1-2015

Recovering Household Preferences for Digital Entertainment

Jin LI
Singapore Management University, jinli@smu.edu.sg

Zhiling GUO
Singapore Management University, ZHILINGGUO@smu.edu.sg

Robert J. KAUFFMAN
Singapore Management University, rkauffman@smu.edu.sg

Citation
LI, Jin; GUO, Zhiling; and KAUFFMAN, Robert J.. Recovering Household Preferences for Digital Entertainment. (2015). *48th Hawaii International Conference on System Sciences HICSS 2015: 5-8 January, Kauai: Proceedings*. 4276-4284.
Available at: https://ink.library.smu.edu.sg/sis_research/2599

2015 48th Hawaii International Conference on System Sciences
http://dx.doi.org/10.1109/HICSS.2015.512

Recovering Household Preferences for Digital Entertainment

[Δ‡]**Jin Li,** [Δ]**Zhiling Guo,** [Δ]**Robert J. Kauffman**
[Δ]Singapore Management University, [‡]City University of Hong Kong
{jinli, zhilingguo, rkauffman}@smu.edu.sg

Abstract

Cable TV return path data made possible by current generation set-top boxes present a new opportunity to analyze household viewing behavior and recover household viewing preferences from it. This research develops a model of household viewing preference that supports quantifying a household's valuation for different categories of digital content within the constraints of the programs to which it subscribes. This study uses a data set of more than 1 million observations on households from a digital entertainment firm that offers basic and premium services. Our estimation is via a Bayesian hierarchical model that employs the Gibbs sampler. The results show that households have relatively homogeneous preferences for entertainment content, but they show heterogeneous preferences for content in the specific packages to which they subscribe. In addition, both HD and premium movies subscriptions have a differentiation effect on enhancing household preferences toward their most preferred content. The findings provide useful insights for understanding household preferences, and are intended to support promotion and content strategy adjustments to improve customer satisfaction.

1. Introduction

With technology advances has come booming demand for digital entertainment [2, 10]. In the digital entertainment industry, satellite and Internet protocol TV (IPTV) revenues continue to grow, and will respectively reach nearly US$100 billion and US$26 billion by 2020. In contrast, global cable TV revenues peaked in 2011, and are expected to reach US$80.3 billion by 2020. Digital cable TV revenues will climb by 18% from their 2013 level to reach US$79.0 billion in 2020 [17]. Although cable TV represents an important channel for digital entertainment consumption, increasing competition from other substitutable services poses a significant challenge for the cable TV industry. Firms face pressure to prevent further subscriber loss and customer churn [14]. Understanding household preferences through their subscription choices and viewing behavior will be important to support firms for targeted marketing.

The cable TV industry can be analyzed in terms of the upstream supply of digital content, bundle design and pricing by digital entertainment vendors, and household-level subscription choices and viewing behavior [7, 9]. The main explanation for the widespread use of bundling is that it facilitates the extraction of consumer surplus by the cable TV operator in comparison to unbundled sales [1, 3]. There is an ongoing policy debate regarding unbundling in the United States. One of the main unbundling scenarios discussed by the Federal Communications Commission is called *themed tiers* [11], in which channels are bundled into different mini-tiers based on *channel genres*. These include such themes as sports and news, among others. This lets households pay for smaller tiers on an *a la carte* basis.

Crawford [8] and Byzalov [4] have evaluated the policy implications of the unbundling pricing strategies. Their empirical results suggest that consumers do not gain much from unbundling, however, their models are based on survey data from households and digital entertainment subscribers. This allows them to infer consumer preferences. The credibility of their models is limited by the quality of their data, and self-reported survey data are often viewed as inaccurate and sometimes untrustworthy. Using a proprietary data set consisting of household-level cable TV return path data from a digital entertainment services vendor's two-way set-up boxes, we provide new evidence about household preferences to support business policy analysis.

To understand household preferences better, we focus on studying household viewing behavior with respect to their bundle choices. Cable TV customers consume the content they prefer, and in this way demonstrate the extent to which they value it [4]. Consumer willingness-to-pay for a bundle of channels is driven by the value they believe they are getting from viewing the channels. Moreover, consuming cable TV programs is similar to consuming movies, music, apparel, groceries, and so on, for which consumers express demand for variety. We propose an additive non-linear utility model to capture the household's consumption of digital content and their variety-seeking behavior across different genres. Our model incorporates households' heterogeneous valuations and preference shifts due to viewing satiation over time for different content categories.

We observe evidence that shows households self-select different bundles based on unobserved viewing preferences. We sought a way to make it possible to recover their preferences through a variety of data analytics approaches. Specifically, we use micro-observations of household viewing behavior – essentially time spent on specific channels – as a basis for inferring valuation. Our view is that households express different marginal utilities of consumption, as well as different rates of satiation relative to their cumulative viewing time for any single genre. Also, their overall utility is additive across genres based on their viewing. The household-level cable TV return path data offer a useful reading of household consumption behavior, within the limits of the firm's ability to directly measure what people are watching on their TVs. From this, we demonstrate that it is possible to recover the distribution of household preferences and the value they place on different genres of bundled TV content.

This article seeks to answer these questions: (1) How do household preferences for various TV program genres differ and how can their preferences be quantified? (2) For households with the same subscription plans, how does consumption of different content contribute differently to households' overall valuation of the TV program bundle? (3) For households with different subscription plans, how does consumption of the same content contribute differently to the overall valuation of TV program bundles?

Through our analysis and empirical results, we are able to identify the genres for which all subscribers show strong preferences. We found that entertainment programs provide the highest utility for almost all households we sampled. We also obtained evidence that households subscribing to premium packages have clear and different preferences. We observe that high definition (HD) program subscriptions have the effect of improving viewing satisfaction at the cost of increasing rate of satiation for the preferred genre. These findings are important to guide the firm in crafting their marketing strategies.

