions without clear intuition but can be answered with counterfactual simulations. We find that the shape of the consumer response of referral incentives is an inverted-U, implying that the firm should neither offer too small of an incentive (MB) because consumers may not find it worthwhile to refer anyone, nor too large of an incentive, because it may limit consumers' motivation to send out higher numbers of referral invites. We find the optimal static incentive amount to be approximately 500MB, which is double the amount observed in our data.

From, a managerial perspective our findings have several implications. The existence of a large proportion of free consumers makes it difficult to assess firm value and future potential for a start-up entrepreneurial firm: the firm observes zero cash flow from free consumers, making it impossible to accurately project the future stream of cash flow for a majority of the consumer base. Without an accurate understanding of consumer value, it is difficult to price the product. Additionally, firms are often reluctant to drastically change the price of premium plans in fear of initiating backlash from existing consumers. Therefore, at best, firms can run small-scale pricing experiments on a limited subset of its consumers in a static setting in order to inform price change. However, because the profitability of these services depend heavily on repeated consumer visits and usage of features, without a model of consumer behavior, it is difficult to predict how consumers will respond in usage behaviors in order to compensate for the change in price. To further complicate the question of pricing, the existence of referrals and social features actually links the behavior of consumers together, therefore making it even more difficult to account for these factors in a pure experimental setting.

Taking a narrow view of the consumer can be highly inaccurate. Past research has shown that it is more expensive to acquire a new consumer than to maintain a current one. Firms that assign negligible worth to free consumers risk losing an important opportunity to maximize benefit from the self-perpetuating consumer base inherent to the freemium model, especially because it takes a long time for a free consumer to upgrade to the premium product. The value of the free consumers can be derived from three possible areas. The first area is their personal usage level. The more that a consumer uses the service, the more likely they will upgrade over time. This is the central assumption that many firms make when they utilize the freemium model, hoping that free consumers will eventually convert into premium consumers. Secondly, for services where social features exist, the social usage of these features can also contribute to consumer value. If two free consumers are sharing with each other over time, the social act of sharing actually contributes to the probability that both consumers will eventually convert to premium consumers. Lastly, free consumers add value via Word-of-Mouth (WOM) referrals. More specifically, in the context where a referral program exists, free consumers bring in other consumers, and over time, other consumers may eventually convert to premium consumers. We provide a method to calculate the customer network value (CNV) of these free consumers, and the firm can account for and use this information in designing products more appropriately for their customers.

Our work has several limitations that can be explored for future research. First, we currently model consumer behavior conditional on adoption of the service. This can be allayed in the future by modeling the consumer's choice of adopting the service. In addition, the data set is from a period where this firm had minimal competition, so we do not model an outside option – allowing consumers to only choose between free and premium plans. It will be interesting to examine the competitive effects of this in the future. On the social usage front, sharing often begins when consumers form links with each other by sending share folder invites. We neither model that process nor distinguish with whom the consumers are sharing. Our specification examines only the magnitude of total inbound social usage and is agnostic to the type of individuals with whom consumers share files.

## Related Literature

Our work intersects multiple domains of literature from a substantive viewpoint: *consumer-to-consumer referrals*, *product sampling*, *product line design and customer lifetime value*. In terms of referral incentives, past works have recognized the importance of managing referral programs (Buttle, 1998; Silverman, 1997). Biyalogorsky et al. (2001) explored the design of optimal referrals, and Ryu and Feick (2007), through experiments, showed that for strong brands, it is good to reward both the sender and the receiver of the referral in order to maximize referral rates, which is true of the referral program of our context.

Another related literature is product sampling. Prior studies on digital goods focused on the fact that they are experience goods, and contended that consumers require time to learn the value of these goods and services (Jain et al., 1995; Heiman and Muller, 1996; Lehmann and Esteban-Bravo, 2006; Heiman et al., 2001; Chellappa and Shivendu, 2005). Therefore, firms can influence the propensity of a consumer to adopt a service by providing free trials. Our context, however, differs temporally. In lieu of offering a limited-time free trial, the freemium model offers a perpetually free product, which can end up serving as a close substitute. Therefore the issue of cannibalization of the premium product is of significant concern. A growing body of literature is emerging that tackles these issues in the form of theoretical models that explore the economics of freemium (Niculescu and Wu, 2013), but given that the dynamic long-term effects are of first-order importance, the paucity of empirical (and even theoretical, with a few exceptions) research is striking.

In the CLV literature (Berger and Nasr, 1998; Gupta et al., 2006a; Fader et al., 2005; Schweidel et al., 2011), firms consider consumers based on a recurring stream of revenues, leading to a “lifetime” value associated with each consumer. Firms can then determine their acquisition and retention strategy based on these value estimates. The value of free customers has also been examined by Gupta et al. (2006b), who evaluate the average of aggregate value of buyers (to sellers and the two-sided platform) in an eBay-like online market place context, where the primary driver of the value of the free consumer comes from the nature of the two-sided platform, and the marketplace obtains fees from sales of goods that buyers bid and purchase from sellers.

From a methodological perspective, our work follows the stream of dynamic discrete-choice structural models (Miller, 1984; Rust, 1987; Wolpin, 1987). To our knowledge, while there are other models of discrete and continuous choice models (Hanemann, 1984; Song and Chintagunta, 2007), we are one of the first studies in Marketing to incorporate multiple discrete and continuous actions in a dynamic structural model and to estimate it using a technique that recovers the value function. While Bajari et al. (2007) and Ryan (2012) use BBL to estimate a dynamic structural model with both discrete and continuous actions, their estimation procedure fails to recover the value function of consumers. In addition, we present an analytic solution to the continuous action using the Euler Equation and Envelope Condition in order to ease the computational burden by avoiding having to numerically maximize over all possible continuous actions per discrete-choice action. Several authors have examined constrained continuous choices, most notably Timmins (2002), who uses discretization in conjunction with grid inversion to obtain the unobservable shock corresponding to the continuous choice. The closest work to ours in terms of discrete-continuous actions and analytic solution is Michelangeli (2008). However, the author’s work differs in two aspects: 1) the model is a finite-horizon dynamic programming model, and the approach is not readily applied to our infinite-horizon context, and 2) the author uses measurement errors associated with the continuous action, and therefore also limits the ability to conduct a wide range of counterfactuals. Because a major focus of our approach lies in the ability to conduct many counterfactual simulations, our model includes the ability to incorporate discrete and continuous actions with structural errors and to recover the value function.

Next, we detail the institutional context and the relevant features of the service. In section three, we describe the details of the data set that we use, as well as model-free evidence that supports our initial conjectures of the value of free consumers. In section four and five, we describe the model and the estimation procedure. In section six, we present the estimation results and the findings of various counterfactual analyses. Lastly, we conclude with a discussion of the managerial implication and limitations of this research.

## 2 Institutional Setting

The freemium company that provided the consumer data is a leading online storage company that stores consumer files in the cloud that synchronize across multiple devices (e.g. laptop, desktop, and mobile phone). The company was founded in the 2000s and currently has hundreds of millions of users world-wide. Later in 2012, major competitors (such as Apple) entered the space by introducing similar versions of the service.[5] During the time period of our data, the cloud storage industry was fragmented amongst smaller providers. However, our focal firm quickly emerged as an important player in the consumer storage and syncing industry, while many potential competitors acted primarily as security and backup services. Therefore, for our purpose, we regard the firm as a monopoly growing a captive user base.

The value proposition of the service is for consumers to store and sync files in the cloud and to share files with other users. Consumers do this by installing an application on their desktop. This application appears as a special folder on the consumer's desktop. Consumers can then add files to their account by simply dragging files into the folder, as one would do with any normal folder. Once the files are added to this folder, a copy of the files are then transferred onto the firm's servers and can be accessed through all of the devices that a consumer has the service software installed, or via an online interface (similar to the workings of a web-mail interface). While a consumer can access their files through different means (e.g. desktop, mobile devices, or web interface), the primary method that consumers use to access the service at the time of our observation is via their desktop, and therefore we focus on this point of usage in our analysis.

Once the files are stored in the account, they take up space that counts towards an account quota. When signing up, all consumers are presented with the choice of three different plans: Free, Premium-Tier-1 and Premium-Tier-2. Free is the basic plan where the consumer receives 2GB of quota; Premium-Tier-1 and Premium-Tier-2 are the premium plans where consumers would receive 50 and 100GB of storage, respectively. These premium plans work on a subscription basis, with the options of payment plans of monthly or yearly. The pricing and payment plan scheme is listed in Table 1. A majority of the upgraded consumers choose the Free and Premium-Tier-1 plans, and therefore we focus our analysis on these two plans, referring to the Premium-Tier-1 plan as the premium plan hereafter.

| **Plan** | **Monthly** | **Yearly** | **Referral Incentive** |
|---|---|---|---|
| Free | - | - | 250MB/Referral Accepted |
| Premium-Tier-1 | \$9.99/Month | \$99/Year | 500MB/Referral Accepted |
| Premium-Tier-2 | \$19.99/Month | \$199/Year | 500MB/Referral Accepted |

Table 1: Price/Referral Incentive by Payment Plans

When a consumer runs out of space, any files that a consumer adds into the service will no longer be uploaded to the server. Most importantly, all file synchronization stops, and since this is the primary value proposition of the

[5] All of these other competing services also use the freemium model.

service, the software is rendered virtually useless to the consumer. This is costly to the consumer, because then a consumer must take the time out to decide which files are not important, and move files out of their full accounts. And even after doing so, it will take time for the account to return to its functionality, as additional time is required for the local folders to resynchronize with the cloud account, proportional to the amount of files removed. Therefore, from conversations with consumers, most users tend to leave “cushion space” in their account. We observe this behavior in a majority of the consumers in the data.

To make space available in the data, a consumer has several choices to make. She can either a) delete files from her account, b) send referral invites to other consumers to join the service, or 3) choose a plan with higher quota. To delete files from her account, a consumer moves or deletes files from their account folder. This change is then synchronized onto the firm’s servers.

A consumer can also earn additional storage space through the referral invite program. In order to use the referral program, a consumer can send out a unique link to other consumers who have yet to join; this unique link includes an identification number that links the invites back to the original sender. There is no limit on the number of referral invites that one can send out, but for the referral to count towards their quota bonus, the friends must join using the attached, unique link. In addition, the referral invite works “both ways”, in that a consumer joining through the original invite also receives an additional 250MB of space. Hence, senders have the incentive to always include a link with their word-of-mouth, and receivers have the incentive to join via the links.[6] Therefore, while there may be some cases that consumers will not be identified as “referred” consumers in our data and bias our results of the effect of WOM and the usage behavior of non-referred consumers, this problem may be at a minimum. A consumer accepts the referral invite by signing up for the service. Once this is done, the original sender receives credit for the invite acceptance and earns the additional space. The level of additional space one receives depends on the plan that one has chosen, as shown in Table 1. While this is a very effective way for consumers to gain space, a consumer can only receive credit for a maximum of 32 acceptances.[7]

Another key feature is for consumers to share files with existing consumers. For instance, if Alice wishes to share files with Bob, Alice would 1) create a sub-folder within their account, 2) send a share folder invite to Bob, and 3) place files into the shared folder. Once Bob accepts the share folder invite, whatever files exists in the share folder is then automatically synchronized across both Alice and Bob’s accounts and will also count towards the quota on both

[6] One might be concerned that this incentive system may encourage consumers who are already planning to join to actively seek out invites from other consumers. If this were the case, then these “willing” consumers may already have a favorable disposition towards the service, and are more likely to behave favorable towards the service (i.e. use the service heavily, send out more invites, more likely to upgrade to a plan later on). We acknowledge that this will bias our results upward and one way to possibly check for the existence of this behavior is to conduct surveys on existing consumer population.

[7] As with any reward system, we have to be aware of of consumers trying to gain more space by “gaming” the system, mainly by creating clone accounts using additional email addresses. The firm is aware of this, and spends significant resources exactly to correct these gaming behavior in consumers by not rewarding false referrals. For instance, they can verify the source of two very different emails from the same consumer by verifying if these clone accounts install the software on the same device using machine footprints such as MAC addresses. Because of the efforts from the company in correcting gaming behavior, we assume that the integrity of our data is not comprised by this behavior.

consumer's account. While it is possible for consumers to share files with consumers who have not adopted using a special "Public" folder, a majority of the sharing is through this aforementioned private share folder feature, and therefore we focus the social usage extension on this. While our model currently does not account for this, we discuss how to account for consumer social usage in section 4.8, the model extension for social usage.

Up to this point, we have focused primarily on the service from the consumer's perspective. It is now helpful to describe the firm's strategy decisions relating to the product: price and length of subscription plans, size of plan quotas, and size of referral incentives. Before the company's public launch, the firm ran pricing experiments on small subsets of its consumers in order to set the current pricing and timing plans. These experiments assumed that the consumer's behavior is static over time, and does not change during the observed period of data. Running counterfactuals off existing data is especially helpful in a setting like this to help inform the product design process. With the incorporation of customer heterogeneity, our methodology allows for them to rank order consumers according to the CLV, and then retroactively study the usage data for the most heavily used features of the most valuable consumers. Then, they can allocate company resources to make certain features easier to use (i.e. easier to share, easier to delete). The data-centric design philosophy is fairly popular among Silicon Valley start-ups, especially with online-gaming firms. Companies like Zynga have teams of data analysts to guide their game design decisions.

