MIS Quarterly

BIG DATA & ANALYTICS IN NETWORKED BUSINESS

MOBILE APP ANALYTICS: A MULTIPLE DISCRETE-CONTINUOUS CHOICE FRAMEWORK

Sang Pil Han
Department of Information Systems, W. P. Carey School of Business, Arizona State University,
PO Box 874106, Tempe, AZ 85287 U.S.A. {shan73@asu.edu}

Sungho Park
Department of Marketing, W. P. Carey School of Business, Arizona State University,
PO Box 874106, Tempe, AZ 85287 U.S.A. {spark104@asu.edu}

Wonseok Oh
College of Business, Korea Advanced Institute of Science and Technology,
85 Hoegiro Dongdaemoon-Gu, Seoul, KOREA 130-722 {wonseok.oh@kaist.ac.kr}

Appendix A

Mobile Apps Used and Websites Visited by Panel Members in the Data

Table A1. Mobile Content Categories: Mobile Apps and Mobile Websites

Category		Count	Example
Communication	Mobile Messengers, Mobile Internet Phone, Email	241	Kakao Talk, Mypeople Messenger, Phone, GO SMS Pro, LINE, NateOn UC, LightSMS, Gmail, LightSMS, Viber, Skype
Entertainment	Book, Cartoon, Adults, Sports, Humor, Magazine	1,054	Naver Webtoon, Live Scores, TIViewer, T store Book, jjComics Viewer, Naver Books, Score Center
Game	Action, Adventure, Board, Puzzle, Racing, Role Playing, Shooting, Simulation, Sports	2,870	Rule The Sky, TinyFarm, Smurfs' Village, Shoot Bubble Deluxe, Hangame, 2012 Baseball Pro, Angry Birds Space, Jewels Star
Map and Navigation	Map, Navigation	340	T map, Google Maps, SeoulBus, Naver Maps, Olle Navi, Subway Navigation,
Lifestyle	Weather, News, Education, Restaurants, Job, Health, Religion, Fashion	1,811	YTN News, MK News, SBS News, Weather, Bible, Newspapers, JobsKorea
Personal Finance	Banking, Stocks, Finance, Real Estate, Ecommerce	414	Smart Trading, M-Stock Smart, KB Star Banking, Auction Mobile, Coupang, Gmarket Mobile
Music and Radio	Radio, Music	341	Music Player, SKY Music, PlayerPro, Mnet, MyMusicOn, FM Radio, Soribada
Photo	Photo Gallery, Camera	297	Gallery, Camera, Cymera, Photo Editor, Instagram
Portal Search	Portal Site, Search Engine	74	Naver, Daum, NATE, Google Search, Junior Naver

Table A1. Mobile Content Categories: Mobile Apps and Mobile Websites (Continued)

Category		Count	Example
Schedule	Scheduler, Memo, Alarm Clock	1,389	Address book, Alarm/Clock, Calendar, Memo, Polaris Office, Polaris Office, Docviewer
Social	Social Networking Service, Board, Blog, Microblog	165	Kakao Story, Facebook, Twitter, Naver Café, Daum Café, Cyworld, Naver Blog, Me2day
Utilities	Productivity, Decoration, Webhard, Widget, Firewall	2,125	Alarm Clock, Calculator, Voice Recorder, HD Browser, Dropbox, Battery Widget
Video	Multimedia, Broadcasting, Movie	290	TV, Youtube, MX Video Player, SKY Movie, Afreeca TV, T-DMB, PandoraTV
Mobile Web	Websites	8,043	Naver.com, Daum.net, Nate.com, Google.co.kr, ppomppu.co.kr, dcinside.com, facebook.com
Sum		19,591	