2. Literature

According to Hsee et al. [12], customers exhibit different preferences when they purchase content versus when they consume it. When someone in a household chooses its bundle subscription, the goal is to select content that is the most beneficial for the members of the entire household. The utility that the choice creates for the members will be constrained by the selected channels in the bundle. A household's observed demand is often used to represent its preference for consumption. A commonly-used measure of demand is the household's viewing time for different channels and TV programs [8, 9]. Chang et al. [5, 6] clustered viewers into different preference groups according to how the members of households allocated their viewing time across different program genres and channels. They reported that households with more specific preferences exhibited lower channel viewing efficiency: they watched fewer programs and a smaller proportion of channels relative to other households with broader preferences. Thus, unbundling channels from themed tiers can benefit consumers by allowing them to selectively choose their most-valued channels.

To understand the policy implications, various authors have explored the usefulness of utility-based structural models for cable TV data. For example, Byzalov [4] proposed a two-stage model of demand for bundles of channels and TV viewing based on a random-utility discrete-choice framework. In the first stage, a household makes a decision to subscribe to a bundle of channels. In the second stage, observations of the viewing choices of each individual in a household will be revealed. In addition, Crawford and Yurukoglu [9] specified a Cobb-Douglas utility function defined on viewing times for different channels. Their household viewing utility model is based on the assumption that the more a household watches a channel, the more the household is willing to pay for it. They estimated the distribution of household preferences for each cable television channel without distinguishing among the levels content diversity across different channels.

A limitation of Byzalov's [4] and Crawford and Yurukoglu's [9] models is that a household's valuation of an hour of TV viewing is independent of the type of programs consumed. In reality though, consumers may derive different levels of utility when they watch an hour of sports versus an hour of news. Another is that their empirical evaluations are both based on weekly consumer recall-based surveys about their past TV viewing experience.

We will attempt to overcome these limitations by modeling household preferences for program content genres, rather than channels. We incorporated household-level preferences and satiation for each genre in our model to help identify their preference differences. In addition, because we have access to detailed observational data of household-level TV program viewing patterns, we are able to obtain insightful estimates of the covariance structure of program genre preferences.

3. Research Context and Data

3.1. Research Context

We utilized a set of TV program genres that were pre-defined by the digital entertainment firm that sponsored this research. They include TV programs in these genres: *Entertainment, Ethnic, Education,* *International, Lifestyle, News, Children, Movies,* and *Sports*. With this scheme, each channel that is offered by the firm can be uniquely assigned to one genre based on the majority of the content offered by the channel. For example, the Disney Channel is categorized in the *Children* genre.

Inferring household preferences based on the channels that the household members watch is standard in the literature [9]. Households pay a monthly subscription fee to access a digital entertainment package that contains different TV channels. Most consumers watch only a small fraction of the channels they are paying for though. Also, the channels they watch often represent only a subset of content genres or categories that are available. Sometimes, channels that are watched for only a short time may generate higher value than channels that are watched for a longer time. For example, high-profile sporting events such as World Cup soccer may be watched only for a short period of time relative to other programs. They may be of extraordinary value to sports fans though. This observation suggests the need to distinguish program content when household viewership patterns are used to infer its preference.

In addition, channels are substitutable. When the TV is turned on, a consumer who wants to watch the news will browse the available news channels to find suitable programs at that time, for example, Channel NewsAsia, CNN, or BBC. A three-digit code is commonly used as a channel identification number for each channel, with the first digit indicating the genre. This makes it easy for consumers to navigate up and down the channel list using the remote control to find relatively substitutable content. So classifying channels by genre is a reasonable proxy for the content attributes of a channel. Inferring preferences based on content, then, will be more accurate than inferring preferences based on channels. And genre-based preference inference is very useful for understanding themed tier pricing strategy.

Since consumers pay a monthly subscription fee, we define the observation period as a month for this research. Consumers typically trade off watching TV against other activities in life. They allocate the limited time they have to selectively viewing TV programs that they like, maximizing their overall utility in the process. The contribution of each genre of TV program they view will be a function of the household's consumption history, summarized by the cumulative viewing time that is observed for each of the program genres they watch. Channels within the same genre are likely to be imperfect substitutes for each other, and each genre also can be thought of as an imperfect substitute for the other genres. One may prefer to watch a sporting event, but watching news about it may suffice, for example.

A key observation about consuming digital content is that there is a point of *satiation*. This is true at the household level, just as it is at the individual level. A viewer who watches news about sporting events is likely to be satiated sooner by the news than by an actual sporting event. We see this with the high pay-per-view prices that viewers are willing to pay for some events (soccer, boxing, F1 racing, etc.). But as an individual's consumption history evolves, one expects to observe some shifts in the person's preferences among alternative genres. This should be observed at the household level too. Thus, we model monthly household consumption of TV program as a dynamic satiation process, in which the viewing of different program genres is accumulated in a way that maximizes the consumption utility of the household.

3.2. Data

Household data for this research represent a seven-month observation period between December 2012 and June 2013. Included are demographic information, subscription details, and viewing records for households. The demographic data include the subscriber's age band, gender, residence region, and dwelling type. The household subscription data details a household's subscription package. Both channel and bundle information is available, such as how many total channels, what basic groups, upsize channels, and HD and add-on channels a household subscribed to. Household viewing data were recorded with time-stamps and viewing durations for all of a household's subscribed channels.[1]

For our data analysis, we aggregated a household's viewing time for each genre by month, and deleted observations with zero total monthly viewing time. This resulted in 1.1 million observations.[2]

4. Modeling Viewing Preferences

4.1. Household Utility: A Key Construct

Let x_{ij} be household i's viewing time for genre j programs, with all $x_{ij} \geq 0$. We define an additive, non-linear utility function for household i across J different program genres: $U_i = \sum_j \beta_{ij}(x_{ij}+1)^{\alpha_{ij}}$. (See Appendix 1 for a table of our modeling notation.)

The parameter $\alpha_{ij} \in (0,1]$ influences household i's rate of diminishing marginal utility of consuming

[1] These data are all anonymized, so no individual or household identities are known or can be inferred.