# 3 Data

The goals of this section are to show the characteristics of our data set and to describe the model free evidence that will help us identify the structural model that will be described in the next section. We obtained a sample of 1,363 anonymous consumers who joined during in the first two years of the firm's using a second-degree snowball sampling methodology. We underwent the following procedures to acquire the total sample of consumers:

1. Randomly sample a seed set of 50 people who have joined in the random sample seed window (Seed Group).

2. Add consumers who are connected (shared files or invited by) to Seed Group (1st Degree Group).

3. Add consumers who are connected to 1st Degree Group (2nd Degree Group).

The random sample seed window includes the first two years after the launch of the service. We then obtain all of these consumers' user activities from their join date until December 31, 2011. Our panel data includes the detailed click-stream data of these consumers over the four year period, which we aggregated into a weekly level. We observe a suite of consumer behaviors that are relevant to our problem.[8] These activities include:

- Total number of files stored and the amount of storage.

[8] In order to protect the confidential nature of the data, usage statistics such as addition, deletion and storage have been normalized to the maximum observed number in the series in Table 2.

- Amount of files deleted.
- Total amount of storage added.
- Number of referral invites that are sent to consumers who have not already joined the service.
- Number of referrals accepted each week.
- Plan choice and payment plan when upgrading.

While we cannot disclose the exact percentage of premium to total consumers, many freemium companies observe premium-to-total consumer ratios ranging from the single digits to over ten percent.

| **Statistic** | **Mean Across Sample** | **SD** | **MIN** | **MAX** |
|---|---|---|---|---|
| Number of Consumers | 1,349 | - | - | - |
| Total # of Observations | 155,279 | - | - | - |
| Time Periods | 115.107 | 18.527 | 93 | 206 |
| Average Addition | 0.0167 | 0.0557 | 0 | 1 |
| Average Deletion | 0.0131 | 0.0502 | 0 | 1 |
| Average Storage | 0.0121 | 0.0542 | 0 | 1 |
| Average Referral Sent | 8.503 | 48.76 | 0 | 966 |
| Average Referral Accepted | 1.831 | 7.626 | 0 | 155 |
| Average Referral Rate (Sent/Accepted) | 0.169 | 0.321 | 0 | 1 |

Table 2: Summary Statistics of Consumers

## 3.1 Model-Free Data Patterns

In this section we examine the data patterns of consumer behavior with the goal of clarifying the key data features that our model needs to characterize and inspire the major design choices of our model. Ultimately, we observe two patterns in the data that justify the value of a free consumer. First, consumers upgrade themselves over time as they become closer to reaching quota. This is the first value of the free customer. Second, even if free consumers never upgrade themselves, they may recruit an additional consumer whom may eventually upgrade.

### 3.1.1 Free Customers Upgrade Over Time From Personal Usage

First, we first examine the ratio of free to premium consumers over time. The left panel of Figure 1 shows that the growth of free consumers greatly outpaces the growth of the premium consumers. This indicates that free consumers cannot be overlooked. There are more consumers who begin as free consumers and convert to premium than there are consumers who join as premium consumers from the beginning.

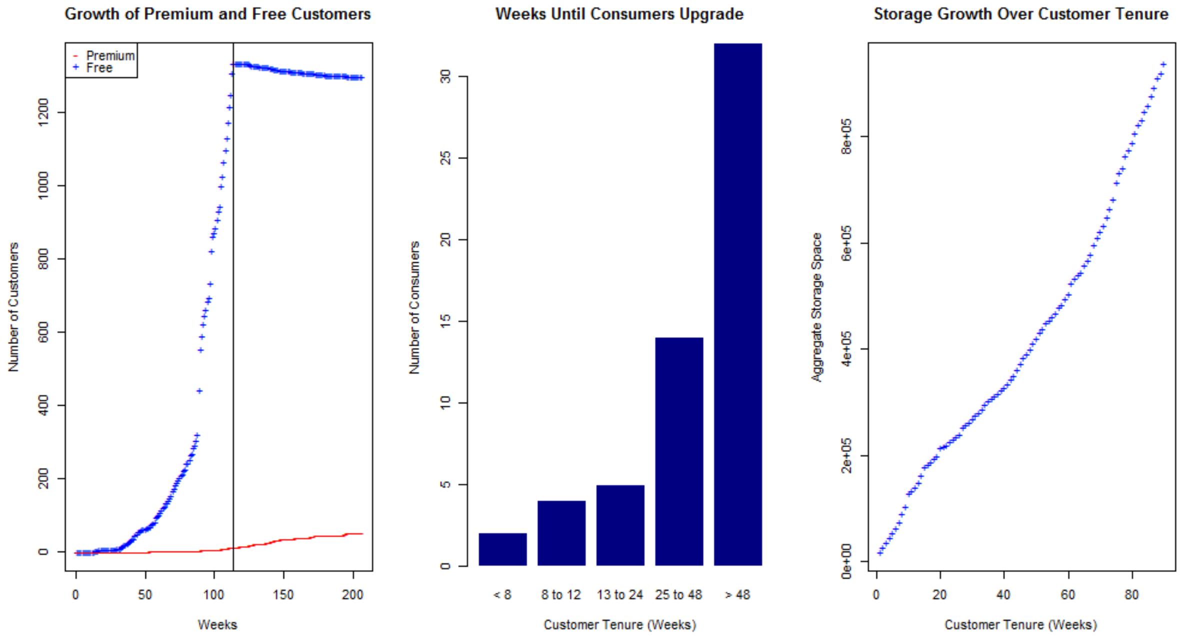


Figure 1: Upgrade Patterns

The middle panel confirms that a majority of consumers upgrade after using the service for several weeks. There are only two consumers out of the entire group of premium consumers in our sample who upgrade within the first six weeks of joining the service. The graph on the right shows the growth of the aggregate consumers storage within their first 90 weeks. This graph is from the perspective of the consumers, in that we see the average storage used per consumer grow over time. This suggests that consumers begin using the service as free consumers and later upgrade to a premium plan once they store enough files.

The customer mix, i.e. the fraction of consumers who choose the premium product is a critically important variable in the freemium business model. We detail the dynamic variation of the consumer mix in Figure 2. We observe a non-monotonic inverted-U-shaped pattern for the fraction of premium users, suggesting that the firm would have to be patient for customer acquisition to be converted to revenues. If the company incurs costs for each “free” customer, then it might well see diminishing profitability when the fraction of premium customers drops over time.

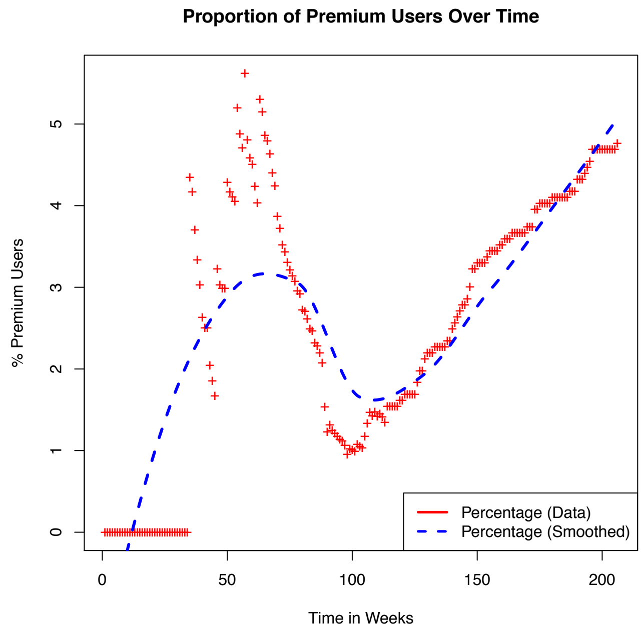


Figure 2: Dynamics of Customer Mix over Time

### 3.1.2 Free Customers Bring in Premium Customers Through Referral Invites

We now examine the consumer referral behavior and whether referral invites recruit consumers who later convert to premium consumers. The goal of referrals is to bring in new consumers, with the hope that these consumers will upgrade to premium. Informal interviews with management indicate that 1) the referral program is effective at acquiring new consumers and 2) the growth of referral parallels with the number of new consumers. The latter reflects the viral nature of referral invites, as the number of consumers has the potential to grow exponentially with a viral coefficient greater than 1.

First, the top left panel in Figure 3 shows the ratio of consumers who joined directly versus those who joined through referrals. From this graph we observe that referral consumers reach almost 40% of the number of consumers who joined directly, and the proportion is increasing over time. This confirms the company's belief that the referral program is quite effective. Furthermore, in the top right graph we juxtapose the total number of referrals sent alongside the number of referred consumers. We see that the number of referred consumers begins small, and that the slope of growth increases dramatically from weeks 60 to 114.[9] This further endorses the significant impact of the referrals. Interestingly, it takes only a few referrals to obtain a large influx of new consumers, suggesting that even small increases in referral rates can have cascading effects.

[9]The number of new consumers stops at week 114 because that is the cut-off point in any new consumers our sampling scheme.

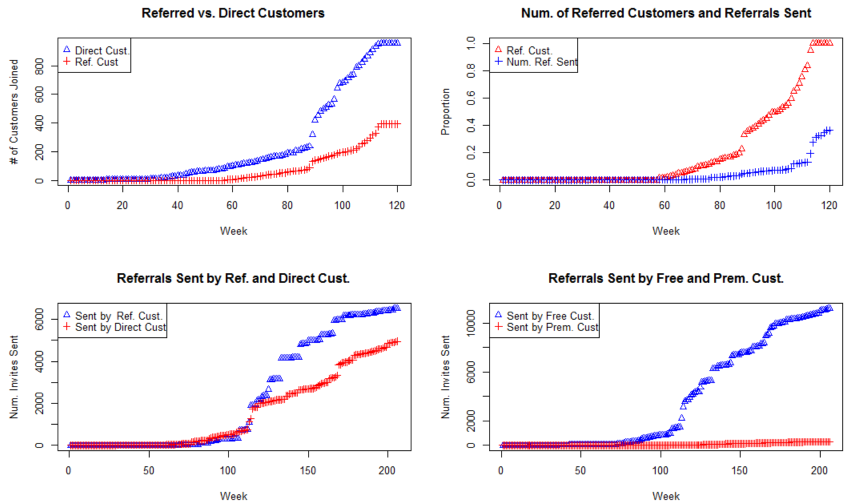


Figure 3: Referral Data Patterns

Along the theme of viral growth, in the bottom left panel we see that the average number of referrals sent by referred consumers is greater than the number of consumers who joined directly. This is added confirmation of the potency of the referral program, and evidence of the rapid snowball effect of the original referral sent by the consumer who joined directly. The bottom right panel, perhaps the most interesting of the four, shows that the total number of referral invites sent by consumers before they upgrade greatly outnumbers the total invites sent by consumers who have upgraded to a premium plan. This confirms our intuition that the success of a referral program will depend on a large number of free consumers willing to send out invites.

Lastly, in Figure 4, we plot the aggregate growth of premium consumers alongside the growth of the premium consumers who are referred by other consumers. Overall, the referred consumers account for over twenty percent of all premium consumers, indicating that there is a significant value to the referral program.

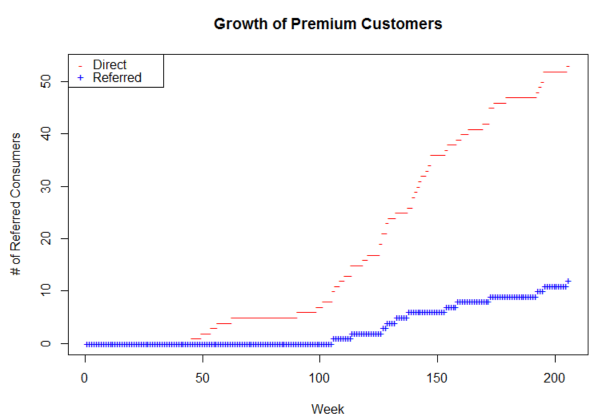


Figure 4: Breakdown of Premium Customers Who Joined Directly vs. Referred.

# 4 Model

The freemium business model is defined as a firm offering at least two differentiated variations of their service, one version with limited features but perpetually free, and the other versions with enhanced features at the cost of a subscription fee. The free and premium plans are identical in terms of quality, and the products differ only in terms of the additional premium features offered (e.g. increased storage capacity).

## 4.1 Motivation for Structural Modeling and Sources of Dynamics

A majority of users of freemium services are non-paying consumers who are initially enticed by the free plan. The most valuable asset of free consumers remains in their potential—the potential to upgrade, the potential to refer friends to join the service. Understanding the factors that influence such consumer behavior over time is key to the success of freemium. A structural model that characterizes the dynamic response of consumer behavior is therefore critical for the following reasons. First, we need to account for three different consumer choices – referral, plan-choice and deletion – in an integrated model of consumer behavior. Customers send out referral invites to share their enthusiasm for the product with friends and to earn additional free space. However, their motivation to upgrade or delete can be diminished by the extra space earned from referrals. In addition, consumer deletion behavior inherently differs according to their chosen plans. Those who have chosen the free plan may have to delete more in order to maintain enough space to store files, and those who have chosen the upgraded plan do not have to delete much due to the wealth of new space. The three decisions are endogenous, and we need a methodology that can account for this.

