Learning from Your Friends' Check-Ins: An Empirical Study of Location-Based Social Networks (Web Appendix)

Liangfei Qiu
Warrington College of Business
University of Florida, Gainesville, FL 32611
E-mail: liangfei.qiu@warrington.ufl.edu

Zhan (Michael) Shi
W.P. Carey School of Business
Arizona State University, Tempe, AZ 85287
E-mail: zmshi@asu.edu

Andrew B. Whinston
McCombs School of Business
University of Texas at Austin, Austin, TX 78705
E-mail: abw@uts.cc.utexas.edu

Appendix A: Calculations of Social Network Centrality Measures

All social network measures describe in Table 1 are calculated using a Python language software package, NetworkX (see Table A.1).[1]

Table A.1. How Social Network Measures are Calculated

degree_centrality(G)	Compute the degree centrality for nodes.
clustering(G[, nodes, weight])	Compute the clustering coefficient for nodes.
closeness_centrality(G[, u, distance, ...])	Compute closeness centrality for nodes.
betweenness_centrality(G[, k, normalized, ...])	Compute the shortest-path betweenness centrality for nodes.

[1] http://networkx.github.io/documentation/networkx-1.9.1/overview.html

Appendix B: Robustness Checks

B.1 Instrument Variable II: Check-ins of Friends' Friends

We conduct additional robustness checks using the average number of check-ins of user *i*'s friends' friends that are not user *i*'s friends and have *less than or equal to x common friends* with user *i*, where *x* = 1, 3, 5. The estimation results are presented in Columns 1, 2, and 3 of Table B.1.

Table B.1 Additional Robustness Checks

	(1)	(2)	(3)	(4)	(5)
	IV(II): x = 1	IV(II): x = 3	IV(III): x=5	Random coefficients	NMF
γ_1	0.625***	0.636***	0.684***	0.714***	0.721***
	[3.473]	[3.522]	[3.733]	[4.106]	[4.352]
γ_2	-0.0176**	-0.0208**	-0.0226**		-0.0231**
	[2.053]	[2.116]	[2.217]		[2.162]
δ_1	-0.508***	-0.529***	-0.542***	-0.462***	-0.485***
	[3.453]	[3.753]	[3.774]	[3.149]	[3.268]
δ_2	-0.0121	-0.0117	-0.0139		-0.0148
	[1.234]	[1.385]	[1.562]		[1.653]
φ_1	0.00409	0.00498	0.00472	0.00571	0.00519
	[0.227]	[0.302]	[0.265]	[0.264]	[0.311]
φ_2	-2.23e-5	-2.53e-5	-2.74e-5		-2.14e-5
	[0.141]	[0.136]	[0.128]		[0.123]
ρ_1	-0.00224	-0.00269	-0.00256	-0.00272	-0.00275
	[0.405]	[0.374]	[0.388]	[0.413]	[0.362]
ρ_2	-1.27e-05	-1.19e-05	-1.25e-05		-1.58e-05
	[0.154]	[0.177]	[0.136]		[0.172]
τ	0.215***	0.227***	0.269***	0.229***	0.265***
	[2.865]	[3.324]	[3.586]	[3.461]	[3.962]
$\boldsymbol{M_i}$	Yes	Yes	Yes	Yes	Yes
Monthly dummies	Yes	Yes	Yes	Yes	Yes
σ_γ				0.253***	
				[2.874]	
σ_δ				0.624***	
				[3.083]	

z or t statistics in brackets, * *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

B.2 Random Coefficients Specification

A random coefficients model allows heterogeneous observational learning effect across users. The main model in the quality learning stage is as follows:

$$U_{ijt} = \alpha_j + \gamma_i R_{ij,t-1} + \varphi_1 S_{ij,t-1} + \beta_j V_{ij,t-1}$$

$$+\delta_i R_{ij,t-1} V_{ij,t-1} + \rho_1 S_{ij,t-1} V_{ij,t-1} + \Theta \boldsymbol{M_i} + \tau G_{jt} + Month_t + \varepsilon_{ijt}.$$