Appendix B

Correlation Matrices of App Choice Incidence and App Use Time

Table B1. Correlation Matrix of App Choice Incidence

	Communi-cation	Game	Map/ Navigation	Entertain-ment	Lifestyle	Personal Finance	Music & Radio	Photo	Portal	Schedule/ Memo	Social Network	Utility	Video	Web
Communication		**0.07**	**0.10**	**0.05**	**0.09**	**0.08**	0.00	**0.20**	**0.10**	**0.51**	**0.11**	**0.36**	**0.11**	**0.02**
Game	**0.07**		**0.06**	**0.17**	**0.10**	**0.05**	**0.08**	**0.10**	**0.09**	**0.10**	**0.10**	**0.12**	**0.13**	**0.06**
Map/Navigation	**0.10**	**0.06**		**0.11**	**0.18**	**0.17**	**0.13**	**0.14**	**0.14**	**0.12**	**0.08**	**0.14**	**0.13**	**0.06**
Entertainment	**0.05**	**0.17**	**0.11**		**0.16**	**0.08**	**0.15**	**0.09**	**0.12**	**0.06**	**0.12**	**0.08**	**0.13**	**0.05**
Lifestyle	**0.09**	**0.10**	**0.18**	**0.16**		**0.17**	**0.11**	**0.16**	**0.15**	**0.12**	**0.13**	**0.15**	**0.14**	**0.07**
Personal Finance	**0.08**	**0.05**	**0.17**	**0.08**	**0.17**		**0.06**	**0.11**	**0.13**	**0.11**	**0.11**	**0.14**	**0.11**	**0.06**
Music & Radio	0.00	**0.08**	**0.13**	**0.15**	**0.11**	**0.06**		**0.14**	**0.11**	**0.05**	**0.13**	**0.07**	**0.15**	**0.05**
Photo	**0.20**	**0.10**	**0.14**	**0.09**	**0.16**	**0.11**	**0.14**		**0.18**	**0.23**	**0.22**	**0.24**	**0.22**	**0.08**
Portal	**0.10**	**0.09**	**0.14**	**0.12**	**0.15**	**0.13**	**0.11**	**0.18**		**0.15**	**0.13**	**0.14**	**0.16**	**-0.03**
Schedule/Memo	**0.51**	**0.10**	**0.12**	**0.06**	**0.12**	**0.11**	**0.05**	**0.23**	**0.15**		**0.11**	**0.36**	**0.15**	**0.03**
Social Network	**0.11**	**0.10**	**0.08**	**0.12**	**0.13**	**0.11**	**0.13**	**0.22**	**0.13**	**0.11**		**0.12**	**0.15**	**0.05**
Utility	**0.36**	**0.12**	**0.14**	**0.08**	**0.15**	**0.14**	**0.07**	**0.24**	**0.14**	**0.36**	**0.12**		**0.16**	**0.05**
Video	**0.11**	**0.13**	**0.13**	**0.13**	**0.14**	**0.11**	**0.15**	**0.22**	**0.16**	**0.15**	**0.15**	**0.16**		**0.10**
Web	**0.02**	**0.06**	**0.06**	**0.05**	**0.07**	**0.06**	**0.05**	**0.08**	**-0.03**	**0.03**	**0.05**	**0.05**	**0.10**	

Note: Bold: significant at the .01 level.

Table B2. Correlation Matrix of App Use Time

	Communi-cation	Game	Map/ Navigation	Entertain-ment	Lifestyle	Personal Finance	Music & Radio	Photo	Portal	Schedule/ Memo	Social Network	Utility	Video	Web
Communication		**-0.02**	0.01	**0.02**	**0.04**	0.00	**0.13**	**0.37**	**0.10**	**0.11**	**0.27**	**0.12**	0.01	**0.05**
Game	**-0.02**		-0.01	**0.03**	**0.04**	-0.01	0.01	**-0.02**	0.01	0.00	-0.01	**0.06**	0.01	-0.01
Map/Navigation	0.01	-0.01		-0.01	**0.03**	0.01	0.00	0.00	0.00	**0.05**	-0.01	**0.04**	0.01	-0.01
Entertainment	**0.02**	**0.03**	-0.01		0.00	**-0.02**	**0.04**	**-0.02**	**0.08**	0.00	0.00	**0.02**	**0.02**	0.00
Lifestyle	**0.04**	**0.04**	**0.03**	0.00		0.06	**0.07**	0.01	**0.08**	**0.07**	0.01	**0.10**	**0.09**	**0.05**
Personal Finance	0.00	-0.01	0.01	**-0.02**	**0.06**		0.00	-0.01	**0.08**	**0.04**	**-0.02**	**0.05**	**0.10**	**0.05**
Music & Radio	**0.13**	0.01	0.00	**0.04**	**0.07**	0.00		**0.07**	**0.06**	**0.03**	**0.08**	**0.05**	**0.09**	**0.06**
Photo	**0.37**	**-0.02**	0.00	**-0.02**	0.01	-0.01	**0.07**		**0.09**	**0.04**	**0.31**	**0.09**	**0.02**	**0.09**
Portal	**0.10**	0.01	0.00	**0.08**	**0.08**	**0.08**	**0.06**	**0.09**		**0.04**	**0.09**	**0.06**	**0.11**	**0.09**
Schedule/Memo	**0.11**	0.00	**0.05**	0.00	**0.07**	**0.04**	**0.03**	**0.04**	**0.04**		0.01	**0.09**	**0.06**	0.01
Social Network	**0.27**	-0.01	-0.01	0.00	0.01	**-0.02**	**0.08**	**0.31**	**0.09**	0.01		**0.06**	**0.04**	**0.08**
Utility	**0.12**	**0.06**	**0.04**	**0.02**	**0.10**	**0.05**	**0.05**	**0.09**	**0.06**	**0.09**	**0.06**		**0.09**	**0.07**
Video	0.01	0.01	0.01	**0.02**	**0.09**	**0.10**	**0.09**	**0.02**	**0.11**	**0.06**	**0.04**	**0.09**		**0.09**
Web	**0.05**	-0.01	-0.01	0.00	**0.05**	**0.05**	**0.06**	**0.09**	**0.09**	**0.01**	**0.08**	**0.07**	**0.09**	