Second, we need a structural model because we wish to conduct counterfactual experiments to simulate the value of the free consumer and to observe the effects of changing firm policies on consumer behavior. With atheoretical models, the outcomes of changes in certain product design variables, such as price, free quota size, and referral incentives, cannot be readily characterized as there are often no variations in these variables during the observed data period. A model based on microfoundations of consumer behavior uses theory about consumer behavior to recover primitives of consumer preferences, which are likely to be invariant to changes in these product design and other policy variables. These preference parameters can then be used to evaluate how consumers would make choices in a counterfactual scenario, enabling us to provide recommendations that are under managerial interest.

A fundamental process we need to account for in our model is the inter-temporal tradeoff in upgrade, referral, and deletion behavior. The source of dynamic behavior comes from a combination of three factors: a) uncertainty in file addition, b) uncertainty in ease of upgrade, deletion and referral, and c) substantial penalty of a full account.

#### 4.1.1 Uncertainty in File Addition

First, consumers face uncertainty when anticipating the number of additional files needed for storage each period, as some weeks require less than others. In order to store files, a consumer requires space, and therefore must select a plan that fits the amount of data she will receive in the current period. When faced with the space constraint, a consumer can either upgrade to a higher-space plan, gain additional space from referring others to join, or delete files to make space.

#### 4.1.2 Uncertainty in Ease of Upgrade, Deletion and Referral Decisions

The upgrade decision is complicated by dynamic factors that facilitate or complicate the decision to upgrade to a premium plan from week to week. We see these examples of weekly unobserved factors from our discussion with current consumers. One such example is a consumer waiting for budget approval so he can pay for the premium plan. For all of the weeks prior to the budget approval, it is "harder" for the consumer to upgrade. However, once the budget is approved, even if a customer's usage is not close to quota, he upgrades. Another example is a consumer who anticipates leaving for an upcoming trip. It is easier for consumers to upgrade while in front of a computer, compared to when they are away during vacation.

If a consumer chooses not to upgrade, she has the choice to gain more space from deleting files. However, in the same manner as upgrades, consumers also faces an uncertainty in the ease of deletion. Users of the service have expressed that certain weeks are easier to delete while other weeks are harder (i.e. deadlines at work or exams at school), and we observe this lumped deletion pattern in the data. This causes the user to continually make the trade-off of whether to upgrade today in order to save the streams of deletions that she has to make in the future. Therefore at a certain point it may be optimal for a consumer to upgrade in order to outweigh the cost of continually deleting in the future.

A consumer faces two forms of uncertainty with regard to referrals. The first is the ease of referral from week to week. Second is the uncertainty of when the invitation will be accepted. Therefore, a consumer cannot simply send out invites the week that she runs out of space, and must consider in advance how likely her invites will be accepted and if her usage will be sufficiently below quota by the time the additional free space is acquired.

#### 4.1.3 Substantial Penalty of a Full Account

Lastly, a final source of dynamics is the need to anticipate periods of increased use. The consequences of a full folder in a customer's account include termination of file syncing and subsequent freezing of the account, which renders the service essentially useless. Therefore, our model should account for the fact that the consumer incurs a substantial cost due to the sudden drop in the entire value proposition of the service.

## 4.2 Consumer Decisions

The main value proposition of the service is to help consumers store and sync files in cloud storage (for now we have yet to incorporate social usage). To do this, consumers need storage space in their accounts. The size of this account depends on the different plans that consumers choose. At the time of the data set, the company offered two different plan sizes relevant to our research: 1) 2 GB for free and 2) 50 GB for \$9.99/month

In each time period $t$ (week), a consumer $i \in \{1,...,N\}$ chooses three decisions to maximize her utility: 1) whether to upgrade to a premium plan or remain a free consumer, 2) how many consumers to send referral invites, and 3) how many MB's of files to delete from her personal folder. The time line of events can be summarized in Figure 5 and is elaborated below.

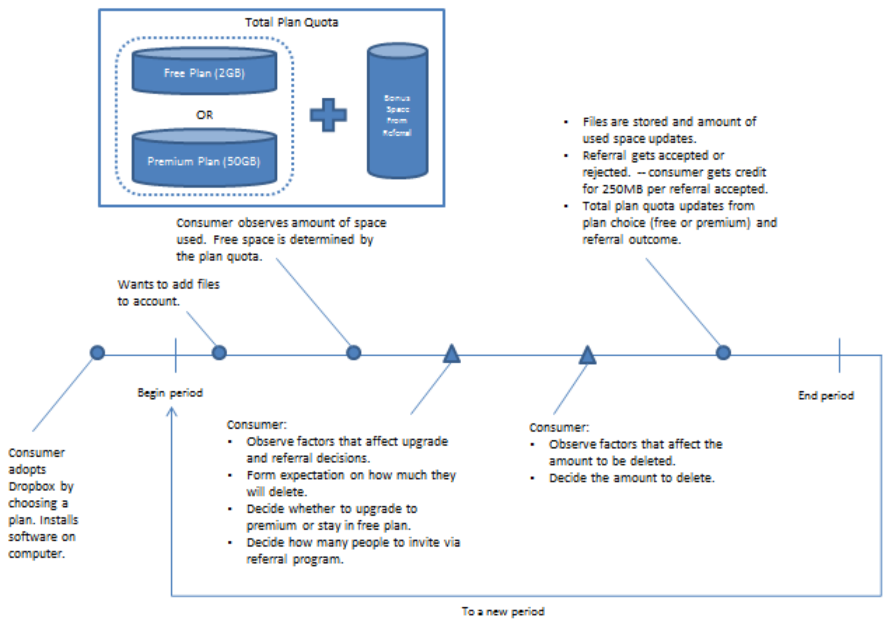


Figure 5: Time Line of Events

At each week $t$, a consumer $i$:

1. Observes exogenous shock $a_{it}$, the amount of files needed to be added.

2. Observes discrete-choice upgrade/referral shock $\varepsilon_{it}(y_{it}, r_{it})$ and chooses decisions $y_{it}$ and $r_{it}$ simultaneously.

3. Observes continuous-choice deletion shock $\nu_{it}$ and chooses decision $d_{it}$.

4. $r^a$ realizes at the end of the week based on the total number of outstanding invitations $(R_{it} - R_{it}^a)$.

5. Update state variables $x_{it}$, $z_{it}$, $R_{it}$, $R_{it}^a$.

At the beginning of the period $t$, a consumer observes $a_{it}$, the amount of files (MB) that needs to be added to her personal folder. We start the period with the exogenous addition because it is a natural departure point for the consumer process. Addition of files is the fundamental value proposition of the service and all other consumer decisions depend on the amount of files that need to be added. Then, the consumer makes the joint decision of plan choice ($y_{it}$) and the number of referral invitations to send ($r_{it}$). For the upgrade option, we specifically look at three options:

$$y_{it} = \begin{cases} 2, & \text{if customer upgrades with the yearly payment option,} \\ 1, & \text{if customer upgrades with the monthly payment option,} \\ 0, & \text{if customer does not upgrade.} \end{cases}$$

For options $y = \{1,2\}$, a consumer upgrades from a free to a premium plan and pays a price $P^m$ and $P^y$, respectively. We model both the plan price and the plan length because they are both policies that can be changed by the company and therefore are subjects of interest in the counterfactual simulations.

Simultaneously, a consumer determines the number of referral invitations to send to other consumers who have yet to sign up. This is modeled as a discrete count variable bounded by $R^{max}$:

$$r_{it} \in \{0,1,...,R^{max}\}.$$

Then, a consumer observes the continuous-choice specific deletion shock $\nu$. This is interpreted as the weekly unobserved factors that make deletion easier or harder. For instance, one week a consumer may find it harder to delete from her folders because she is traveling for work, so $\nu$ would be a low value. Another week could be spring cleaning, which makes it easier for a consumer to delete unnecessary files, rendering $\nu$ high.

After these decisions and shocks are realized, the consumer then realizes the value of $r^a$, the number of accepted referrals in the current period. $r^a$ follows a binomial distribution, with the parameters: 1) $p^a$, the empirical acceptance probability in the population, and 2) $R_{it} - R^a_{it}$ the number of outstanding referral invites, where $R_{it}$ is the cumulative number of referral invites sent, and $R^a_{it}$ is the cumulative accepted referral invites. Lastly, the four state variables $x_{it}$, $z_{it}$, $R_{it}$, and $R^a_{it}$ are updated at the end of the period. $x$ is the cumulative amount of space used in a consumer's account, and $z$ is the number of premium plan weeks left in their account. By default, $z$ is 0 for all consumers who are in the free plan.

### 4.3 Period Utility Function

Now we describe how each part of the time line component contributes to a consumer's period utility. Note that the consumer's decision making process is not solely based on this utility, but is based on inter-temporal trade-offs as described in §4.6. We suppress the $i$ subscript for expositional clarity, even though the model is at an individual level. At each time period $t$ , a consumer gains utility from having files stored in their account and from having folders that are free of files they no longer need. In addition, they incur a cost for the effort to delete files, to pay to upgrade to a premium account, and to send out referral invites to their friends. We express the components in the following equation:

$$u(\mathbf{D_t},\mathbf{S_t},\varepsilon_\mathbf{t},\nu;\Theta) = \overbrace{\theta x_t}^{\text{Storage Utility}} + \overbrace{d_t \nu_t + \alpha d_t^2}^{\text{Deletion Utility}} + \overbrace{\rho r_t^2}^{\text{Referral Utility}} + \underbrace{(\alpha^p P^m 1[y_t = 1] + \alpha^p P^y 1[y_t = 2])}_{\text{Upgrade Utility}} + \varepsilon_t(y_t, r_t) \tag{1}$$

where $\mathbf{D}_t$ is the vector of decision variables such that $\mathbf{D_t} = (y_t, r_t, d_t)$, $\mathbf{S_t}$ is the vector of observable state variables such that $\mathbf{S_t} = (x_t, z_t, R_t, R_t^a)$ . $\varepsilon_t$ is the vector of discrete-choice private shocks related to the joint upgrade and referral decision, where $\varepsilon_t = (\varepsilon_t(y = 0, r = 0), ..., \varepsilon_t(y = 2, r = r^{max}))$. $\Theta$ is the vector of structural parameters to be estimated such that $\Theta = (\theta, \alpha, \alpha^p, \rho)$. The storage utility term, a *flow utility*, contributes to a consumer's utility each period that files are stored in the account.[10] The deletion, referral and upgrade utility terms, all *action utilities*, only contribute to a consumer's utility when the actions are taken each period. Below we examine each of these components of the consumer's utility.

**Utility from Using the Service (Storage Utility)**

At each period, a consumer receives utility from using the service. $x_t$ denotes the cumulative MB used for personal folder storage. While one can assume various functional forms on this benefit, we assume a linear benefit specification on $x_t$ for parsimony, since we wish to simply capture the relationship that consumers gain more utility from having more files in their folder.

**Utility from Deletion**

The utility specification must satisfy three aspects:

1. Consumer incurs a cost of deleting files.

2. A consumer can only delete as much as there are files in the folder.

[10]The flow utility here is similar to a consumer receiving flow utility in each period after purchasing a durable good (e.g. a car or television).

3. A consumer is forced to delete the amount of files that causes the account to exceed plan quota.

We employ a flexible quadratic utility form for deletion to capture the potential convex cost to deletion. The maximum amount that a consumer can delete at any period is capped by $x_t + a_t$, the amount of files in a consumer's account at time $t$. The constraints on how much a consumer can delete is enforced implicitly in the constraint correspondence in the value function specified in the dynamics section below.

**Utility from Referring Other Customers**

The utility from referring other consumers must reflect three aspects:

1. For each accepted referral, the consumer gains the referral bonus quota $m$ MB's of space.
2. A consumer faces an uncertainty about the acceptance of each referral.
3. The consumer incurs a transactional and reputational cost for inviting another consumer.

$r_t$ is the number of referral invites that a consumer sends at time $t$. The benefit of an accepted referral is reflected in the state variable $Q_{it}$. The uncertainty that customer faces with regard to how many referrals will be accepted in each period is captured via the binomial distributed shock $r^a$, with parameters $n = R_t$ and $p = p_r$, that the customer realizes at the end of each period. Lastly, $\rho$ is the coefficient of the convex transactional cost that a customer incurs for sending out an additional invite. The convex cost reflects the fact that it becomes increasingly difficult for customers to think of an additional friend to invite in any given week.

## 4.4 Alternate Configurations of Decision Timing

While decision timing is unlikely to significantly alter the results of a dynamic infinite horizon model, we still discuss the implications of other possible orders of the consumer decisions. First we consider the placement of the deletion decision. There are two possibilities: 1) place the deletion decision before the upgrade/referral decisions or 2) model all three decisions simultaneously. First, if we were to place deletion first, we would be assuming that customers have the same deletion behavior regardless of whether they opt for status quo or they gain more space from upgrading/referring. Given the institutional context, this is inconsistent with their understanding of actual consumer deletion behavior – a key reason why customers refer or upgrade is because they want more storage space. Therefore, customers do not need to delete as much in a premium plan than a free plan.