In this model, each user has fixed coefficients γ_i and δ_i in her utility function, but these coefficients depend on both her observable characteristics and unobserved random factors. We denote the average values of γ_i and δ_i across users as γ_1 and δ_1, and assume the following specification for user heterogeneity:

$$\begin{bmatrix}\gamma_i \\ \delta_i\end{bmatrix} = \begin{bmatrix}\gamma_1 \\ \delta_1\end{bmatrix} + \begin{bmatrix}\Pi_\gamma \\ \Pi_\delta\end{bmatrix} M_i + \begin{bmatrix}\omega_{ij\gamma} \\ \omega_{ij\delta}\end{bmatrix},$$

where M_i is a 4×1 vector of user *i*'s observable characteristics, including the degree centrality, closeness centrality, betweenness centrality, and individual clustering coefficient summarized in Table 1, and $\omega_{ij\gamma}$ and $\omega_{ij\delta}$ represent unobserved individual heterogeneity at restaurant *j* and follow normal distribution $N(0, \sigma_\gamma^2)$ and $N(0, \sigma_\delta^2)$, respectively. The first source of consumer learning heterogeneity is from observable factors: $\begin{bmatrix}\Pi_\gamma \\ \Pi_\delta\end{bmatrix} M_i$. The parameter matrices Π_γ and Π_δ measure the observable heterogeneity of the learning effect across consumers that depends on individual centrality measures in the location-based social network. The second source of consumer learning heterogeneity is from unobserved random factors: $\begin{bmatrix}\omega_{ij\gamma} \\ \omega_{ij\delta}\end{bmatrix}$. The parameters σ_γ^2 and σ_δ^2 measure the unobserved heterogeneity of observational learning effect across consumers, meaning that different users will be affected differently by observational learning. The results are shown in column 4 in Table B.1.

B.3 Nonnegative Matrix Factorization (NMF)

The intuition of our NMF method is that consumers' tastes and the underlying social network among consumers are simultaneously determined by some hidden lower dimensional feature space. Following Ma et al. (2008), and Shi and Whinston (2013), we modify the structural model by changing the restaurant heterogeneity α_j in equation (2) to restaurant-consumer heterogeneity α_{ij}, and assume that:

$$\alpha_{ij} = \pi_{j0} + \pi_{j1}c_{i1} + \cdots + \pi_{jq}c_{iq} + \cdots + \pi_{jQ}c_{iQ},$$

where c_{iq} represents consumer *i*'s latent preference factors, and the venue-specific π_{jk} are parameters to be estimated. The latent factors c_{iq} are obtained by factorizing the social network graph, and π_{jq} are jointly estimated with other parameters by maximum likelihood estimation (MLE). In order to reduce the computational burden of MLE, we choose $Q = 3$. The computation is carried out by applying the standard NMF procedures in Shi and Whinston (2013). In column 5 of Table B.1, we find that the effect of observational learning is still significant after controlling for the individual-level NMF latent factors.

Appendix C: A Simple Analytical Model of the Role of Friends' Check-Ins

In this appendix, we use a simple analytical model to illustrate that (i) observing one friend's check-in is a weak quality signal, and observing multiple friends' check-ins is a much stronger quality signal; (ii) the marginal effect of friend's check-ins is decreasing (concavity); and (iii) the impact of an additional friend's check-in is smaller when a consumer is more certain about the restaurant quality before observing friends' check-ins.

Suppose that the true quality of a restaurant is a random variable, V. A focal consumer's prior belief on V is given by:

$$V \sim N(V_0, 1/\rho_V),$$

where V_0 is the mean of the prior, and ρ_V is the precision of the prior belief. The focal consumer *i* can observe her friends' check-ins at the restaurant in a location-based app. More specifically, a friend's check-in is modelled as a private signal about the restaurant quality:

$$S_j = V + \varepsilon_j,\ \ \varepsilon_j \sim N(0, 1/\rho_\varepsilon),$$

where ρ_ε is the precision of consumer i's information for her friends' check-ins $j = 1, 2, \ldots, n$ (suppose that consumer *i* receives *n* friends' check-ins). Consumer *i* makes a Bayesian inference using her prior belief and the signals contained in her friends' check-ins:

$$\mathbf{E}[V|I_i] = \frac{\rho_V}{n\rho_\varepsilon + \rho_V} V_0 + \frac{\rho_\varepsilon}{n\rho_\varepsilon + \rho_V} \sum_{j=1}^{n} S_j\,,\ \ \mathbf{Var}[V|I_i] = 1/(n\rho_\varepsilon + \rho_V),$$

where I_i is the information set of consumer i. Essentially, consumer i's conditional expectation, $\mathbf{E}[V|I_i]$, is a weighted average of the prior mean and the signals contained in her friends' check-ins. If consumer *i* observes only one friend's check-in, her variance of ex-post belief on the restaurant quality is $1/(\rho_\varepsilon + \rho_V)$, which is much larger than the variance of ex-post belief when she observes multiple check-ins, $1/(n\rho_\varepsilon + \rho_V)$, where $n > 1$. In other words, when a consumer receives one friend's check-in, she is still uncertain about the restaurant quality. Receiving multiple friends' check-ins will reduce quality uncertainty. Therefore, we can obtain that (i) observing one friend's check-in a weak quality signal, and observing multiple friends' check-ins is a much stronger quality signal.