[2] A majority of the households viewed five to eight out of the nine

[2] A majority of the households viewed five to eight out of the nine genres. This is evidence of variety-seeking behavior in digital content consumption.

genre j program. The coefficient $\beta_{ij} > 0$ measures household i's baseline utility for genre j's content when $x_{ij} = 0$. Since $\alpha_{ij} \leq 1$, the sub-utility function is concave in genre consumption, the smaller α_{ij} is, the higher the rate of satiation.[3] This utility specification can accommodate a wide variety of situations, including consuming all genres, a subset of genres, and only one genre.

Household viewing behavior is often constrained by time. Denote $c_i > 0$ as household i's maximum amount of monthly viewing time. Households maximize their utility by allocating their limited time to various genre-based program content, which leads to the consumption constraint $\sum_{j=1}^{J} x_{ij} \leq c_i$.

To develop a statistical specification of the model, we use a random utility approach and introduce a multiplicative lognormal error for the utility of each genre via $\tilde{U}_{ij} = U_{ij}\varepsilon_{ij}$. We further define the marginal utility of each genre as $\frac{\partial \tilde{U}_i}{\partial x_{ij}} = \frac{\partial U_i}{\partial x_{ij}}\varepsilon_{ij}$, where ε_{ij} is an independently and identically log-normally distributed error term. The lognormal error term specification is to enforce the positivity of the marginal utility. Thus, we have $ln\,(\frac{\partial \tilde{U}_i}{\partial x_{ij}}) = ln\,(\frac{\partial U_i}{\partial x_{ij}}) + \eta_{ij}, \eta_i \sim MVN(0, \Sigma)$, where η_i is the vector of error terms of the J genres and Σ is an identity matrix of size J. Random utility assumes that household knows about the values of η_i, which represents omitted factors that influence its marginal utility. We cannot observe them though.

4.2. Genre Viewing Likelihood

Household i maximizes its utility $\tilde{U}_i$, which includes the deterministic part U_i and the realization of the random utility error, subject to the household's consumption constraint and non-negative consumption $x_{ij} \geq 0$. Solving this household optimization problem we obtain the following set of first-order conditions:

$$V_{ij} + \eta_{ij} = ln\,(\lambda) \text{ if } x_{ij}^* > 0 \quad (1)$$

$$V_{ij} + \eta_{ij} < ln\,(\lambda) \text{ if } x_{ij}^* = 0 \quad (2)$$

where λ is the Lagrange multiplier, and $V_{ij} = ln\big(\alpha_{ij}\beta_{ij}(x_{ij}^* + 1)^{\alpha_{ij}-1}\big)$. (See Appendix 2 for solution details.) The optimality conditions define a mapping from η_i to x_i^*, where x_i^* is the vector of optimal consumption choices in a household for the J genres. The distribution of η_i offers a basis for deriving the distribution of x_i^*.

Define $v_{ij} = V_{i1} - V_{ij}$ and $\xi_{ij} = \eta_{ij} - \eta_{i1}$, for $j = 2, \dots, J$. The probability of m out of J genres being consumed is:

$$Pr\big(x_{ij}^*, m\big) = \int_{-\infty}^{v_{iJ}} \cdots \int_{-\infty}^{v_{i,m+1}} \phi\big(v_{i2}, \dots, v_{im}, \xi_{i,m+1}, \dots, \xi_{iJ} \big| 0, \Omega\big) \cdot |K|\, d\xi_{i,m+1} \cdots d\xi_{iJ}\text{ , for } m = 2, \dots, J-1,$$

where $\phi(\cdot)$ is normal density with mean 0 and covariance matrix Ω, and $|K|$ is the determinant of the Jacobian matrix with $K_{jk} = \frac{\partial v_{i,j+1}(x_i^*)}{\partial x_{i,k+1}^*}$, for $j, k = 1, \dots, m-1$.

For one genre (m = 1) or all genres (m = J), the probabilities are: $Pr\big(x_{ij}^*, 1\big) = \int_{-\infty}^{v_{iJ}} \cdots \int_{-\infty}^{v_{i2}} \phi(\xi_{i2}, \dots, \xi_{iJ})\, d\xi_{i2} \cdots d\xi_{iJ}$, and $Pr\big(x_{ij}^*, J\big) = \phi\big(v_{i2}, \dots, v_{iJ}\big|0, \Omega\big) \cdot |K|$.

This consumption probability equation suggests that, for genres with positive viewing time in a household, the demand is a non-linear function of ξ_{ij} = $v_{ij}(x_{ij}^*)$. For program genres that are not viewed, all $\xi_{ij} < v_{ij}(0)$ will produce the solution. Thus, we integrate the normal distribution of ξ_{ij} up to value $v_{ij}(0)$. The joint distribution in the consumption probability equation then can be evaluated by factoring it into discrete and continuous parts [13].

The likelihood function can be expressed as the product of all N households' likelihood functions over T periods as $L = \prod_{i=1}^{N}\prod_{t=1}^{T} Pr(x_{it}, m_{it})$, where x_{it} is consumer i's observed consumption at time t and m_{it} is the number of non-zero components in the vector x_{it}. Evaluation of the likelihood involves high-dimensional integrals of normal distributions. We present our findings in the next section.

5. Results

We defined five groups in our data based on the following criteria, and for all of which there is a basic subscription package required. They include: (1) no HD or add-ons; (2) with HD but no add-ons; (3) no HD but with at least one *Movies*-related add-on package; (4) no HD but with at least one *Sports*-related add-on package; and (5) no HD, with at least one *Movies*-related and one *Sports*-related add-on package. Group 1 is the least expensive. Group 2 lets households select a high-quality basic package. And Groups 3 to 5 represent premium packages with higher prices.