In addition, we consider the ordering of placing the $r^a$ variable after all of the major decisions. The reason for this is that a majority of the realized acceptances, as observed in the data, do not come immediately after the invitations are sent out. Therefore, it is important to capture this uncertainty in referral acceptance that customers face when they make the upgrade, referral and deletion decisions. There are two other places we can place this variable: 1) at the beginning of the period, or 2) between the upgrade/referral decision and deletion decision. In the first case, since we have a dynamic model, we end up with the same Bellman equation as if we had placed the realization in the end. As a result, the implications of the model stay the same. If we were to place $r^a$ between the upgrade/referral and deletion decision, then we would specifically assume that the referrals per period must be realized before a customer makes the deletion decision. This is a strong assumption, since it assumes that the uncertainty in referral acceptance has no effect on a customer's deletion decision. Aside from being a stronger assumption than what is currently assumed, this also makes the solution of the value function more complex. As a result, we settled on our current specification of the time line as a fairly reasonable assumption.

## 4.5 State Evolution

The state variable $x_t$ keeps track of the total amount of files in MB's that is stored in a consumer's account. This is updated via the linear law of motion $x_{t+1} = x_t + a_t - d_t$, which is simply the sum of the amount of files observed at the beginning of the period and the observed addition amount, subtracted by the amount deleted in the particular period.

The state variable $z_t$ keeps track of the number of periods until the consumer's next payment. $z_t$ is set to 52 if the consumer chooses plan 2, and it is set to 4 if he chooses plan 1. $z_t$ decreases by 1 each period. The state evolution for $z_t$ is specified as:

$$z_t = \begin{cases} 0, & z_{t-1} = 0 \wedge y_t = 0 \\ 52, & z_{t-1} = 0 \wedge y_t = 2 \\ 4, & z_{t-1} = 0 \wedge y_t = 1 \\ z_{t-1} - 1, & z_{t-1} > 0 \end{cases}$$

Only a portion of the referral invites are actually accepted, and we keep track of the total number of successful invites as $R_t^a$. $R^{max}$ is the empirical number of maximum per-period invitations in the data. Whenever an invite is accepted, a consumer gains an additional $m$ MB's of space to their quota. More specifically, we define the variable $Q_{it}$ as a function of the total number of successful invites ($R_{it}^a$), the baseline amount of space ($Q^{free}$), referral bonus capacity ($m$), and incremental amount of space provided by the premium plan ($Q^{premium}$):

$$Q_{it} = Q^{free} + mR_{it-1}^a + \mathbf{1}[z_{it} \geq 1 \vee y_{it} = 1 \vee y_{it} = 2]Q^{premium}$$

The following Table summarizes the state variables and the corresponding laws of motion.

| State Variable and Shocks | Description | Type | Law of Motion |
|---|---|---|---|
| $x_{it}$ | Cumulative MB used in Personal Folders | Observed | $x_{it+1} = x_{it} + a_{it} - d_{it}$ |
| $R_{it}$ | Cumulative number of referrals sent | Observed | $R_{it+1} = R_{it} + r_{it}$ |
| $R^a_{it}$ | Cumulative number of accepted referrals | Observed | $R^a_{it+1} = R^a_{it} + r^a_{it}$ |
| $z_{it}$ | Number of premium weeks left | Observed | $z_{it+1} = z_{it} - 1$ |
| $Q_{it}$ | Total Quota | Observed | $Q_{it+1} = Q^{free} + mR^a_{it} + \mathbf{1}[z_{it} \geq 1]Q^{premium}$ |
| $\varepsilon_{it}(y,r)$ | Upgrade and referral decision shock | Unobserved | Type I Extreme-Value(0,1) |
| $\nu_{it}$ | Deletion decision shock | Unobserved | Log-Normal(0,1) |
| $a_{it}$ | Addition shock | Observed | Log-Normal($\mu_a$,$\sigma_a^2$) |

Next, we describe the full dynamic model.

## 4.6 Dynamics in Consumer Decisions

The dynamics in the consumer's decisions stem from the inter-temporal tradeoff of the consumer's current benefit versus the future benefits of upgrading to a premium account, deleting files to gain free space, and referring other consumers due to social and practical benefits. We therefore model this tradeoff as the sum of discounted future period utilities:

$$\max_{(y,r)} \mathbf{E}_{a,r^a,\varepsilon}\left[\sum_{t=0}^{\infty}\beta^t \mathbf{E}_\nu\left[\max_d u(\mathbf{D}_{it},\mathbf{S}_{it},\varepsilon_{it},\nu_{it};\Theta)\right]|\mathbf{S}_{it}\right] \tag{2}$$

where $\beta$ is the assumed discount factor for all consumers, and the utility function is specified in Equation 1. The solution to the above dynamic programming problem is the same as the solution to the Bellman equation, which is hereby referred to as the value function:

$$V(\mathbf{S},\varepsilon,a;\Theta) = \max_{y,r\in\Gamma(z)} \mathbf{E}_\nu\left[\max_{d\in H(x,z,a,R^a)} u(\mathbf{D},\mathbf{S},\varepsilon,\nu;\Theta) + \beta\mathbf{E}_{\mathbf{S}',\varepsilon',a'}\left[V(\mathbf{S}',\varepsilon',a';\Theta)\right]\right] \tag{3}$$

The integrated Bellman equation, $EV$, expresses the fixed-point that we solve to derive the solution of the expected value function with the discrete-choice shocks $\varepsilon$ as well as $a$ integrated out:

$$EV(\mathbf{S};\Theta) = \mathbf{E}_{a,r^a,\varepsilon}\left[\max_{y,r\in\Gamma(z)} \mathbf{E}_\nu\left[\max_{d\in H(x,z,a,R^a)} u(\mathbf{D},\mathbf{S},\varepsilon,\nu) + \beta EV(\mathbf{S}';\Theta)\right]\right] \tag{4}$$

The difference in interpretation for the expected value function is that it is the value function prior to consumers observing all shocks, and therefore is expressed only as a function of the state variables $\mathbf{S}$. $\Gamma(z)$ is the the choice set for the discrete decisions $y$ and $r$, specified as $(y,r) \in \Gamma(z)$ where:

$$\Gamma(z) = \begin{cases} \{0,1,2\} \times \{0,...,R^{max}\}, & \text{z=0} \\ \{0\} \times \{0,...R^{max}\}, & \text{otherwise} \end{cases},$$

again reflecting the fact that when a consumer is already on the premium plan ($z > 0$), she has no choice to make regarding the product and so the plan is trivially set to 0.

The choice set for the continuous deletion decision $d$ is specified as $H(x,z,a,R^a)$:

$$H(x,z,a,R^a) = [\max(0, x+a-Q(z,R^a)),\ x+a],$$

reflecting the fact that consumers cannot delete more than the total amount stored, and must delete a sufficient amount so that they do not exceed their quota, e.g. when the consumer has stored $x = 1.5$ GB and wants to add $a = 1$ GB, and has a baseline quota of $Q(0,0) = 2$ GB, then, she must delete at least 0.5 GB $(x+a-Q(0,0))$ of data in order to stay within the limit. Observe that if the consumer had chosen to upgrade earlier in the period to an annual premium plan, she could have chosen $y = 2$ resulting in $z = 52$ and a corresponding quota of $Q(52,0) = 100$ GB, allowing for a much higher degree of flexibility.

## 4.7 Identification

$\alpha^p$ is the price coefficient and is identified primarily by $y_t$, the upgrade decision. It would be highly negative if the average number of upgrade from week to week is low. Note that given the fact that the price does not change in the entire observation period, we identify this parameter from the population of consumers who eventually upgrade. In addition, the dynamics in the $x_t$'s in conjunction with the $y_t$'s also help identify this parameter. If the magnitude of $\alpha^p$ were high, then we would see more occurrences of $x_t$'s that are close to the free quota, meaning the consumers will only upgrade when they are close to quota. However, $\alpha^p$ would be low if, on average, upgrades occur when the average level of $x_t$'s are low.

$\theta$ is the marginal utility of storage. It would be high in magnitude if the average level of $x_t$ is high in conjunction with low levels of $d_t$. The parameter would be low in magnitude if the average level of $x_t$'s is low, and we see high levels of $d_t$. This parameter is identified via the dynamics in $x$'s and $d$'s. We would not be able to separately identify this parameter from $\alpha$, the cost of deletion, if the problem were static.

$\alpha$ is the cost to deletion. This parameter is identified from the variation in a consumer's weekly deletion behavior. The parameter would be highly negative if the average amount of deletion is low, meaning it is very costly for consumers to delete. On the other hand, the parameter would be low if the average $d$ were high from week to week, meaning that it is not very costly for consumers to delete.

Lastly, $\rho$ is the quadratic cost to referral. The parameter is identified from the weekly variation in $r$, the consumers' referral behavior. The parameter would be highly negative if the average amount of referrals is low, meaning that it is costly for consumers to invite other consumers. The parameter would be low in magnitude if the average amount of referrals is high, meaning that it is not very costly for consumers to send out invites.

## 4.8 Model Extension: Social Usage

We now describe the extension to include sharing (social usage) in our model. Social usage is defined as when consumers use the service to share files with other existing consumers. The incorporation of social usage in our model makes dynamics even more important, as consumers face additional uncertainty regarding how many files other consumers will add to their jointly shared folders. In order to avoid the scenario where another consumer suddenly shares a large amount of data that maximizes both parties' quota, thereby suspending the accounts, consumers must have the foresight to ensure adequate free space by early upgrading, sending out referrals, and deleting files. In order to delineate the decisions from the old model, we now redefine several variables in order to make the distinction between personal and social decisions and shocks. We place a $p$ and $s$ superscript to the variables $x_{it}$,$a_{it}$, $d_{it}$, and $\nu_{it}$, to make the distinction between personal and social versions of the storage state variable, addition shock, deletion decision, and the deletion-specific shocks. For instance, $a^p_{it}$ now denotes the amount of files added to personal folders and $a^s_{it}$, the amount added to social folders. We now outline the extensions to different parts of our models in order to account social usage.

### 4.8.1 Modification to Consumer Decisions and Time Line

In each time period $t$ (week), a consumer $i \in \{1,...,N\}$ chooses one additional decision to maximize her utility – $d^s_{it}$, the amount of files to delete from shared folders. There are three components that are added to the original time line, and now the time line is:

1. Observes exogenous shocks $a^p_{it}$, the amount of files that need to be added to a consumer's personal (any non-shared) folders, and $a^s_{it}$, the amount of files added to their shared folder by the shared neighbors.

2. Observes discrete-choice upgrade/referral shock $\varepsilon_{it}(y_{it}, r_{it})$ and chooses decisions $y_{it}$ and $r_{it}$ simultaneously.

3. Observes continuous-choice deletion shocks $\nu^p_{it}$ and $\nu^s_{it}$, and chooses decisions $d^p_{it}$, and $d^s_{it}$ simultaneously. $\nu^p_{it}$ and $\nu^s_{it}$ are factors, unobserved to the researcher, that affect a consumer's personal ($d^p_{it}$) and social deletion decisions ($d^s_{it}$).

4. $r^a$ realizes at the end of the week based on the total number of outstanding invitations ($R_{it} - R^a_{it}$).

5. Update state variables $x_{it}^{p}$, $x_{it}^{s}$, $z_{it}$, $R_{it}$, $R_{it}^{a}$, where $x_{it}^{p}$ and $x_{it}^{s}$ are now the total amount of space stored in the personal and social folders.

The modified time line can be seen in Figure 6.

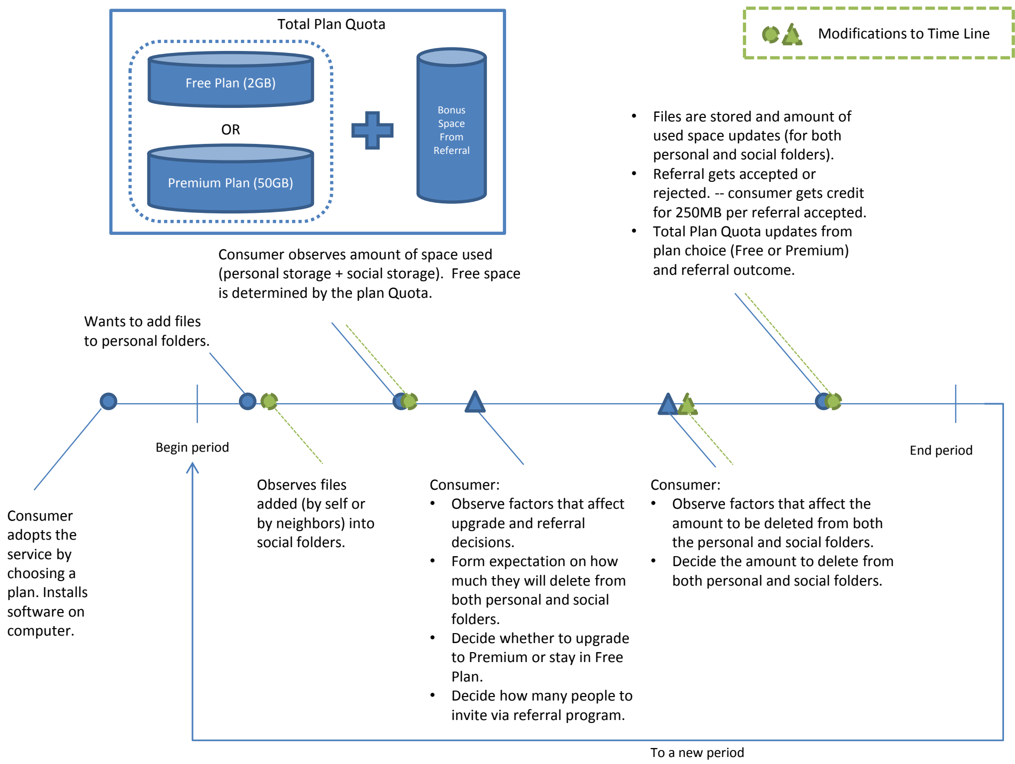


Figure 6: Modified Time Line of Events.