Then, we define the information value of friends' check-ins as the reduction of quality uncertainty:

$$\mathbf{Var}[V] - \mathbf{Var}[V|I_i] = (1/\rho_V) - [1/(n\rho_\varepsilon + \rho_V)],$$

where $\mathbf{Var}[V]$ is the ex-ante quality uncertainty (the variance of prior belief on quality), and $\mathbf{Var}[V|I_i]$ is the ex-post quality uncertainty (the variance of ex-post belief on quality). We can show that

$$\frac{\partial\{\mathbf{Var}[V] - \mathbf{Var}[V|I_i]\}}{\partial n} > 0,$$

which implies that an additional friend's check-in can help reduce quality uncertainty. As the number of check-ins, n, increases, the reduction of quality uncertainty, $\mathbf{Var}[V] - \mathbf{Var}[V|I_i]$ becomes larger ($\mathbf{Var}[V|I_i]$ decreases with n). Again, it confirms our result (i). Furthermore, we also find:

$$\frac{\partial^2\{\mathbf{Var}[V] - \mathbf{Var}[V|I_i]\}}{\partial n\,\partial n} < 0.$$

It implies that the reduction of quality uncertainty increases with the number of check-ins, n, but (ii) the marginal effect is decreasing (concavity). Finally, we can obtain:

$$\frac{\partial^2\{\mathbf{Var}[V] - \mathbf{Var}[V|I_i]\}}{\partial \rho_V\,\partial n} < 0,$$

which implies that the impact of an additional friend's check-ins is smaller when the precision of the prior belief is higher. The intuition is that if consumer i has very precise prior information on the restaurant quality before observing her friends' check-ins, the information contained in friends' check-ins is less valuable. Therefore, we obtain that (iii) the impact of an additional friend's check-in is smaller when a consumer is more certain about the restaurant quality before observing friends' check-ins.

Appendix D: Additional Estimation Results under Different Time Frames

In this appendix, we provide the estimation results using only after Feb 15, 2011 data in column 1 in Table D.1. The results are consistent. Because we have only one snapshot of the location-based network, a potential bias is that the friendships that formed after Feb 15, 2011 would affect the estimation of the observational learning effect. More specifically, $R_{ij,t-1}$ (the number of a focal user's friends' check-ins) used in our estimation may be smaller than the true number of friends' check-ins because we do not account for the friendships that the focal user formed after Feb 15, 2011 ($S_{ij,t-1}$ will not be affected by the potential friendship formation because it is the number of anonymous people's check-ins). Therefore, we may over-estimate the impact of observational learning using the data after Feb 15, 2011. However, we think the bias should not be large because of two reasons: (i) Before early 2010 (the early stage of the application), the user-base grew quickly. After Feb 15, 2011, the network operated on this application had entered a relatively stable period. (ii) We have also estimated our model using the data from August 2010 to July 2011 (column 2 in Table D.1). On the one hand, this shorter time frame is closer to the snapshot time of the network graph, Feb 15, 2011. On the other hand, if the bias mentioned earlier is significant, we should observe that the magnitude of observational learning estimated using the data from August 2010 to July 2011 is significantly different from the magnitude of observational learning estimated using the data after Feb 15, 2011. The reason is that $R_{ij,t-1}$ after Feb 15, 2011 will tend to be smaller than the true number of friends' check-ins, while $R_{ij,t-1}$ before Feb 15, 2011 will tend to be greater than the true number of friends' check-ins. However, we find that the magnitude of observational learning estimated using the data from August 2010 to July 2011 is very similar to the magnitude of observational learning estimated using the data after Feb 15, 2011, which indicates that the potential bias may not be large.