Note: Bold: significant at the .01 level.

Appendix C

Details on the Estimation and the Scalability of the Proposed Model

In this appendix, we provide a guideline for the estimation of the proposed model. The estimation follows the general steps of simulated maximum likelihood estimation. Below we outline each step of the procedure.

1. Determine initial values for all model parameters (α_0, $\bar{\beta}$, Ξ_β, Σ_β, Π_β, Γ_β, Λ_β, $\bar{\lambda}$, Ξ_λ, Σ_λ, Π_λ, Γ_λ, Λ_λ).

2. Draw H random values for each of ψ_h, υ_h, φ_h, and ν_h from the standard normal distribution. Note that H is the number of panel members in the data. Statistical packages usually provide random number generators and these can be used. However, we use the Halton draws to generate random numbers because it is known to improve the simulation performance. See Chapter 9 of Train (2009) for detailed discussion on this.

3. Using the parameter values and random numbers drawn in the previous step, calculate the following:

$$M! \cdot \left(\prod_{j=0}^{M} \frac{1-\alpha_{hjt}}{q_{hjt}+\tau_j} \right) \cdot \left(\sum_{j=0}^{M} \frac{q_{hjt}+\tau_j}{1-\alpha_{hjt}} \right) \cdot \frac{\prod_{i=0}^{M} e^{V_{hit}}}{\left(\sum_{j=0}^{J} e^{V_{hjt}} \right)^{M+1}}$$

   Note that this is the integrand in (10).

4. Repeat steps 2 and 3 many times and take averages of computed integrands. This average is the simulated likelihood for the given parameter values. In the model estimation, we repeat the process 100 times using the Halton draws to obtain the simulated likelihood.

5. Maximize the simulated likelihood value (computed through step 1 to step 4) by updating model parameters. GAUSS package is used for

this numerical maximization. Specifically, we use a quasi-Newton numerical optimization algorithm developed by Broyden, Fletcher, Goldfarb, and Shanno (BFGS). The BFGS algorithm is an iterative method for solving unconstrained nonlinear optimization problems. The BFGS method is one of the most popular quasi-Newton numerical optimization algorithms. Moreover, this algorithm is suited to problems with large numbers of variables and has proven to have good performance even for non-smooth optimization problems.

We have 23,222 observations (1366 users × 17 weeks) and 363 parameters to be estimated. The computation time of likelihood value (step 1 through step 4) is 6.82 seconds with a desktop PC running at 2.8 GHz. Note that, like many other packages, GAUSS supports multi-thread computing and we calculate the likelihood using eight threads with eight-core processor. Multi-threading can reduce the computing time almost linearly. If one use 16 thread with a 16-core processor running at the same speed, the computation time will be reduced by a half.

To see how computation time changes as the number of observations changes in our model, we reduce the number of observation by a half (n = 11,611). We find that the computation time is 2.97 second (44% of 6.82 seconds).

Currently, we consider 15 alternatives in the main model. The total estimation time of this model is about seven days. If the number of alternatives increases, the model parameters will increase almost linearly and we can expect the similar increase in the likelihood computation time. However, the number of iterations needed for the numerical optimization (step 5) will increase due to the elevated problem complexity, resulting in substantial increase in total estimation time. It seems that there are no general results specifying the relationship between the number of alternatives (or model parameters) and nonlinear optimization time.