Next, we discuss the modification to the utility specification to incorporate social usage.

### 4.8.2 Modification to the Utility Specification

Let $\psi(x_t^p, x_t^s;\theta)$ be the utility contribution of storage, specified as

$$\psi(x_t^p, x_t^s;\theta) = \theta_1 x_t^p + \theta_2 x_t^s,$$

where $\theta = (\theta_1, \theta_2)$, is the vector of coefficients of the benefit of file storage. Following our reasoning in the model section, we assume linear, first-order effects of storage of files in personal and social folders.

The modification to the utility contribution from personal and social deletion is straight-forward, as we employ a flexible quadratic utility for both deletion actions. Let $\chi^d(d_t^p, d_t^s, \nu_t^p, \nu_t^s; \alpha)$ be the deletion utility, specified as

$$\chi^d(d_t^p, d_t^s, \nu_t^p, \nu_t^s; \alpha) = d_t^p \nu_t^p + \alpha_1 \left(d_t^p\right)^2 + d_t^s \nu_t^s + \alpha_2 \left(d_t^s\right)^2,$$

where $\alpha = (\alpha_1, \alpha_2)$, is the vector of coefficients of the personal and social deletion cost. Including the above modifications, we arrive a new version of the utility specification from Equation 1:

$$\begin{aligned} u(\mathbf{D_t}, \mathbf{S_t}, \varepsilon_\mathbf{t}, \nu_\mathbf{t}^\mathbf{p}, \nu_\mathbf{t}^\mathbf{s}; \Theta) = \psi(x_t^p, x_t^s; \theta) + \chi^d(d_t^p, d_t^s, \nu_t^p, \nu_t^s; \alpha) - \rho r_t^2 + \\ \alpha^p P^m 1[y_t = 1] + \alpha^p P^y 1[y_t = 2] + \varepsilon_t(y_t, r_t). \end{aligned} \tag{5}$$

The maximum amount that a consumer can delete at any period is capped by $\left(x_t^p + a_t^p + x_t^s + a_t^s\right)$, the total amount of files in a consumer's personal and social folders at time $t$. As before, the constraints on how much a consumer can delete is enforced implicitly in the constraint correspondence in the value function specified in the Bellman equation. The state vector $\mathbf{S_t}$ and decision vector $\mathbf{D_t}$ are now $\mathbf{S_t} = \left(x_t^p, x_t^s, z_t, R_t, R_t^a\right)$ and $\mathbf{D_t} = \left(y_t, r_t, d_t^p, d_t^s\right)$.

#### 4.8.3 Modification to the Dynamic Model

There are two changes to the dynamic model specification. The first is in the laws of motion for the state variables $x_t^p$ and $x_t^s$, and the second change is the specification of the expected value function from Equation 4.

First, since the storage state variable has now been split into two state variables $x_t^p$ and $x_t^s$, their laws of motion must be specified as well. We take the linear transition law from before and apply it to both variables, such that $x_{t+1}^p = x_t^p + a_t^p - d_t^p$ and $x_{t+1}^s = x_t^s + a_t^s - d_t^s$.[11] The consumer in this shared usage model thus has an additional decision, on how much to delete from social usage folders. She implicitly makes a tradeoff between deleting personal documents and files versus shared files, which could potentially have implications for usage by others.

Second, in the expected value function, we have to take additional expectations $a^s$ and $\nu^s$, which are the shocks corresponding to the social additions and deletion of files that are socially shared. The inner-maximization problem is over the vector of deletion decisions $d = (d^p, d^s)$.

$$EV(\mathbf{S};\Theta) = \mathbf{E}_{a^p, a^s, r^a, \varepsilon}\left[\max_{y, r \in \Gamma(x,z)} \mathbf{E}_{\nu^\mathbf{p}, \nu^\mathbf{s}}\left[\max_{d \in H(x^p, x^s, a^p, a^s, z, R^a)} u(\mathbf{D}, \mathbf{S}, \varepsilon, \nu^p, \nu^s) + \beta EV(\mathbf{S}'; \Theta)\right]\right].$$

[11] Since $x_t = x_t^p + x_t^s$, we do not need to store $x_t$ as an additional state variable.

### 4.8.4 Additional Analyses to be Conducted

With the social usage extension, a number of additional research questions can then be answered.

- **Value of the Free Consumer:** what is a more accurate assessment of the value of a free consumer?
- **Value of the Social Feature:** what is the value of the ability of sharing with consumers over time? Does this value increase or decrease over time? How does it compare with the referral program?

We hypothesized that one of the major sources in which free consumers deliver value is their act of sharing with other consumers. The more that a consumer shares with another, the higher the probability that each consumer will upgrade. Therefore, by having the social usage portion incorporated into our model, additional counterfactuals can actually be conducted to assess a more accurate value of the free consumer and how this value is deconstructed with regard to personal usage, referrals, and social usage. We also expect to investigate how this value changes over time. As consumers share with an increasing number of people over time, we might expect that the value of the social feature may increase over time as well. Incorporating the social feature into this model allows us to test this empirically.

## 4.9 Model Extension: Heterogeneity

In this section, we outline the specification of observed consumer in our utility model. As before, we index consumers by $i$, such that $i \in \{1,...,N\}$. We extend all but the price coefficient parameter to be individual-specific according to the following specification:

$$\begin{aligned} u(\mathbf{D_{it}}, \mathbf{S_{it}}, \varepsilon_{\mathbf{it}}, \nu_{\mathbf{it}}; \Theta) = \theta_i x_{it} + \alpha_i d_{it}^2 + d_{it} \nu_{it} + \\ \alpha_i^p P^m 1[y_{it} = 1] + \alpha_i^p P^y 1[y_{it} = 2] + \varepsilon_{it}(y_{it}, r_{it}) + \rho_i r_{it}^2, \end{aligned} \tag{6}$$

where we define the vector of parameters $\Theta_i = \{\Xi_i, \alpha^p\}$, and $\Xi_i$ is the vector of individual-level parameters such that $\Xi_i = \{\theta_i, \alpha_i, \alpha_i^p, \rho_i\}$. We further define a distribution over $\Xi_i$ such that:

$$\Xi_i = \begin{pmatrix} \theta_i \\ \alpha_i \\ \alpha_i^p \\ \rho_i \end{pmatrix} \sim MVN\left( \begin{pmatrix} \mu_\theta \\ \mu_\alpha \\ \mu_{\alpha^p} \\ \mu_\rho \end{pmatrix}, \Sigma \right),$$

where $\mu_\theta$,$\mu_\alpha$, $\mu_{\alpha^p}$, and $\mu_\rho$ are parameters that represent the population-level mean for $\theta_i$, $\alpha_i$, $\alpha_i^p$, and $\rho_i$, and $\Sigma$ is the population-level covariance matrix. We assume independent, diffuse normal priors on the $\mu$'s ($P(\theta_i, \alpha_i, \alpha_i^p, \rho_i) =$

$P(\theta_i)P(\alpha_i)P(\alpha_i^p)P(\rho_i)$ and a diffuse Inverse-Wishart prior on $\Sigma$.

$\theta_i$, $\alpha_i$, $\alpha_i^p$, and $\rho_i$ are identified from the individual-level variations across time in storage ($x_{it}$), distance to quota ($Q_{it} - a_{it} - x_{it}$), deletion activity ($d_{it}$), and referral activity ($r_{it}$) in our panel data. Our panel data has a considerable length in consumer activity, as we observe individual-level consumer behavior over a four-year period. Even aggregated at the weekly level, all consumers have at least 93 weeks of data.

With regard to $a_{it}$, the difference now is that the consumer simply forms an expectation over the empirical distribution of her own $\{a_{ik}\}_1^t$ draws, up to the most current week, as opposed to forming the expectation over the empirical distribution over the entire population. In addition, we estimate the $p_{ir}$ parameter, which is the individual-level referral acceptance probability. This is identified from the variation in the history of referrals sent and referrals accepted for an individual consumer.

As in the homogeneous case, consumers form expectations over the private draws of $\varepsilon_{it}$, which are i.i.d. type-1 extreme value, and $\nu_{it}$, log-normal distributed, which are both common across the population. Standard assumptions from the single-agent dynamic structural literature, such as conditional independence of shocks, still hold.

# 5 Estimation

The structural parameters $\Theta = (\theta, \alpha, \alpha^p, \rho)$ represent the benefit to storage, deletion cost, price coefficient, and the referral cost. Our model and setting present several challenges in estimation: a) large state space, b) discrete-continuous decisions, and c) jagged likelihood. We considered several possible estimation approaches; we explain why we decided to use the Bayesian Imai-Jain-Ching (Imai et al., 2009) method (IJC), and we discuss how IJC alleviates the aforementioned challenges.

We begin with a choice from two common classes of estimation approaches: iteration-based methods in the tradition of Rust (1987) or simulation based two-step methods that follow the tradition of Hotz and Miller (1993); Hotz et al. (1994). The advantage of the first method is that we obtain an estimate of the value function at the end of the estimation process, but this comes at a higher computational cost than the simulation-based methods. While the simulation-based methods are computationally light, such as BBL (Bajari et al., 2007) and POB (Pakes et al., 2007), their accuracy heavily depends on being able to correctly recover the primitives of the agent's policy function in the first step, as any errors in the first step will propagate into the second step and potentially become magnified through the simulation process.

The fundamental idea of the iteration-based estimators is to nest a fixed-point iteration step within the maximization step of MLE. First, one solves the value function of the consumer dynamic programming problem via a fixed-point iteration of the Bellman equation for a given parameter guess. The solution to the fixed-point is a contraction-mapping and therefore, under regularity conditions, we are guaranteed to find a unique solution to the value function. In the second step of the procedure, conditional on solving the value function, the problem is a traditional maximum-likelihood estimation problem, and one can proceed using tradition optimization routines to obtain a consistent estimate of the structural parameters. The algorithm iterates through these two steps for every guess of the parameter value. This procedure, referred to as the Nested Fixed Point estimator (NFXP), is the work horse in estimating many dynamic discrete-choice structural models. Rust (1987) provides proofs of convergence and the properties of the estimator.

Our setting presents a few immediate estimation challenges. First, the NFXP estimator is computationally demanding because it fully solves the Bellman equation at every guess of the parameter value. In addition, the fixed-point iteration must be solved across all states, and therefore the computational time for each iteration of the NFXP increases as the size of the state space grows. The IJC algorithm alleviates this computational challenge. It does so by 1) evaluating the fixed-point iteration once per guess of the parameters and stores a collection of these values and 2) approximating the value function by using a history of past stored value functions weighted by kernels. Another state-of-the-art estimation algorithm one can use is the Mathematical Program with Equilibrium Constraints (MPEC) algorithm (Su and Judd, 2012). While MPEC also has a lighter computational burden of estimating dynamic discrete choice model over the traditional NFXP, it requires user-supplied specifications of the constraint sparsity patterns in order to fully take advantage of the speed gains. Since we wanted an approach that offers the computational advantages as well as the benefits of Bayesian estimation, we opted for IJC. One of these benefits is the ability to specify a rich specification of heterogeneity, which we describe in section 4.9. This is another advantage over BBL, where we would have been restricted to a limited heterogeneity specification.

Furthermore, at each evaluation of the fixed-point iteration, we reduce the number of states, that we have to evaluate the value function over, by using shape-preserving splines. The spline approximates the value function by only having to evaluate the value function over a small subset of states (knots), and then "fills-in" the rest of the value function over all states. Lastly, we gain an additional computational saving by avoiding the calculation of the numerical integral over the entire support of $a$ and $\nu$. To do this, we use two independent Gaussian quadratures to approximate the integrals over a subset of the support over $a$ and $\nu$. This idea is similar to that of using splines, where the Gaussian quadrature makes a polynomial approximation of the function over a small number of nodes, a subset of the entire support of $a$ and $\nu$.

Another challenge to tackle is the discrete-continuous choice aspect of the model. IJC, like NFXP, is designed to estimate dynamic models with discrete-choice controls. In order to handle the continuous-choice control in our problem, we combine the IJC algorithm with a likelihood modification derived from the Euler equation in the spirit of the continuous-choice dynamic models described in the macroeconomics literature. We obtain log-likelihood values of the continuous-choice shock $\nu$ in a grid-inversion fashion as Timmins (2002). We discuss the details of the grid-inversion in the likelihood specification section below.