Table D.1 The Effect of Observational Learning: Different Time Frames

	(1) After Feb 15, 2011	(2) August 2010 – July 2011
γ_1	0.872***	0.854***
	[4.663]	[4.376]
γ_2	-0.0267**	-0.0284**
	[2.231]	[2.215]
δ_1	-0.714***	-0.685***
	[4.387]	[4.214]
δ_2	-0.0127*	-0.0118
	[1.868]	[1.687]
φ_1	0.00573	0.00668
	[0.324]	[0.475]
φ_2	-2.21e-05	-2.35e-05
	[0.186]	[0.324]
ρ_1	-0.00252	-0.00241
	[0.386]	[0.377]
ρ_2	-1.19e-05	-1.46e-05
	[0.104]	[0.139]
τ	0.212***	0.243***
	[3.054]	[3.389]
$\boldsymbol{M_i}$	Yes	Yes
Monthly dummies	Yes	Yes

z or *t* statistics in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix E: Self-Selection Bias

To account for potential self-selection biases, we consider a Heckman-type model in this appendix. For instance, younger people may be more likely to register for the app and then post mobile check-ins. Older people may physically visit the restaurant, but do not leave mobile check-ins because they have not registered for the app. Following Cameron and Trivedi (2009), we assume that Y_{ijt} (whether a consumer posts a mobile check-ins at a restaurant) is observed only when a consumer has registered for the app:

$$Y_{ijt} = \begin{cases} Y_{ijt}^* & \text{if } w_i^* > 0 \\ - & \text{if } w_i^* \leq 0 \end{cases},$$

where Y_{ijt} is observed only when $w_i^* > 0$, and w_i^* is a latent variable. The selection equation is given by:

$$w_i^* = \beta_0 + \beta_1 Gender_i + \beta_2 Age_i + u_i.$$

Assume that the errors ε_{it} and u_{it} follow a bivariate normal with mean zero and covariance matrix $\begin{bmatrix} 1 & \sigma_{\varepsilon u} \\ \sigma_{\varepsilon u} & \sigma_u^2 \end{bmatrix}$. The estimation results are presented in Table E.1. We find that the results are robust.

Table E.1 The Effect of Observational Learning: Heckman-Type Model

	(1)
	Heckman-Type Model
γ_1	0.712***
	[3.894]
γ_2	-0.0203**
	[2.054]
δ_1	-0.486***
	[3.533]
δ_2	-0.00852
	[1.632]
φ_1	0.00643
	[0.489]
φ_2	-2.33e-05
	[0.206]
ρ_1	-0.00229
	[0.361]
ρ_2	-1.27e-05
	[0.116]
τ	0.185***
	[2.894]
$\boldsymbol{M_i}$	Yes
Monthly dummies	Yes

z or t statistics in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix F: Additional Estimation Results Using Demographic Information

In this appendix, we present the results estimated using only consumers who report their demographic information. The results are consistent (see Table F.1).

Table F.1 The Effect of Observational Learning: User Demographic Information

	(1) User Demographic Information
γ_1	0.842***
	[4.359]
γ_2	-0.0263**
	[2.057]
δ_1	-0.583***
	[3.625]
δ_2	-0.0107*
	[1.823]
φ_1	0.00557
	[0.749]
φ_2	-2.55e-05
	[0.216]
ρ_1	-0.00237
	[0.462]
ρ_2	-1.38e-05
	[0.155]
τ	0.236***
	[3.171]
Gender (1=female; 0=male)	0.0312
	[1.186]
Age (years)	-0.00107
	[0.688]
$\boldsymbol{M_i}$	Yes
Monthly dummies	Yes

z or *t* statistics in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix G: Relaxing the i.i.d Assumption of the Error Term

In our main model, we assume that ε_{ijt} follows an *i.i.d* distribution. In other words, even if consumer *i* visits restaurants *j* and *k* in the same time period *t*, ε_{ijt} and ε_{ikt} are independent. This might be a strong assumption in certain contexts. In this appendix, we relax this assumption and allow that ε_{ijt} is correlated with ε_{ikt}, where $k \neq j$. We assume that correlation between ε_{ijt} and ε_{ikt} is $\rho_{\varepsilon 1}$ (a parameter to be estimated). The estimation results are presented in column 1 in Table G.1.