We also want to direct interested researchers to Chandra Bhat's web page (http://www.caee. utexas.edu/prof/bhat/FULL_CODES.htm) where his GAUSS and R codes for MDCEV model and its variants are available along with technical notes.

Appendix D

Model Estimation Results in Mobile Competitive Analysis

The results in Table D1 show that mobile users' baseline utility for Kakao Talk (a communication app) is the highest and that their baseline utility for Rule the Sky (a game app) is the lowest among the top nine apps. Consistent with the descriptive findings, the baseline utility for Kakao Talk is the highest. Demographic variables account for the substantial heterogeneity in the baseline utilities for all mobile users. For example, older users exhibit a lower intrinsic preference for Kakao Talk (a communication app), Facebook (a social network app), Mellon (a music/radio app), and Naver (a portal search app). Moreover, although users exhibit inertia in their app choices in all apps, they show stronger tendencies of inertial use of game apps (Rule the Sky and TinyFarm). Finally, the baseline utilities for Naver, Kakao Story, and YouTube significantly increase over the sample period.

The results in Table D2 show that the satiation level is the highest for YouTube and the lowest for Tiny Farm among the popular apps. A game app (TinyFarm) shows the lowest satiation, and this finding is consistent with our results based on the app category-level data. We find that there is a substantial user heterogeneity in terms of satiation levels. For example, older users show significantly lower satiation levels for TinyFarm, but exhibit a significantly higher satiation level for Kakao Talk and Mellon. The state dependence coefficients are all significantly negative and this indicates that users show notably lower satiation levels (or longer usage time) this week if the app was used last week. This can be arguably interpreted as the learning reinforcement effect, and/or habituation in app use. The estimated time trend coefficient of TinyFarm is significant and positive, capturing mobile users' decreasing time allocation for TinyFarm over time. This might be explained by boredom with the same game over time.

Table D1. Estimates for Baseline Utility Parameters: App-Level Analysis

		Demographic Variables								
	Constant	Age 30's	Age 40's & Over	Female	Income Mid-class	Income Upper-class	Education High School Graduate	Education University Graduate	State Dependence	Time Trend
KakaoTalk	**-19.88**	**-0.30**	**-0.41**	**0.36**	**-0.13**	**-0.12**	-0.07	-0.04	**2.29**	0.007
	(0.40)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.07)	(0.06)	(0.07)	(0.004)
Naver	**-22.04**	**-0.21**	**-0.31**	**0.12**	**-0.08**	-0.04	**-0.27**	**-0.12**	**3.80**	**0.010**
	(0.39)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.07)	(0.06)	(0.04)	(0.004)
Kakao Story	**-21.21**	-0.06	**-0.41**	**0.22**	-0.06	**-0.12**	**0.18**	0.05	**2.65**	**0.007**
	(0.39)	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.06)	(0.05)	(0.03)	(0.003)
RuleTheSky	**-25.14**	0.11	-0.09	**0.34**	**-0.29**	**-0.31**	0.11	-0.08	**6.49**	0.014
	(0.44)	(0.14)	(0.18)	(0.12)	(0.13)	(0.15)	(0.21)	(0.18)	(0.15)	(0.013)
Facebook	**-21.88**	**-0.73**	**-0.69**	**0.10**	0.04	0.03	**-0.45**	**-0.32**	**3.66**	0.003
	(0.39)	(0.06)	(0.06)	(0.04)	(0.05)	(0.05)	(0.08)	(0.06)	(0.05)	(0.005)
TinyFarm	**-24.32**	-0.17	**-0.55**	**0.21**	-0.05	-0.15	0.13	-0.01	**5.66**	0.014
	(0.41)	(0.10)	(0.13)	(0.08)	(0.09)	(0.11)	(0.14)	(0.10)	(0.10)	(0.010)
Mellon	**-22.12**	**-0.63**	**-0.90**	**0.12**	0.10	**0.17**	**-0.24**	**-0.40**	**3.17**	-0.001
	(0.39)	(0.07)	(0.08)	(0.05)	(0.06)	(0.07)	(0.10)	(0.07)	(0.05)	(0.006)
Daum	**-23.10**	0.01	-0.10	0.00	-0.09	-0.09	**-0.30**	**-0.23**	**4.52**	0.007
	(0.40)	(0.08)	(0.09)	(0.06)	(0.06)	(0.07)	(0.11)	(0.09)	(0.06)	(0.006)
YouTube	**-20.97**	-0.07	-0.09	**0.17**	0.00	**-0.12**	**-0.28**	**-0.26**	**1.29**	**0.009**
	(0.39)	(0.05)	(0.05)	(0.03)	(0.04)	(0.04)	(0.07)	(0.05)	(0.03)	(0.004)
Others	**-16.84**	-0.09	**-0.21**	0.05	**-0.12**	**-0.11**	**-0.38**	**-0.25**	–	0.002
	(0.39)	(0.05)	(0.06)	(0.04)	(0.05)	(0.05)	(0.08)	(0.06)	–	(0.002)