Lastly, a consequence of using the grid inversion technique is that the likelihood can be jagged and multi-modal. Therefore, not only does the likelihood not have a analytic derivative, but the jagged likelihood can cause traditional gradient-based optimization methods to become fixed at local maxima. Therefore, in practice, the model from Timmins (2002) is best estimated using comparison methods with multiple starting values. Once the optimizer is in the locality of a globally optimal region, only then can one trust gradient-based methods to find the global optimum. This can be computationally demanding and can take a bit of a coordination effort in order to ensure one finds the global optimum. In addition, if the number of parameters of one's discrete-continuous model is high, then estimating this model using traditional gradient methods would be practically infeasible. The IJC method can handle this challenge since Markov Chain Monte Carlo's (MCMC) stochastic optimization nature is robust to complex likelihood shapes that are highly non-monotonic and can handle parameter space with high dimensionality Imai et al. (2009). We have verified this in our own context with extensive Monte Carlo simulations, and find that global optimum is achieved regardless how jagged the likelihood is.

To tackle all of the estimation challenges, we use a modified version of the Bayesian Imai-Jain-Ching (IJC) algorithm, in conjunction with several state-of-the-art numerical computation techniques, such as Gaussian quadratures and splines, to estimate a discrete-continuous choices dynamic structural model in a Bayesian fashion. The IJC algorithm is a variant of MCMC. It builds upon MCMC methods, based on full likelihood estimation, in that it uses Gaussian kernels and a stored histories of stored pseudo-value functions to approximate the true value function. It provides the benefits of MCMC while alleviating the heavy computational burden of a estimating full-solution Bayesian dynamic discrete choice model with forward looking agents that requires solving the Bellman equation at each MCMC iteration.

The IJC algorithm is a modified version of MCMC, and it follows these four steps at every MCMC iteration $k$:

1. Draw proposed parameter values, $\Theta^{*k}$.

2. Evaluates pseudo-Expected Value Functions (pseudo-EVF) at currently proposed parameters and the last accepted parameters, $E\tilde{W}(D,\cdot,\Theta^{*k})$,$E\tilde{W}(D,\cdot,\Theta^{*k-1})$. These pseudo-EVF's are approximations to the Expected Value Function in Equation 4, and they are constructed using previously stored pseudo-Value Functions, from $H^k = \{\Theta^{*l}, \tilde{W}(\cdot,\cdot,\Theta^{*l})\}_{l=1}^{l=k-1}$, via kernel methods. This is the key innovation of IJC.

3. Calculates pseudo-likelihood values at the currently proposed parameters and the last accepted parameters, $\tilde{L}(\Theta^{*k}, E\tilde{W}(\cdot,\cdot,\Theta^{*k}))$ and $\tilde{L}(\Theta^{*k-1}, E\tilde{W}(\cdot,\cdot,\Theta^{*k-1}))$, using the pseudo-EVF's calculated in the previous step. These likelihood values are used in a traditional Metropolis-Hastings step, to decide whether to accept or reject $\Theta^{*k}$. Since the prior and the pseudo-likelihood are not conjugate, we cannot obtain a closed-form distribution on the posterior, and therefore we cannot use a Gibbs sampler.

4. Create a new pseudo-Value Function $\tilde{W}(\cdot,\cdot,\Theta^{*k})$ by evaluating the Bellman operator on Equation 3. This is then added to the history of past proposal parameters and pseudo-Value Functions, $H^k$. In the specific context of a dynamic discrete-choice problem from Ching et al. (2012), the pseudo-Value Function is referred to as the pseudo-Emax function.

### 5.0.1 Likelihood Specification

Now we explain the formation of the likelihood specification. With the standard conditional independence assumption, the individual likelihood for the model can be specified as:

$$L_i(\Theta) = \prod_{t=1}^{T} P(y_t, r_t, d_t | x_t, z_t, R_t, R_t^a, a_t; \Theta) = \prod_{t=1}^{T} P(y_t, r_t | x_t, z_t, R_t, R_t^a, a_t; \Theta) P(d_t | y_t, r_t, x_t, z_t, R_t, R_t^a, a_t; \Theta),$$

where $P(y_t, r_t | x_t, z_t, R_t, R_t^a, a_t; \Theta)$ is the likelihood contribution from the consumer's discrete choices: plan-choice ($y$) and number of referral invites ($r$). We are able to factor the joint likelihood $P(y_t, r_t, d_t | x_t, z_t, R_t, R_t^a, a_t; \Theta)$ into the products of $P(y_t, r_t | x_t, z_t, R_t, R_t^a, a_t; \Theta)$
and $P(d_t | y_t, r_t, x_t, z_t, R_t, R_t^a, a_t; \Theta)$ due to the timing assumption: consumers make the $y$ and $r$ decisions simultaneously, before the decision $d$. Assuming $\varepsilon(y,r)$ to be distributed type I extreme value, the functional form of the likelihood can be expressed as:

$$P(y_t, r_t | x_t, z_t, R_t, R_t^a, a_t; \Theta) = \left[ \frac{exp(V_{jk}(x_t, z_{t,}, R_t, R_t^a, a_t; \Theta))}{\sum_m \sum_n exp(V_{mn}(x_t, z_t, R_t, R_t^a, a_t; \Theta)} \right]^{1[y_t = j, r_t = k]},$$

where the discrete-choice specific Value Function $V_{jk}(x, z, R, R^a; a; \Theta)$ is:

$$V_{jk}(x, z, R, R^a; a; \Theta) = \mathbf{E}_\nu \left[ \max_{d \in H(x,z,a,R^a)} u(x, z, R, R^a; \nu; y = j, r = k, d; \Theta) + \beta EV(x', z', R' R^{a'}; \Theta) \right].$$

Next, the $P(d_t | y_t, r_t, x_t, z_t, R_t, R_t^a, a_t; \Theta)$ is the likelihood contribution from the continuous choice: deletion ($d$). Since we have a monotonic relationship between $d$ and $\nu$, it is possible to invert values of $\nu$ from observed values of $d$ through $g(\cdot)$, the first order condition from the deletion sub-problem. The likelihood can then be formed as:

$$P(d_t | y_t, r_t, x_t, z_t, R_t, R_t^a, a_t; \Theta) = P(\nu_t = g^{-1}(y_t, r_t, x_t, z_t, R_t, R_t^a, a_t, d_t) | y_t, r_t, x_t, z_t, R_t, R_t^a, a_t; \Theta) \left| \frac{\partial g^{-1}(\cdot)}{\partial d} \right|.$$

From Timmins (2002), the continuous-choice shock $\nu$ can be inverted from the policy function $d^* = g(y, r, x, z, R, R^a, a, \nu)$ for given values of all the other actions and state variables. Once the $\nu$ is recovered, it can then be evaluated at the density function of the specified distribution of $P(\nu|\cdot)$ in order to get the likelihood contribution.

For our particular specification of the utility function and the transitions, the model is part of the class of dynamic programming problems called Linear-Quadratic problems. For this class of models, one can derive an analytic solution to the continuous choice portion of the value function by using a combination of the first order conditions and the Envelope Theorem. The resulting Euler equation gives the analytic solution for $d^*$, the optimal amount of deletion, and it can be specified as follows:

$$d^*(y,r) = \max\left(\min\left(\frac{\beta\theta - \nu(1-\beta)}{2(1-\beta)\alpha},\ x+a\right),\ x+a-Q(z,r)\right).$$

The min and max statements simply provide the boundary constraints, derived from the correspondence constraint, on the optimal $d^*$ amount. The min on the $x+a$ simply ensures that a consumer cannot delete more than the current amount in the consumer's account, and the max on $x+a-Q(z,r)$ ensures that consumers are forced to delete excess files that puts the account usage over current quota. The derivation of this expression, the $\nu$ inversion expression $g^{-1}(\cdot)$, and the Jacobian term $\left|\frac{\partial g^{-1}(\cdot)}{\partial d}\right|$ can be found in the appendix.

## 5.1 Estimation Challenge with the Social Usage Extension

The biggest challenge with the social usage modification is the increased computational burden in the estimation process. The increase in the state space and the two additional integrals increase the number of functional evaluations that we need to compute at each iteration of the IJC algorithm. One possible additional innovation to lighten the computational complexity of the estimation is to treat the state variables $x^p$ and $x^s$ as continuous state variables. Fortunately, IJC allows for the natural extension of incorporating continuous state variables. The conversion into continuous state variables means that we only have to compute the expected value function at one value of $x^p$ and $x^s$ at each IJC iteration. Since we draw values of $x^p$ and $x^s$ from uniform distributions spanning the full support of these state variables, we will have values of the expected value function over a full range of values of $x^p$ and $x^s$ over many IJC iterations. Note that this approach implicitly assumes smoothness of the value function in the continuous state, and given our institutional setting, this seems like a reasonable assumption. As long we can store a sufficient number of these values, IJC can approximate the expected value function by interpolating over a weighted sum of the past history of expected value function evaluations using a kernel method. Currently, we are in the process of estimating and performing robustness checks on this extension.

# 6 Results

## 6.1 Parameter Estimates

In this section, we present the results of our estimation. We obtained these values through 10,000 iterations of the IJC algorithm with three independent chains using random initial values. Convergence of all chains are assessed visually and using the Gelman-Rubin statistic (Gelman and Rubin, 1992). We present the parameters in Table 3, and they are:

- $\theta$: consumer's benefit to storage
- $\alpha$: consumer's deletion cost
- $\alpha^p$: consumer's price coefficient
- $\rho$: consumer's referral cost

| Variable | Estimate |
|---|---|
| $\theta$: Benefit to Storage | 0.376 |
| | (0.375, 0.377) |
| $\alpha$: Deletion Cost | -0.055 |
| | (-0.0577, -0.0539) |
| $\alpha^p$: Price Coefficient | -1.917 |
| | (-2.500, -1.421) |
| $\rho$: Referral Cost | -11.339 |
| | (-11.802, -10.942) |

Table 3: Bayesian IJC Estimates (95% HPD in parentheses)

All of the signs of the parameters are as expected. In addition, the 95% highest posterior density intervals for all of the variables do not include zero, indicating that all the variables significantly contribute to consumer's utility in terms of upgrade, deletion and referral decisions. We now explain the intuitive implications of each parameter, with the first parameter contributing as a flow utility and the last three contributing as action utilities.

First we examine the parameter $\theta$, which is the benefit to storage. This parameter is the linear benefit to a consumer having files stored in their account folder. The positive coefficient indicates that consumers, on average, receive positive flow utility for having megabytes of files stored in their folders overtime. This positive coefficient indicates that consumers get value from having files stored over time as opposed to simply adding files into the folder *temporarily* and then using the service to purely transfer files between different computers and mobile devices.

$\alpha$ denotes the cost of deletion. The negative coefficient indicates that consumers have convex cost to deletion. This means it becomes incrementally more costly for consumers to delete files as the amount of files needed to be deleted increases. In other words, consumers prefer many weeks where they delete a modest amount as opposed to a few weeks of a large amount of deletion. This type of "smoothing" behavior indicates that the firm may wish to think about ways to profit from a different storage accounting scheme where, in lieu of establishing a quota for the total amount of storage per month, the firm can adjust an upload/download bandwidth scheme. Other existing freemium companies such as Evernote use such an approach.

Next, $\alpha^p$ is the price coefficient. This parameter denotes how price sensitive the average consumers may be. The negative value of this estimate is as expected and indicates the magnitude of the costs that consumers must bear when upgrading to a premium plan. Lastly, $\rho$ is the cost of referral for consumers. This cost could be attributed to the cognitive, social, and other costs of actually sending out invitations to friends. The quadratic nature of this term reflects the fact that, as each consumer sends more and more referrals per week, it becomes harder to think of more friends who do not already have invitations. We see that the highest posterior density intervals for this coefficient do not overlap zero, indicating that there are significant efforts that consumers have to bear in order to send out referral invites. The benefit for each referral is accounted by an expectation of referral bonus quota included in the correspondence constraint ($H(\cdot)$) in the dynamic problem of Equation 4. The quota increases according to a specified referral incentive amount (250 MB).

## 6.2 Counterfactuals

In this section, we present the results of the counterfactual simulations generated from the estimated parameters. The goal of these "what-if" analyses is to see the effects of changing key design variables on profits and consumer usage behavior. These design parameters include changing the price charged for the premium plans, and the magnitude and timing of referral incentives. We conduct the simulations with two objectives in mind: profit maximization and referral maximization.