Table G.1 The Effect of Observational Learning: Correlations in ε_{ijt}

	(1)	(2)
γ_1	0.664***	0.702***
	[3.728]	[3.842]
γ_2	-0.0227**	-0.0215**
	[2.145]	[2.304]
δ_1	-0.417***	-0.435***
	[3.163]	[3.338]
δ_2	-0.0143	-0.0125
	[1.653]	[1.533]
φ_1	0.00522	0.00674
	[0.287]	[0.385]
φ_2	-2.03e-5	-2.16e-5
	[0.137]	[0.166]
ρ_1	-0.00212	-0.00237
	[0.327]	[0.364]
ρ_2	-1.32e-05	-1.65e-05
	[0.127]	[0.138]
τ	0.221***	0.242***
	[3.217]	[3.542]
$\boldsymbol{M_i}$	Yes	Yes
Monthly dummies	Yes	Yes
$\rho_{\varepsilon 1}$	0.147*	0.127*
	[1.852]	[1.833]
$\rho_{\varepsilon 2}$		0.253**
		[2.213]

z or *t* statistics in brackets, * *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Second, consumer *i*'s visit decisions may be correlated over time. If the correlation over time is not captured by the covariates we have already controlled for (e.g. consumers' network characteristics, restaurant online rating, monthly time dummies, etc.), then it will be left in the error terms. In other words, ε_{ijt} may be correlated with $\varepsilon_{ij,t-1}$. We incorporate this possibility and re-estimate the model when the correlation between ε_{ijt} and $\varepsilon_{ij,t-k}$ is $(\rho_{\varepsilon 2})^k$ (a parameter to be estimated). The estimation results are presented in column 2 in Table G.1.

Appendix H: Robustness Checks on Instruments

In this appendix, we try sensitivity tests by defining different severe weather instruments. In our main model, we defined a day to be a cold day or a hot day if the minimum temperature on that day is below 0 degree C or the maximum temperature on that day is above 35 degree C, and define a day to be a heavy rain day/a heavy snow day if the precipitation rate is greater than 4 mm per hour/4 cm per hour. In this appendix, we conduct two sensitivity tests. We define a day to be a cold day or a hot day if the minimum temperature on that day is below –2 (or +2) degree C or the maximum temperature on that day is above 33 (or 37) degree C, and define a day to be a heavy rain day/a heavy snow day if the precipitation rate is greater than 3 (or 5) mm per hour/ 3 (or 5) cm per hour. Then, we calculate the number of severe weather days in each time period (month) and construct new weather instruments: $Cold_{t-1}$, Hot_{t-1}, $Rain_{t-1}$, and $Snow_{t-1}$. We find that the estimation results are robust under the two sensitivity tests (see columns 1 and 2 in Table H.1).

Then, we add a robustness check on a new instrumental variable: friends' birthdays. Following Ke and Yang (2016), we use a focal consumer's friends' birthdays as an instrumental variable for her friends' check-ins, and examine how these birthday-induced check-ins by her friends affect her own visit (check-in) decisions. More specifically, our empirical strategy is to instrument for the number of friends' check-ins made at restaurant *j* in period t – 1, $R_{ij,t-1}$, with the number of consumer *i*'s friends who have their birthdays in month t - 1. The basic idea is that it is more likely for people to dine out on or around their birthdays for celebration. Therefore, we expect to see a larger number of friends' check-ins around friends' birthdays. Being totally random, birthdays should be an exogenous source of variation, which can avoid many possible confounds: friends' birthdays in month t – 1 should not directly affect a focal consumer's willingness to visit restaurant j in month t, except through $R_{ij,t-1}$. A high F statistic (27.31) suggests that friends' birthdays are not weak instruments. The estimation results are shown in column 3 in Table H.1. The findings are consistent with our previous results and suggest a significant effect of observational learning.

Table H.1 The Effect of Observational Learning: Instrumental Variables

	(1)	(2)	(3)
	IV: Weather I	IV: Weather II	IV: Birthdays
γ_1	0.723***	0.736***	0.623***
	[4.201]	[4.415]	[4.032]
γ_2	-0.0226**	-0.0227**	-0.0261**
	[2.082]	[2.054]	[2.183]
δ_1	-0.453***	-0.432***	-0.512***
	[3.565]	[3.547]	[3.336]
δ_2	-0.0124*	-0.0121*	-0.0123*
	[1.836]	[1.806]	[1.792]
φ_1	0.00432	0.00423	0.00546
	[0.514]	[0.637]	[0.587]
φ_2	-2.29e-05	-2.17e-05	-2.64e-05
	[0.212]	[0.183]	[0.321]
ρ_1	-0.00174	-0.00166	-0.00416
	[0.328]	[0.445]	[0.892]
ρ_2	-1.69e-05	-1.72e-05	-1.87e-05
	[0.102]	[0.116]	[0.152]
τ	0.233***	0.249***	0.267***
	[3.207]	[3.338]	[3.541]
$\boldsymbol{M_i}$	Yes	Yes	Yes
Monthly dummies	Yes	Yes	Yes

z or *t* statistics in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix I: The Effect of Observational Learning after Multiple Visits