We randomly selected 500 households for each

[3] For example, if a particular genre has a high value of β_{ij} and a value of α_{ij} close to 1 to household i, then this household has high baseline preference and low satiation. We would expect the household to spend a large amount of time viewing genre j channels. But if $\alpha_{ij}, j = 1, \dots, J$, are very small and β_{ij} are not too different for household i, which imply a high satiation rate and similar baseline preference for different genres, we would expect a relatively even viewing time distribution across multiple genres.

group from our large data set. Since the five groups are mutually exclusive, this resulted in 5 × 500 = 2,500 households and 7 × 2,500 = 17,500 observations. Eliminating observations with no program viewing time in all genres produced our final data set for model estimation. Households that subscribe to more expensive packages have longer average viewing times. This justifies the assumption that household valuation in terms of willingness-to-pay for cable TV packages is driven by the utility they get from viewing TV programs offered by the package.

Model estimation involves two sets of parameters. The first set is household i's baseline sub-utility vector for all nine genres β_i. The second set is the household's satiation rate α for consuming certain genre's content. To simplify estimation, we set β_{ij} to be both household and genre-specific, while the satiation rate α_j is genre-specific. We set $\gamma_{ij} = ln(\alpha_j \beta_{ij})$ as the log transformation of baseline marginal utility evaluated at $x_{ij} = 0$. This is the highest marginal utility derived for each genre.

Because of the large number of household-specific parameters, we adopted the Markov chain Monte Carlo (MCMC) estimation method. We used the Gibbs sampler and Metropolis-Hasting methods to generate the parameters recursively based on a hierarchical Bayesian estimation framework [15, 16]. The estimation methods are shown in the Appendix 3.

Table 1 shows the five group estimates of $\bar{\gamma}$ and α. For the purpose of identification, we choose *Entertainment* as the base category and set its γ_i value to 0. The estimates of γ_i for other genres are relative values to *Entertainment*. A positive value indicates a household's higher marginal utility derived from consuming that genre's content. It provides a natural ranking of the household's baseline preferences for different genres. The estimates of the α's reveal different rates of satiation in genre consumption.

Table 1. Common parameter estimates

	$\bar{\gamma}$				
	Group 1	**Group 2**	**Group 3**	**Group 4**	**Group 5**
Entertainment	0	0	0	0	0
Ethnic	0.16(.119)	-3.05(.123)	-2.17(.155)	-0.81(.129)	-2.77(.150)
Education	-0.90(.078)	-1.29(.104)	-1.71(.100)	-1.34(.077)	-1.80(.123)
International	-1.96(.074)	-3.96(.101)	-2.62(.123)	-1.92(.112)	-2.31(.161)
Lifestyle	-1.94(.074)	-2.42(.111)	-2.51(.098)	-2.55(.087)	-2.34(.107)
News	-1.50(.068)	-3.75(.092)	-2.87(.095)	-1.93(.094)	-2.53(.125)
Children	-1.39(.098)	-3.28(.133)	-2.91(.117)	-2.13(.105)	-3.21(.123)
Movies	-3.71(.076)	-5.26(.085)	-0.31(.106)	-3.27(.116)	0.38(.128)
Sports	-2.76(.077)	-5.11(.106)	-4.16(.105)	-0.93(.100)	-1.78(.108)
	α				
	Group 1	**Group 2**	**Group 3**	**Group 4**	**Group 5**
Entertainment	0.49(.006)	0.22(.006)	0.28(.005)	0.37(.007)	0.29(.010)
Ethnic	0.56(.011)	0.49(.010)	0.46(.010)	0.47(.009)	0.44(.008)
Education	0.56(.009)	0.29(.009)	0.40(.006)	0.48(.008)	0.39(.008)
International	0.57(.010)	0.49(.009)	0.50(.011)	0.52(.008)	0.44(.014)
Lifestyle	0.57(.010)	0.35(.007)	0.39(.011)	0.51(.010)	0.37(.009)
News	0.59(.010)	0.49(.008)	0.49(.012)	0.52(.010)	0.45(.009)
Children	0.63(.011)	0.49(.008)	0.51(.012)	0.54(.009)	0.53(.008)
Movies	0.72(.012)	0.64(.008)	0.31(.010)	0.55(.013)	0.22(.008)
Sports	0.63(.010)	0.53(.016)	0.56(.011)	0.51(.009)	0.44(.008)

Note: Posterior standard deviations are shown in parentheses.

In the basic packages estimation, *Ethnic* program has a coefficient of 0.16, and shows a slightly higher valuation than *Entertainment*. This largely confirms that the majority of our sampled households in the basic package and they derive relatively high marginal utility by watching *Ethnic* programs. Their marginal utilities for *Movies* and *Sports* are not as high as other genres, largely because bulk of the content belongs to the add-on packages that require a premium subscription to be accessed. The satiation factor for this group ranges from 0.49 to 0.72. The smaller the α_j, the greater the satiation effect, so *Entertainment* has a higher rate of satiation.

Compared with Group 1, households in Group 2 are HD subscribers. They prefer *Entertainment* and *Education* rather than *Ethnic* programs. Although they also derive relatively low marginal utilities in *Movies* and *Sports*, the coefficient estimates of -5.26 (vs. -3.71) and -5.11 (vs. -2.76) suggest this group obtains even lower average marginal utility from consuming those contents than the base group. The range of HD subscribers' marginal utilities gets wider, and the rate of satiation becomes faster (from 0.22 to 0.64). This provides evidence that the HD feature may help differentiate the household's marginal utilities for different genres. All else equal, the households' viewing time distribution is likely to be concentrated around a few of the most preferred genres.

For households that subscribe to *Movies* add-ons, *Sports* add-ons, and both, we see clear preferences for those genres. For example, Group 3's households that subscribed to *Movies* add-on channels derived the second highest marginal utility from them (-0.31), very close to the first-ranked *Entertainment* genre. This group has a very low marginal utility for *Sports* though (-4.16). The same observations are true for Groups 4 and 5. Compared with the base subscription group, the viewing utilities for *Sports* (0.38) and *Movies* (-1.78) are substantially higher for households that subscribe to the premium packages in Group 5. Their marginal utilities for *Ethnic* (-2.77) and *Children* (-3.21) are the lowest; this suggests that these households are likely not to focus on *Ethnic* programs and probably do not have kids.