### 6.2.1 Impact of Referrals on Value of a Customer

In this section, we run a counterfactual to gain a better understanding of the value of a free consumer. More specifically, we find the value of the free consumer from their referral effects. In order to calculate this, we conduct a counterfactual where the referral program does not exist.[12] We generate 1,000 identical free consumers, each starting out by choosing the free plan, and then we simulate their behavior for 115 weeks (average number of weeks observed in the data). We then compare the number of consumers who upgrade in this setting, and compare it with the baseline setting where the referral program does exist.[13] The difference in the fraction of consumers who upgrade is characterized as the average

[12] We operationalize this by 1) making $\rho$ take on a large negative magnitude, 2) making the referral bonus quota per referral to be 0, and 3) setting the $R^{max} = 0$

[13] We simulate the baseline setting by using the estimated primitives and setting all incentives to mirror the conditions in the actual data. We do this instead of subtracting the counterfactual from the actual data in order to minimize errors from simulation and model fit. In an additional effort to minimize simulation error, we use the the same sequence of draws for $a$, $\nu$, $\varepsilon$ for both the baseline counterfactual and the No-Referral counterfactual. To calculate the baseline counterfactual, we do the following three steps:

1. Start 1,000 baseline consumers on the free plan, then simulate for 115 periods for each consumer.
2. For each period, we sum up the total number of referrals accepted.

value of a free consumer based on referrals, and accrues even if these consumers never upgrade. We can then interpret this value as being a lower bound since some of these consumers will upgrade over time.

| | **Baseline Scenario** | **No-Referral Scenario** |
|---|---|---|
| Total Number of Premium Consumers | 182 | 64 |
| Organic Premium Consumers | 132 | 64 |
| Referred Premium | 50 | 0 |
| Total Referrals Sent | 914 | 0 |
| Referrals Gained from Referred | 250 | 0 |
| Total Deletion Amount in Free Plan (Organic) | 1,975 GB | 2,568 GB |
| Total Deletion Amount (Both) | 4,358 GB | 3,094 GB |
| Total Storage Amount in Free Plan Per Consumer (Organic) | 13.45 GB | 12.91GB |

Table 4: Comparison of Referral and No-Referral Program Counterfactuals

We now examine the results of this counterfactual in Table 4. There are a few interesting observations. First, the amount of consumers who upgraded from the referred batch alone is almost as many as the number of premium consumers in the No-Referral scenario. This is interesting because it indicates that the value of the consumer from referrals alone is substantial, as the referral program brings in consumers that accounts for 27% of the total number of premium consumers.

Second, we observe that in order to compensate for the amount of space that would have been gained from the referrals, consumers in the No-Referral scenario delete more, and therefore deletion is a closer substitute to referral than upgrades are. This is as expected, since consumers need to make space, and without the means to gain space from referrals their only other option is to delete more. We see this by comparing the total amount of deletion in the original 1,000 consumers who joined by themselves (Organic) during the periods when they are in the free plan. We see an increase of more than 25% in deletion in the No-Referral scenario. The implication of this is that when consumers have a lower amount of storage per period, each consumer will also have a lower probability of choosing the premium plan per period. By summarizing the history of total storage for all consumers in the free plan (excluding space gained from referrals) and comparing this amount across both the Baseline scenario and the No-Referral scenario, we see that this is indeed the case. What is interesting about this is that the Organic consumers end up storing 3% more files with the existence of a referral program. This is true even after we subtract the additional amount of files they store in the extra storage space they gain from the referral program.

Last and perhaps most importantly, there are twice as many consumers from the Organic consumers who upgrade in the baseline condition, meaning that the extra space they get from the referrals actually makes them more likely to upgrade in the long run, even if they do not end up using the extra space from referrals. What is counterintuitive about this is that typically referral programs in freemium firms are seen as an impediment to consumer upgrading to a premium version. It is viewed as a pure customer acquisition tactic via Word-of-Mouth. However, we see evidence

3. For each incremental referral accepted each week, we simulate an additional consumer for the remainder of the periods.

that the presence of referral programs, if geared to drive engagement and usage, may work to increase the number of consumers choosing a premium plan. We explore the implications of changing referral incentives in the optimal referral counterfactuals found in section 6.2.3.

Synthesizing these observations, we estimate the value of a free consumer *per month* to be an average of $2 (182*$9.99/1000 Consumers). This value comes from generating 1,000 free consumers and simulating their various consumer behavior over the average number of observed periods in our data. 35% of this value comes from the free consumers using the service in isolation and who would eventually upgrade in absence of any other effects. 27% of this value is from the free consumers bringing in other consumers who eventually choose a premium plan. Another 38% of this value comes from a synergistic effect of consumers using the service for more storage than they would have without the additional space they gain from the referrals. This translates to the fact that a total of 65% of the free consumer value comes from the existence of a referral plan.

#### 6.2.2

#### 6.2.3 Referral Maximization

Most firms are interested in the freemium business model since, when paired with referral incentives, it has the potential to help the firm grow its consumer base rapidly. Since the most important stated goal for any early stage company is to gain traction in obtaining a large user base, it is of interest for firms to understand how to make this process more effective. First, we explore what would happen if we were to change the referral incentives offered for each accepted referral invite per consumer. The default amount that is given to each referral during our observation is 250 MB per invite accepted. We vary the incentive across a wide range, from as little as 50 MB per referral to as large as 3 GB per referral. The results of the counterfactual are presented in Figure 7.

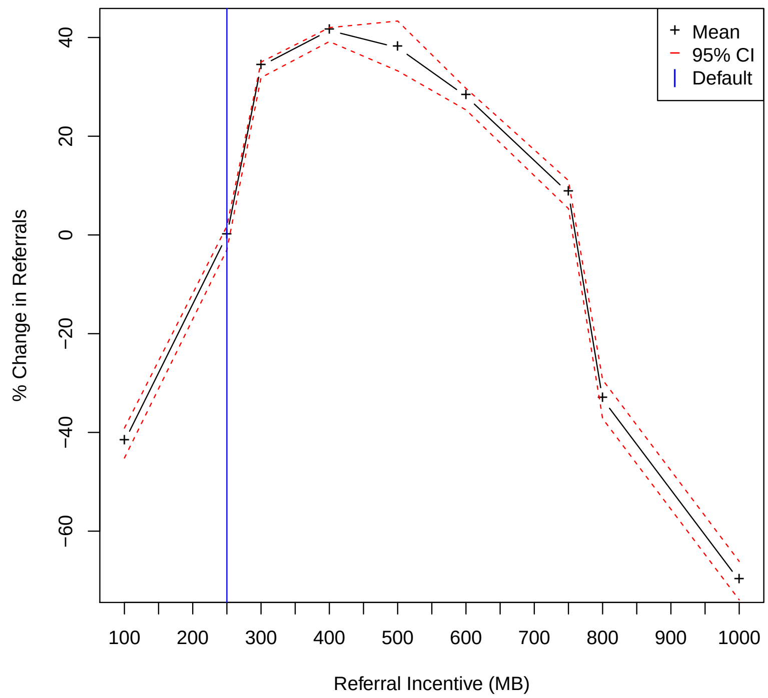


Figure 7: Change in Referral Incentives on Percent Change in Average Referrals Sent (MB/Referral Accepted)

First, note that the indicated confidence intervals for each of the changes in referral incentives are different from each other. This indicates that there are significant differences in the percent change in average referrals sent given change in referral incentives. Furthermore, as we increase referral incentives, we see an increase in the number of referrals sent, and after 450-500MB of incentive, the average number of referrals sent actually decreases. This indicates that the maximal amount of referral incentive lies around 500 MB if our goal is to increase the average number of referrals sent. If the firm gives too much space for the referral incentive, they are not encouraging more, but rather fewer referrals. This is because if a consumer can gain enough free space with only one referral, why go through all the trouble of inviting more? In addition, the increase in referrals also comes at the cost of decreasing upgrades. As consumers get more free space from referrals, they are less likely to upgrade, as indicated by Figure 8.

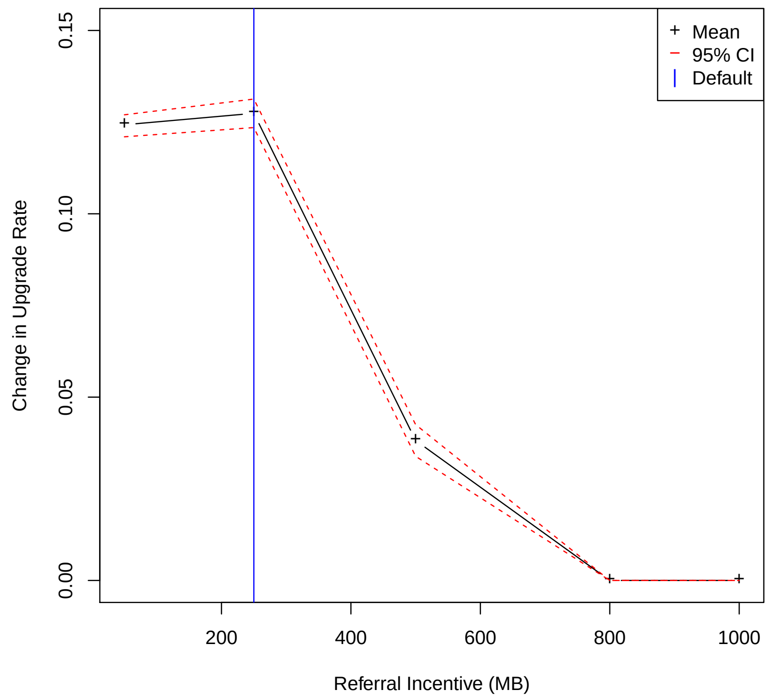


Figure 8: Change in Referral Incentives on Percent Change in Upgrade Rate

#### 6.2.4 Dynamic Trajectories of Referral Incentive

In the previous section, we assumed that the trajectory of referral incentives is static. This means that once the firm has decided a particular referral incentive, it is given to the consumer forever. For the following simulations, we explore the case where the firm offers incentives in different orders and gives higher incentives during a promotional period. We design the counterfactual along two dimensions: 1) announcement of bonus, and 2) sequential order of bonus. To conduct the simulation, we generate 1,000 consumers for 100 weeks, all observing the same exogenous shocks in order to reduce simulation error. We simulate the scenarios where the firms start out with the default referral incentive of 250 MB per referral accepted, and then ramp up to 500 MB per referral accepted at week 50 and the reverse (ramp down). We also simulate whether the firm announces this referral change at the beginning of week 1 or not. Table 5 summarizes the number of consumers who choose a premium plan in the 100 weeks, and the total amount of deletion and referrals sent across all people.

| | No. of Premium Consumers | Amount of Deletion | Amount of Referrals |
|---|---|---|---|
| Announce - Ramp Up | 109 | 2.97 GB | 982 |
| Announce - Ramp Down | 182 | 3.06 GB | 606 |
| No Announcement - Ramp Up | 132 | 3.08GB | 708 |
| No Announcement - Ramp Down | 129 | 3.09GB | 670 |

Table 5: Dynamic Trajectory Results

An examination of Table 5 suggests that the maximum in terms of referral maximization and profit maximization both are with the Announce strategies, indicating that it is potentially better to announce your incentive change ahead of time depending on your goal. If the goal of the firm is to grow a user base and encourage consumers to send out referrals, then the best strategy is the Announce-Ramp-Up strategy. We can see that there is almost a 50% increase over the Announce-Ramp-Down strategy. However, this is at the cost of drastic decrease in upgraded consumers. This makes sense because on average many more consumers are gaining additional space from 1) increase space per referral accepted, and 2) increased overall number of referrals sent. In terms of profit optimization, we see that the strategy that yields the highest number of premium consumer is the Announce-Ramp-Down. This can almost be seen as a "introductory offer" referral bonus, where in the first year of a consumer's tenure, they will have an additional incentive to send out referrals, and after the first year, they will resume a "default" incentive. What is interesting with this strategy is that one would have expected this to yield a higher total lifetime referral number, but instead, it is yielding more consumers to upgrade simply because the initial additional bonuses are actually motivating consumers to use the service more often (via increased storage). The results from the dynamically changing referral incentives show that there are more options to experiment with *how* to give the referral incentive change. Optimizing profit and referrals along the dimensions of 1) when to announce the incentive change and 2) when to actually give the incentive are questions worthy of future investigation.

# 7 Discussion and Limitations

In this study, we present a dynamic structural model of consumer upgrade, usage, and referral behaviors as a framework to develop a deeper understanding of the freemium business model. We hypothesize that the value of consumers come from three sources: 1) consumers eventually convert to premium over time, 2) consumers bring in others via the referral program who then convert to premium consumers, and 3) consumers convert to premium product because of increasing usage of social features. We find that the value of a free consumer to be $2 per month. In addition, we find that over 65% of the $2 value is attributed to the existence of the referral program alone – indicating that the effect of a referral program is quite significant. We look further into the possibility of profit maximization and referral maximization with additional counterfactual simulations. We discover two important findings: 1) existence of a static optimal incentive for referral is double the current amount, and 2) the best way to optimize referrals is to announce and ramp-up a referral incentive change. In addition to the substantive finding, we provide a way to account for the network value of consumers via counterfactual simulations. We also incorporate the dynamic structural model literature by demonstrating a way to incorporate both discrete and continuous actions into an integrated model that allows for the recovery of the value function.

Our findings can inform managers in several ways. First, we confirm the critical importance of the referral program in its contribution not only to the growth of user base (via more referrals), but also to the dynamic lifetime value of consumers, which help to accurately assess the value of the firm. Secondly, the referral incentive is a viable managerial control to experiment with and increase according to our static optimal counterfactual.