Note: Standard errors in parentheses. Bold: significant at the .05 level. The estimated value of α_0 is -5.37 (standard error: 0.19). The state dependence parameter of "Others" cannot be separately identified from its constant because the "Others" option is always selected by all users.

Table D2. Estimates for Satiation Parameters: App-Level Analysis

		Demographic Variables								
	Constant	Age 30's	Age 40's & Over	Female	Income Mid-class	Income Upper-class	Education High School Graduate	Education University Graduate	State Dependence	Time Trend
KakaoTalk (Communication)	**2.83**	**0.57**	**0.79**	**-0.31**	-0.04	**-0.08**	**0.44**	**0.42**	**-0.80**	**-0.010**
	(0.08)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.07)	(0.002)
Naver (Portal)	**4.15**	**0.14**	**0.11**	**-0.11**	**-0.13**	-0.03	**0.22**	**0.11**	**-1.18**	**-0.014**
	(0.07)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.06)	(0.04)	(0.05)	(0.003)
Kakao Story (Social)	**4.48**	0.02	**0.40**	**-0.71**	0.06	**0.24**	**-0.22**	**-0.24**	**-0.55**	-0.002
	(0.05)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.04)	(0.003)
RuleTheSky (Game)	**2.74**	-0.19	**-0.56**	-0.07	0.10	**-0.22**	0.28	-0.20	**-1.14**	0.011
	(0.20)	(0.11)	(0.14)	(0.09)	(0.10)	(0.11)	(0.15)	(0.12)	(0.17)	(0.009)
Facebook (Social)	**3.92**	**0.36**	**0.40**	**-0.28**	0.03	**-0.11**	**0.39**	**0.51**	**-1.01**	-0.004
	(0.07)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.07)	(0.04)	(0.05)	(0.004)
TinyFarm (Game)	**2.06**	**-0.43**	**-1.04**	0.00	-0.02	-0.02	-0.01	0.01	**-0.47**	**0.014**
	(0.12)	(0.08)	(0.11)	(0.06)	(0.07)	(0.08)	(0.10)	(0.08)	(0.09)	(0.007)
Mellon	**4.46**	**0.50**	**0.80**	**-0.17**	-0.18	**-0.31**	**0.55**	0.22	**-1.62**	-0.017
(Music/Radio)	(0.13)	(0.12)	(0.14)	(0.08)	(0.10)	(0.12)	(0.19)	(0.12)	(0.09)	(0.010)
Daum	**4.23**	**-0.27**	-0.04	0.02	**-0.19**	**-0.18**	0.19	**0.36**	**-1.13**	**-0.023**
(Portal)	(0.11)	(0.08)	(0.08)	(0.06)	(0.06)	(0.07)	(0.11)	(0.09)	(0.08)	(0.006)
YouTube	**4.89**	-0.03	**0.23**	**-0.14**	0.05	-0.03	-0.11	0.06	**-0.90**	**-0.013**
(Video)	(0.07)	(0.06)	(0.07)	(0.05)	(0.05)	(0.06)	(0.09)	(0.07)	(0.04)	(0.005)
Others	**1.25**	**0.14**	**0.16**	0.01	0.00	0.02	-0.04	0.04	–	0.002
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	–	(0.002)

Note: Standard errors in parentheses. Bold: significant at the .05 level. The state dependence parameter of "Others" cannot be separately identified from its constant because the "Others" option is always selected by all users.

References

Train, K. 2009. *Discrete Choice Methods with Simulation* (2nd ed.), New York: Cambridge University Press.