A sample covariance-correlation matrix for Group 1 is shown in Table 2 to provide more information. The same matrix can be produced for each sample, which will offer similar qualitative insights.

The covariance matrix among all genres is presented in the lower triangle of the table, and the correlations between genres are in the upper triangle. The largest variances are observed for the *Ethnic* (3.84) and *Children* (3.20) genres, indicating substantial heterogeneity in the sample. This is reasonable since not all households speak the language that the

Ethnic package focuses on, and not all households have children. Also, the correlations between genres range from 0.08 to 0.63. The least correlated genres are *Ethnic* and *Lifestyle*, and *Children* and *Lifestyle*. The most correlated genres are *Education* and *Lifestyle.* This indicates that when a household prefers to view *Lifestyle* programs, it is highly likely that the household prefers the *Education* programs as well.

Table 2. Covariance and correlations

Covariance and Correlations Estimates of $\{\gamma_i\}$ for Group 1								
	Et	*Ed*	*I*	*L*	*N*	*C*	*M*	*S*
Et	3.84 (.30)	0.25	0.14	0.08	0.32	0.39	0.35	0.41
Ed	0.71 (.16)	2.19 (.18)	0.17	0.63	0.16	0.19	0.29	0.35
I	0.36 (.13)	0.32 (.10)	1.62 (.16)	0.25	0.24	0.33	0.34	0.44
L	0.21 (.15)	1.34 (.15)	0.45 (.11)	2.08 (.19)	0.21	0.08	0.29	0.18
N	0.72 (.14)	0.26 (.10)	0.34 (.09)	0.35 (.10)	1.31 (.13)	0.13	0.18	0.34
C	1.37 (.20)	0.51 (.14)	0.74 (.13)	0.21 (.14)	0.27 (.12)	3.20 (.27)	0.30	0.41
M	0.83 (.16)	0.52 (.14)	0.54 (.11)	0.51 (.13)	0.25 (.10)	0.67 (.14)	1.51 (.15)	0.38
S	1.12 (.16)	0.73 (.12)	0.78 (.11)	0.37 (.12)	0.54 (.10)	1.03 (.15)	0.65 (.11)	1.94 (.16)

Note: Lower triangle is for covariance and the upper one is for correlations. Posterior standard deviations are shown in the parentheses. The abbreviations in the column headers and rows, as follows: *Et*: *Ethnic; Ed*: *Education*; *I*: *International*; *L*: *Lifestyle*; *N*: *News*; *C*: *Children*; *M*: *Movies*; *S*: *Sports.*

Figures A1 and A2 in Appendix 4 show sub-utility curves for the genres for households that subscribe to the basic packages with and without the HD option. We see that both *Movies* and *Sports* provide low base utilities (the intercepts). Subscribers that select the HD option show a faster diminishing rate of marginal utility than households without it. Figures A3 and A4 show the sub-utility curves for subscribers with at least one premium add-on package. *Movies* and *Sports* provide high base utilities for these households. Meanwhile, similar to the HD option, *Movies* add-ons, as shown in Figure A3, help differentiate the sub-utilities of different genres. In contrast, *Education, Lifestyle, International* and *New*s are relatively neutral genres that are not strongly preferred or disliked by any group.

6. Managerial and Strategy Implications

Our viewership-based utility model can help to understand customer viewing preferences and satiation levels for different subscription bundles. It produces useful insights for marketing managers to make customized promotions to targeted households based on their genre preferences. The customer's satisfaction level can be increased, as can access to a higher proportion of preferred channels and content. From the vendor's perspective, customized promotions help to reduce marketing cost and increase marketing efficiency and effectiveness.

Churn management. To tackle the problem of customer defection or *churn*, companies usually rank their customers' propensity to churn, and implement retention policies for targeted customers with a high propensity to leave. Understanding customers through our model and providing customized services may reduce their potential churn rate, as this may improve their satisfaction for the vendor's service offerings. On the other hand, making appropriate recommendations and promotions to customers also may enable the provider to co-create value by adjusting household subscriptions. This is beneficial since it will be possible to elicit information on the customer's true preferences and create higher satisfactions for the digital entertainment content that is offered. In addition to reducing the churn rate, this may also have another beneficial consequence: raising the customer's willingness-to-pay.

Upselling. Based on a household's current digital entertainment subscription, the service provider can promote channels based on the extent to which the household's genre preferences can be identified. *Upselling* is encouraging the customer to buy higher quality versions or extensions of what they are already consuming. For example, we observed each household's total viewing time and the viewing time allocation for each genre to which they subscribed. For those households that show strong preferences for a specific genre and were still not satiated for this genre, it may be that these households would like to have access to even more content. For current subscription plans beyond the basic bundles, customers can choose to subscribe to more channels within the same genre or purchase premium add-on packages that are similar to the genres of the basic bundles. Channels with their preferred genre can be flagged and add-on packages can be recommended to encourage more consumption.

Cross-selling. To detect customer preferences and extend them by providing channels with genres they enjoy is a form of *cross-selling* of digital entertainment. It can be supported through data analytics that identify the various channels and programming that will extend the household's preference by inferring what other content genres they will like. For example, we observed a high correlation between the consumption of the *Education* and *Lifestyle* genres in the basic subscription group. Currently, for our data, this type of subscription requires households to choose at least three themed tiers from a total of seven tiers that are available. Each tier can be uniquely mapped to a specific genre. For households that subscribe to the *Education* genre but not *Lifestyle* genre programs, it may be beneficial to recommend to them the *Lifestyle* package.

7. Conclusion

We developed a structural model of household decision-making for digital entertainment delivered to the home via cable TV services. From data that we obtained, and for which we extracted five samples of 500 households, we demonstrated a method that permits inferences about household preferences for cable TV program bundles and their viewing choices.