A limitation of this work is that we currently do not model consumers switching to the outside option, so therefore we can only assume the results hold for a firm acting as a monopoly with a captive user base. Since the acquisition of this data set, the industry has become an oligopoly, so we need to treat the pricing results with some caution. Future work may incorporate the pricing choices of competitors in order to arrive at a better pricing strategy. Another limitation of our work is that our model does not account for consumer heterogeneity. Therefore, we lose the ability to segment consumers by ranking the most valuable free consumers and then examining their choice behavior. Additionally, a model with heterogeneity will allow us to compare the differences in the decomposition of the their consumer value into the personal, referral, and social components among the most valuable consumers and the least valuable consumers. This may lead to more insights about the intrinsic nature of networked consumer lifetime value. Fortunately, the chosen Bayesian IJC framework allows for a natural extension for incorporating heterogeneity. In addition, the panel nature and the abundance of our data should ensure that our model will less likely be over-parametrized. Lastly, we have yet to account for social usage of consumers. A major usage component of the service is to help consumers share files with each other. Therefore, one can imagine that a portion of the customer lifetime value can be apportioned to the mutual sharing between consumers. Our model does not currently account for this shared usage, and one might speculate that our current estimate of a free consumer may be under-biased. We acknowledge this current limitation of our model and therefore that the results we present may been interpreted as a potential lower bound on the value of free consumers. However, by showing that the value of these free consumers do exist, we show that this area of research is worthy of future investigation.

# References

Bajari, Patrick, C Lanier Benkard, Jonathan Levin. 2007. Estimating dynamic models of imperfect competition. *Econometrica* **75**(5) 1331–1370.

Berger, Paul D, Nada I Nasr. 1998. Customer lifetime value: marketing models and applications. *Journal of interactive marketing* **12**(1) 17–30.

Biyalogorsky, Eyal, Eitan Gerstner, Barak Libai. 2001. Customer referral management: Optimal reward programs. *Marketing Science* **20**(1) 82–95.

Buttle, Francis A. 1998. Word of mouth: understanding and managing referral marketing. *Journal of strategic marketing* **6**(3) 241–254.

Chellappa, Ramnath K, Shivendu Shivendu. 2005. Managing piracy: Pricing and sampling strategies for digital experience goods in vertically segmented markets. *Information Systems Research* **16**(4) 400–417.

Ching, Andrew T, Susumu Imai, Masakazu Ishihara, Neelam Jain. 2012. A practitioner's guide to bayesian estimation of discrete choice dynamic programming models. *Quantitative Marketing and Economics* **10**(2) 151–196.

Fader, Peter S, Bruce GS Hardie, Ka Lok Lee. 2005. Rfm and clv: Using iso-value curves for customer base analysis. *Journal of Marketing Research* 415–430.

Gelman, Andrew, Donald Rubin. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* **7** 457–511.

Gupta, Sunil, Dominique Hanssens, Bruce Hardie, Wiliam Kahn, V Kumar, Nathaniel Lin, Nalini Ravishanker, S Sriram. 2006a. Modeling customer lifetime value. *Journal of Service Research* **9**(2) 139–155.

Gupta, Sunil, Carl Mela, Jose Vidal-Sanz. 2006b. The value of a free customer. *Harvard Business School Working Paper* .

Hanemann, W. Michael. 1984. Discrete/continuous models of consumer demands. *Econometrica* 541–61.

Heiman, Amir, Bruce McWilliams, Zhihua Shen, David Zilberman. 2001. Learning and forgetting: Modeling optimal product sampling over time. *Management Science* **47**(4) 532–546.

Heiman, Amir, Eitan Muller. 1996. Using demonstration to increase new product acceptance: Controlling demonstration time. *Journal of Marketing Research* 422–430.

Hotz, V. J., R. A. Miller. 1993. Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies* **60** 497–529.

Hotz, V Joseph, Robert A Miller, Seth Sanders, Jeffrey Smith. 1994. A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies* **61**(2) 265–289.

Imai, Susumu, Neelam Jain, Andrew Ching. 2009. Bayesian estimation of dynamic discrete choice models. *Econometrica* **77**(6) 1865–1899.

Jain, Dipak, Vijay Mahajan, Eitan Muller. 1995. An approach for determining optimal product sampling for the diffusion of a new product. *Journal of Product Innovation Management* **12**(2) 124–135.

Lehmann, Donald R, Mercedes Esteban-Bravo. 2006. When giving some away makes sense to jump-start the diffusion process. *Marketing Letters* **17**(4) 243–254.

Michelangeli, Valentina. 2008. Does it pay to get a reverse mortgage. *Doctoral Thesis Manuscript. Department of Economics, Boston University.* .

Miller, Robert A. 1984. Job matching and occupational choice. *The Journal of Political Economy* 1086–1120.

Needleman, S., A. Loten. 2012. When freemium fails. *The Wall Street Journal* .

Niculescu, Marius, DJ Wu. 2013. Economics of free under perpetual licensing: Implications for the software industry. *Available at SSRN 1853603* .

Pakes, Ariel, Michael Ostrovsky, Steven Berry. 2007. Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *The RAND Journal of Economics* **38**(2) 373–399.

Rust, J. 1987. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica* **55** 999–1033.

Ryan, Stephen P. 2012. The costs of environmental regulation in a concentrated industry. *Econometrica* **80**(3) 1019–1061.

Ryu, Gangseog, Lawrence Feick. 2007. A penny for your thoughts: Referral reward programs and referral likelihood. *Journal of Marketing* **71**(1) pp. 84–94. URL `http://www.jstor.org/stable/30162131`.

Schweidel, David A, Eric T Bradlow, Peter S Fader. 2011. Portfolio dynamics for customers of a multiservice provider. *Management Science* **57**(3) 471–486.

Silverman, George. 1997. How to harness the awesome power of word of mouth. *DIRECT MARKETING-GARDEN CITY-* **60** 32–37.

Song, Inseong, Pradeep K Chintagunta. 2007. A discrete: Continuous model for multicategory purchase behavior of households. *Journal of Marketing Research* 595–612.

Su, Che-Lin, Kenneth L Judd. 2012. Constrained optimization approaches to estimation of structural models. *Econometrica* **80**(5) 2213–2230.

Timmins, Christopher. 2002. Measuring the dynamic efficiency costs of regulators' preferences: Municipal water utilities in the arid west. *Econometrica* **70**(2) 603–629.

Wolpin, Kenneth I. 1987. Estimating a structural search model: the transition from school to work. *Econometrica* **55**(4) 801–817.

# A Appendix

## A.1 Derivation of Optimal Deletion and Jacobian

Take Equation 4, the Expected Value Function, and we define $V^d(\mathbf{S}, a, \nu; y, r; \Theta)$ to be the solution to the inner-maximization problem over the continuous action $d$:

$$EV(\mathbf{S};\Theta) = \mathbf{E}_{a,r^a,\varepsilon}\left[\max_{y,r\in\Gamma(z)} \mathbf{E}_\nu \left[\underbrace{\max_{d\in H(x,a,z,R^a)} u(\mathbf{D},\mathbf{S},\nu) + \beta EV(\mathbf{S}';\Theta)}_{\equiv V^d(x,z,R,R^a;a,\nu;y,r;\Theta)}\right] + \varepsilon(y,r)\right],$$

and therefore specifying $V^d(\cdot)$ in integral form is:

$$V^d(x,z,R,R^a;a,\nu;y,r;\Theta) = \max_{d\in H(x,a,z,R^a)} u(\cdot) + \beta \int_{a'}\int_{r^{a\prime}}\int_{\varepsilon'} \max_{y',r'} \int_{\nu'} V^d(x',z',R',R^{a\prime},a',\nu';y',r';\Theta)$$
$$dP\nu' + \varepsilon'(y',r')dP\varepsilon' dPr^{a\prime} dPa'. \tag{7}$$

Assuming optimal discrete choice solutions of $y$, $r$, $y'$ and $r'$, now we take the derivative of 7 to obtain a first-order condition with respect to the continuous action $d$:

$$\textbf{FOC}\text{: } \frac{\partial u(\cdot)}{\partial d} + \beta \frac{\partial x'}{\partial d} \int_{a'}\int_{r^{a\prime}}\int_{\varepsilon'} \max_{y',r'} \left\{ \int_{\nu'} V_x^d(x',z',R',R^{a\prime};a',\nu',y',r';\Theta)dP\nu' + \varepsilon'(y',r')\right\}$$
$$dP\varepsilon' dPr^{a\prime} dPa' = 0.$$

Where $V_x^d(\cdot)$ is the derivative of Equation 7 with respect to x. In addition, the Envelope Condition holds at optimal value of $d$, therefore we have an expression for $V_x^d(\cdot)$:

$$\textbf{EC}\text{: } V_x^d(\mathbf{S};a,\nu,y,r;\Theta) = \frac{\partial u(\cdot)}{\partial x} + \beta \frac{\partial x'}{\partial x} \int_{a'}\int_{r^{a\prime}}\int_{\varepsilon'} \max_{y',r'} \left\{ \int_{\nu'} V_x^d(\mathbf{S}';a',\nu',y',r';\Theta)dP\nu' + \varepsilon'(y',r')\right\}$$
$$dP\varepsilon' dPr^{a\prime} dPa'.$$

Given that the law of motion for the state space $x$ is $x' = x + a - d$, we have that $\frac{\partial x'}{\partial d} = -1$ and $\frac{\partial x'}{\partial x} = 1$. Combining the FOC, EC, $\frac{\partial x'}{\partial d}$, and $\frac{\partial x'}{\partial x}$, we arrive at the following equations:

$$\frac{\partial u(\cdot)}{\partial d} = \beta \int_{a'} \int_{r^{a\prime}} \int_{\varepsilon'} \max_{y',r'} \left[ \int_{v'} V_x^d(x',z',R',R^{a\prime};a',\nu',y',r';\Theta) dP\nu' + \varepsilon(y',r') \right] dP\varepsilon' dPr^{a\prime} dPa' \tag{8}$$

$$V_x^d(x,z,R,R^a,a,\nu,y,r,\Theta) - \frac{\partial u(\cdot)}{\partial x} = \frac{\partial u(\cdot)}{\partial d}. \tag{9}$$

Now we have everything we need to find the analytic expression. Since the form of the utility function is quadratic and transition law is linear, $V^d(\cdot)$ is of a linear form in the state variables and shocks:

$$V^d(x,z,R,R^a,a,\nu,y,r;\Theta) = \kappa \cdot \Gamma,$$

where $\kappa = (\kappa_1, ..., \kappa_8)$ , a vector of constants, and $\Gamma = (\mathbf{S}, a, \nu)$, a vector of the state variables, and exogenous shocks. From this, advance the time period by one, we arrive at:

$$\begin{aligned} V^d(x',z',R',R^{a\prime},a',\nu',y',r';\Theta) =& \kappa_1(x+a-d) + \kappa_2(\mathbf{1}[y=2]\cdot 52 + \mathbf{1}[y=1]\cdot 4 + \mathbf{1}[y=0]\cdot(z-1)) + \\ & \kappa_3(R+r) + \kappa_4(R^a+r^a) + \kappa_5 a' + \kappa_6 \nu' + \kappa_7 y' + \kappa_8 r'. \end{aligned}$$

Taking the derivative with respect to x would simply yield:

$$V_x^d(x',z',R',R^{a\prime},a',\nu',y',r';\Theta) = \kappa_1. \tag{10}$$

Now we have an expression for $V_x^d(x',z',R',R^{a\prime},a',\nu',y',r';\Theta)$, we come back and solve for this below. Next, recall the utility function from Equation 1, taking the derivative w.r.t $d$ and $x$, we plug back into Equation 8 to arrive at:

$$\begin{aligned} 2\alpha d + \nu =& \beta \int_{a'} \int_{r^{a\prime}} \int_{\varepsilon'} \left\{ \max_{y',r'} \int_{\nu'} \kappa_1 dP\nu' + \varepsilon(y',r') \right\} dP\varepsilon' dPr^{a\prime} dPa' \\ \kappa_1 =& \frac{2\alpha d + \nu}{\beta}. \end{aligned}$$

where $K$ is the total number of $y$-$r$ combination of choices. Now, combine all of the above into Equation 9 yields:

$$
\begin{aligned}
\kappa_1 - \theta =& 2\alpha d + \nu \\
\frac{2\alpha d + \nu}{\beta} - \theta =& 2\alpha d + \nu \\
2\alpha d + \nu =& \beta(2\alpha d + \nu + \theta) \\
\frac{\beta\theta}{1-\beta} =& 2\alpha d + \nu \qquad (11)
\end{aligned}
$$

At this point, we can solve for $d$ to get the optimal deletion policy function $d^*$:

$$
d^* = \frac{\beta\theta - \nu(1-\beta)}{2(1-\beta)\alpha} \qquad (12)
$$

The policy function only depends on the exogenous shock $\nu$ and primitives. However, the derivation from above assumes that our solution is not a corner solution. Therefore we include the boundary conditions, which depends on the states, to ensure the optimal $d^*$ value is in the feasible region:

$$
d^* = \max\left(\min\left(\frac{\beta\theta - \nu(1-\beta)}{2(1-\beta)\alpha},\ x+a\right),\ x+a-Q(z,R^a)\right).
$$

To obtain the $\nu$ inversion expression $g^{-1}(\cdot)$ that we use in the likelihood expression, we simply solve the Equation 11 in terms of $\nu$ instead of $d$. Thus we have:

$$
\nu \equiv g^{-1}(\cdot) = \frac{\beta\theta}{(1-\beta)} - 2\alpha d.
$$

To obtain the Jacobian $\left|\frac{\partial g^{-1}(\cdot)}{\partial d}\right|$, we simply take the derivative of the above $\nu$ inversion expression and obtain:

$$
\left|\frac{\partial g^{-1}(\cdot)}{\partial d}\right| = 2\alpha
$$