We learned that, conditional on their subscriptions, households tend to self-select for the purchase of different bundles depending on their unobserved viewing preferences. All subscribers showed strong preferences for programs in the *Entertainment* genre. Also, households that subscribed to the more basic packages seem to value *Ethnic* program bundles – with cultural tie-ins, the household's primary language, and other attributes – the highest. In contrast, the households that subscribed to premium packages clearly preferred *Movies* and *Sports*, and they probably have little willingness-to-pay or preference for *Children*-related programs. We further found that HD subscriptions tend to be associated with the enhancement of a household's preferences, along with an increasing rate of satiation for their preferred genre. Also, both the HD feature and *Movies* add-ons have a *sub-utility differentiation effect*. Households that choose the HD program option or subscribe to premium packages, including *Movies* add-ons, seem to value their viewing experiences in the most preferred genres much higher than in other genres.

Our model and results can help a digital entertainment vendor to predict customer viewing valuation and satiation for different bundle subscriptions. Based on such a better understanding of its customers, a vendor can optimize its marketing strategies via churn management, upselling and cross-selling.

There are several directions for future research that are based on the current limitations of this work. Our findings suggest that household demographics are likely to be strong predictors of consumer bundle choices. For example, specific ethnic groups that are represented at the household level tend to choose the basic bundles, and households with no children tend to subscribe to the premium packages. Adding more detailed household demographics to the model will be helpful to achieve more fine-grained insights.

Our current model does not consider households' bundle subscription decisions, but instead takes the household subscriptions as given. We have asserted that households optimize their digital entertainment viewing experiences by allocating their limited viewing time to the consumption of programs from different genres. To address this limitation, it may be appropriate to build a two-stage conditional choice model in which households choose subscription bundles in the first step. Then, given a household's subscription, its members will make decisions about the programs they watch. On the other hand, different bundle subscriptions can be viewed as treatment effects. An upsizing or HD effect can be incorporated into a unified model to make estimation and prediction. Based on this extension, the vendor can get clear about how the bundle option affects household viewing valuation and satiation.

We focused on the retail market in this research. The operation of the wholesale market for cable TV programming content also should have an effect on the operator's pricing and bundling choices. Future work to extend our analysis can move from the household to the digital entertainment vendor side of the market. For example, considering bundle design will help the vendor to identify different household-level price sensitivity and support pricing strategy. We can also leverage viewership data to identify different bundles through clustering rather than using the pre-defined bundle options. For this, a counterfactual policy simulation can be used to demonstrate the business value of new clustered bundle configurations. This will be interesting since it allows exploration of how the configuration of cable TV packages, optimal bundle and unbundled pricing strategies, and upstream content acquisition costs and constraints affect the benefits that households can obtain, as well as the extent that cable operators will be profitable.

Acknowledgement and Data Disclaimer

This research was done under a binding non-disclosure agreement with the corporate sponsor, requiring the anonymization of all data that were used and analyzed. No personally-identifiable information about customers has been disclosed or shared. This research was supported by the Singapore National Research Foundation under the International Research Centre @ Singapore Funding Initiative, administered by the Interactive Digital Media Programme Office (IDMPO).

References

[1] Adams, W.J., Yellen, J.L. Commodity bundling and the burden of monopoly. *Qtrly. J. Econ.*, 90(3), 1976, 475-498.

[2] Bakker, G. The evolution of entertainment consumption and the emergence of cinema, 1890-1940. Working paper, London Sch. Econ., London, UK, 2007.

[3] Bakos, Y., Brynjolfsson, E. Bundling information goods: pricing, profits, and efficiency. *Mgmt. Sci.,* 45(12), 1999, 1613-1630.

[4] Byzalov, D. Unbundling cable television: an empirical investigation. Working paper, Temple University, Philadelphia, PA, 2010.

[5] Chang, R.M., Kauffman, R.J., Son, I. Consumer micro-behavior and TV viewership patterns: data analyt-

ics for the two-way set-top box. *Proc. 14th Intl. Conf. Elec. Comm*., ACM Press, New York, NY, 2012.
[6] Chang, R., Ghosh P., Jung G., Kauffman R.J., Zhang P. Do household cable TV viewing patterns demonstrate efficiency and concentration? 7th China Summer Wkshp. Info. Mgmt., Tianjin, China, June 2013.
[7] Chipty, T. Vertical integration, market foreclosure, and consumer welfare in the cable television industry. *Amer. Econ. Rev.*, 91(3), 2001, 428-453.
[8] Crawford, G. S. The discriminatory incentives to bundle in the cable television industry. *Quant. Mktg. Econ.*, 6(1), 2008, 41-78.
[9] Crawford, G.S., Yurukoglu, A. The welfare effects of bundling in multichannel television markets. *Amer. Econ. Rev.,* 102(2), 2012, 643-685.
[10] Eskicioglu, A.M., Town, J., Delp, E.J. Security of digital entertainment content from creation to consumption. *Info. and Mgmt.*, 44(1), 2003, 53-59.
[11] Federal Communication Commission. Further report on the packaging and sale of video programming to the public. Washington, DC, 2006.
[12] Hsee, C.K., Yang, Y., Gu, Y., Chen, J. Specification seeking: how product specifications influence consumer preference. *J. Cons. Res.*, 35(6), 2009, 952-966.
[13] Kim, J., Allenby G., Rossi P. Modeling consumer demand for variety. *Mktg. Sci.*, 21(3), 2002, 229-250.
[14] Liebowitz, S.J., Zentner A. Clash of the titans: does Internet use reduce television viewing? *Rev. Econ. Stat.*, 94(1), 2012, 234-245.
[15] Netzer, O., Lattin, J.M., Srinivasan, V. A Hidden markov model of customer relationship dynamics. *Mktg. Sci.*, 27(2), 2008, 185-204.
[16] Rossi, P.E., Allenby, G.M. Bayesian statistics and marketing. *Mktg. Sci.* 22(3), 2003, 304-328.
[17] The Digital TV Europe.net. Fast growth 'over' for the global pay TV industry. May 28, 2014.

Appendix 1. Modeling Notation

NOTATION	DEFINITION	COMMENTS
J	Number of genres	9 genres
N	Number of households	500 households
T	Number of observing periods	7 periods
α_{ij}	Household i's diminishing rate of marginal utility for genre j programs	$\alpha_{ij} \in (0,1]$
β_{ij}	Household i's baseline utility for genre j's program	$\beta_{ij} > 0$
γ_{ij}	Log transformation of household i's highest marginal utility for genre j programs	$\gamma_{ij} = ln(\alpha_j \beta_{ij})$
x_{ij}	Household i's monthly viewing time for genre j	$x_{ij} \geq 0$
c_i	Household i's total monthly viewing time	$c_i > 0$
U_{ij}	Household i's utility for viewing genre j programs	
U_i	Household i's utility across J program genres	$U_i = \sum_{j=1}^{J} U_{ij}$
ε_{ij}	Lognormal error terms for household i's random utility from viewing genre j program	
Σ	Identity matrix	
η_{ij}	Normal error terms for household i's log transformation of marginal utility from viewing genre j programs	$\eta_{ij} = ln\varepsilon_{ij}$ $\eta_i \sim MVN(0, \Sigma)$
ξ_{ij}	Difference for normal error terms	$\xi_{ij} = \eta_{ij} - \eta_{i1}$
$\widetilde{U}_{ij}$	Household i's random utility from viewing genre j program	$\widetilde{U}_{ij} = U_{ij}\varepsilon_{ij}$
V_{ij}	Log transformation of household i's marginal utility from viewing genre j program	
v_{ij}	Difference of household i's marginal utility	$v_{ij} = V_{i1} - V_{ij}$
λ, μ_j	Lagrange multipliers	
m	Number of genres viewed by a household	$m = 1, \ldots, J$
K	Jacobian matrix produced in the change of variable process	

Appendix 2. For Likelihood Function

The Lagrangian for the maximization problem is $\Lambda\left(x_{ij}, \lambda, \mu_j\right) = \widetilde{U}_i + \lambda\left(c_i - \sum_{j=1}^{J} x_{ij}\right) + \sum_{j=1}^{J} \mu_j x_{ij}$, where λ and μ_j, for $j = 1, \ldots, J$, are Lagrange multipliers. Differentiating w.r.t. x_{ij} gives $\frac{\partial \widetilde{U}_i}{\partial x_{ij}} - \lambda + \mu_j = 0$. At the optimum, the consumption constraint will be binding. Complementary slackness implies that $\sum_{j=1}^{J} x_{ij}^* = c_i$ and $\lambda > 0$. The relevant conditions are $\frac{\partial \widetilde{U}_i}{\partial x_{ij}} - \lambda = 0$ if $x_{ij}^* > 0$, or $\frac{\partial \widetilde{U}_i}{\partial x_{ij}} - \lambda < 0$ if $x_{ij}^* = 0$. Substituting into the Lagrangian and taking logs, these conditions can be rewritten as equations (1) and (2) in the main text. The equality constraint $\sum_{j=1}^{J} x_{ij}^* = c_i$ induces a singularity in the distribution of x_i^*. For this, we adopted the standard approach to difference the conditions in Equations 1 and 2 with respect to one genre. Since $c_i > 0$, households view channels in at least one genre. We assume the first genre is consumed. For estimation, we rearranged the genre order of each observation to make sure the first genre's viewing time is positive. This has reduced the dimensions of the vector of parameters to be estimated by 1.

Appendix 3. Hierarchical Bayesian Estimation

The *Gibbs sampler* is the most basic MCMC method used in Bayesian statistics. By transforming $ln\left(\alpha_j \beta_{ij}\right)$ into γ_{ij}, we can recursively generate draws of parameters from these conditional distributions: (1) $\gamma_i \mid X_i, \alpha, \bar{\gamma}, V_\gamma$; (2) $\bar{\gamma} \mid \{\gamma_i\}, V_\gamma$; (3) $V_\gamma \mid \{\gamma_i\}, \bar{\gamma}$; and (4) $\alpha \mid X, \{\gamma_i\}$. Since both $\bar{\gamma}$ and V_γ have conjugate distributions, we can use Gibbs sampling to make draws. X_i is a matrix to denote household i's observations of historical viewing time vectors for all genres, and X is a matrix for all individuals' X_i. As γ_i and α do not belong to conjugate distributions, we use Metropolis-Hastings to make draws for the parameters.

To generate $\boldsymbol{\gamma_i}$. The conditional distribution of γ_i can be expressed as: $f(\gamma_i \mid X_i, \alpha, \bar{\gamma}, V_\gamma) \propto N(\bar{\gamma}, V_\gamma) L(X_i|\gamma_i, \alpha)$ $\propto V_\gamma^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\gamma_i-\bar{\gamma})'^{V_\gamma^{-1}}(\gamma_i-\bar{\gamma})\right] L(X_i|\gamma_i, \alpha)$. We use the random walk Metropolis-Hastings algorithms to make new draws. The proposal density used here has the form $\gamma_i^* = \gamma_i^{(k)} + s\Delta$, where s is a positive scale parameter and Δ is multivariate normal with mean vector 0 and identity covariance matrix. We set $s = 0.75$. Based on the ratio for the random walk of Metropolis-Hastings, the acceptance probability can be expressed as:

$Pr(Accept) = min\left\{\frac{exp\,[-\frac{1}{2}(\gamma_i^*-\bar{\gamma})'V_\gamma^{-1}(\gamma_i^*-\bar{\gamma})]L(X_i|\gamma_i^*,\alpha)}{exp\,[-\frac{1}{2}(\gamma_i^{(k)}-\bar{\gamma})'V_\gamma^{-1}(\gamma_i^{(k)}-\bar{\gamma})]L(X_i|\gamma_i^{(k)},\alpha)}, 1\right\}$. So γ_i^* will be accepted as a new draw $\gamma_i^{(k+1)} = \gamma_i^*$ with this probability; and $\gamma_i^{(k+1)} = \gamma_i^{(k)}$ otherwise.

To generate $\boldsymbol{\bar{\gamma}}$. Note that $\bar{\gamma} \mid \{\gamma_i\}, V_\gamma \sim MVN(u_{\bar{\gamma}}, V_{\bar{\gamma}}), V_{\bar{\gamma}} = [(Z'Z \otimes V_\gamma^{-1}) + V_0^{-1}]^{-1},\ u_{\bar{\gamma}} = V_{\bar{\gamma}}[(Z' \otimes V_\gamma^{-1})\Gamma^* + V_0^{-1}u_0]$. Also, $Z = (1,1, \dots 1)'$ is a vector of 1's with length N, $\Gamma^* = vec(\Gamma')$, and $\Gamma = (\gamma_1', \gamma_2', \dots, \gamma_N')$ is a N*dim(γ) matrix. The *diffuse priors* of u_0 and V_0 are defined as: $u_0 = (0,0, \dots 0)'$, a vector of 0s with length of dim(γ); and $V_0 = 100 I_{dim\,(\gamma)}$.

To generate $\boldsymbol{V_\gamma}$. In Bayesian statistics, the Wishart distribution is the conjugate prior of the inverse covariance matrix of a multivariate normal random vector. We specify the prior for V_γ as $V_\gamma^{-1} \sim Wishart(G, g)$. It follows that $V_\gamma \mid \{\gamma_i\}, \bar{\gamma} \sim IW(\sum_{i=1}^{N}[(\gamma_i-\bar{\gamma})(\gamma_i-\bar{\gamma})'] + G^{-1}, g+N)$, where g and G are prior hyperparameters. We set $g = dim(\gamma) + 5$, with $G = I_{dim\,(\gamma)}$.

To generate $\boldsymbol{\alpha}$. The conditional distribution of α can be expressed as $f(\alpha|X, \{\gamma_i\}) \propto N(\alpha_0, V_{\alpha_0}) L(X|\alpha, \{\gamma_i\}) \propto$ $V_{\alpha_0}^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\alpha-\alpha_0)'^{V_{\alpha_0}^{-1}}(\alpha-\alpha_0)\right] \cdot L(X|\alpha, \{\gamma_i\})$. Similar to the method we use to generate γ_i, we set the related scale parameter as 0.01. The probability of acceptance of a draw is:

$$Pr(Accept) = min\left\{\frac{exp\,[-\frac{1}{2}(\alpha^*-\alpha_0)'V_{\alpha_0}^{-1}(\alpha^*-\alpha_0)]L(X|\alpha^*,\{\gamma_i\})}{exp\,[-\frac{1}{2}(\alpha^{(k)}-\alpha_0)'V_{\alpha_0}^{-1}(\alpha^{(k)}-\alpha_0)]L(X|\alpha^{(k)},\{\gamma_i\})}, 1\right\}.$$

Appendix 4. Subscription Group Sub-Utility Curves

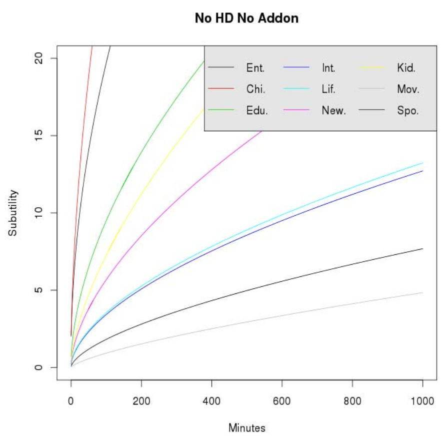

Figure A1: Subscription Group 1

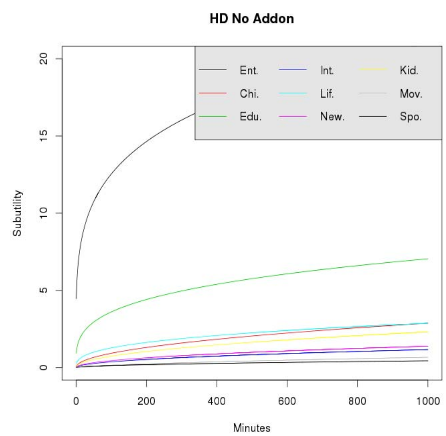

Figure A2: Subscription Group 2

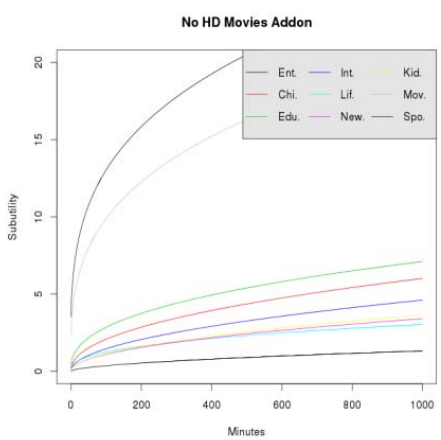

Figure A3: Subscription Group 3

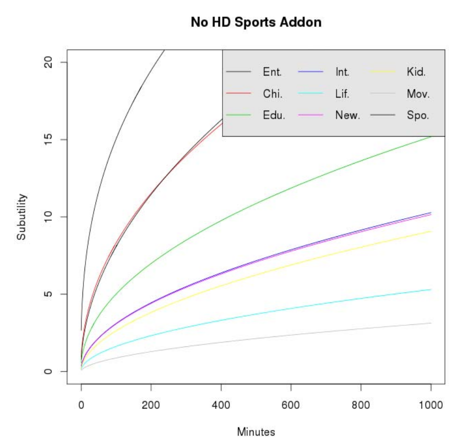

Figure A4: Subscription Group 4