In this appendix, we revise our model in the quality learning stage to incorporate consumer learning after multiple visits as follows:

$$U_{ijt} = \alpha_j + \gamma_1 R_{ij,t-1} + \varphi_1 S_{ij,t-1} + \sum_{k=1}^{K} \beta_{jk} V_{ijk,t-1}$$

$$+ \sum_{k=1}^{K} \delta_k R_{ij,t-1} V_{ijk,t-1} + \sum_{k=1}^{K} \rho_k S_{ij,t-1} V_{ijk,t-1} + \Theta \boldsymbol{M_i} + \tau G_{jt} + Month_t + \varepsilon_{ijt},$$

where the binary variable $V_{ijk,t-1} = 1$, if consumer i's number of self check-ins at restaurant j up until period t – 1 is k, k = 1, 2, …, K; $V_{ijk,t-1} = 0$, otherwise. In this model specification, $V_{ijk,t-1}$ indicates how many times the focal consumer has visited restaurant j. The terms, $\sum_{k=1}^{K} \beta_{jk} V_{ijk,t-1}$, $\sum_{k=1}^{K} \delta_k R_{ij,t-1} V_{ijk,t-1}$, and $\sum_{k=1}^{K} \rho_k S_{ij,t-1} V_{ijk,t-1}$, imply that even after the focal consumer has visited the restaurant K times, she may keep learning restaurant quality. In our estimation, we set K = 3. The estimation results are presented in Table I.1. We find that consumers indeed keep learning after multiple visits, but the marginal learning effect from friends' check-ins decreases with the number of visits. More specifically, the impact of a friend's check-in before having visited a restaurant is $\gamma_1 = 0.829$; the impact of a friend's check-in after having visited a restaurant once is $\gamma_1 + \delta_1 = 0.829 - 0.417 = 0.412$; the impact of a friend's check-in after having visited a restaurant twice is $\gamma_1 + \delta_2 = 0.829 - 0.561 = 0.268$; and the impact of a friend's check-in after having visited a restaurant three times is $\gamma_1 + \delta_3 = 0.829 - 0.589 = 0.240$. Our results suggest that learning becomes insignificant after having visited a restaurant twice.

Table I.1 The Effect of Observational Learning after Multiple Visits

	(1)
	Learning after Multiple Visits
γ_1	0.829***
	[4.332]
δ_1	-0.417***
	[3.245]
δ_2	-0.561***
	[3.924]
δ_3	-0.589***
	[3.893]
φ_1	0.0114
	[0.735]
ρ_1	0.00326
	[0.467]
ρ_2	0.00553
	[0.762]
ρ_3	0.00584
	[0.826]
$\boldsymbol{M_i}$	Yes
Monthly dummies	Yes

z or *t* statistics in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix J: Modeling Correlations Between q_j and p_j

In this appendix, we examine a more flexible structure that allows us to model the correlation between q_j and p_j explicitly. We assume that:

$$q_j = q_{0j} + \varepsilon_{qj}, p_j = p_{0j} + \varepsilon_{pj},$$

where q_{0j} and p_{0j} are constants to estimate, and the random terms ε_{qj} and ε_{pj} follow a bivariate normal with mean zero and covariance matrix $\begin{bmatrix} \sigma_p^2 & \sigma_{pq} \\ \sigma_{pq} & \sigma_q^2 \end{bmatrix}$. In this specification, σ_{pq} captures the possible covariance between q_j and p_j. The detailed estimation results can be found in Table J.1. We find that the covariance σ_{pq} is positive (0.000300), but not statistically significant (p value > 0.10). We can also obtain the correlation between q_j and p_j: $\rho_{pq} = \frac{\sigma_{pq}}{\sigma_p \sigma_q} = \frac{0.000300}{0.068*0.046} = 0.096$.

Table J.1 The Effect of Observational Learning: Correlations Between q_j and p_j

	(1) Main Model
γ_1	0.732***
	[4.032]
γ_2	-0.0203**
	[2.058]
δ_1	-0.618***
	[3.732]
δ_2	-0.0112*
	[1.883]
φ_1	0.00326
	[0.429]
φ_2	-1.27e-05
	[0.127]
ρ_1	-0.00216
	[0.387]
ρ_2	-1.34e-05
	[0.176]
τ	0.239***
	[3.402]
$\boldsymbol{M_i}$	Yes
Monthly dummies	Yes

z or *t* statistics in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix K: Estimation Procedures and Counterfactual Analyses

We estimate our model using the method of maximum likelihood estimation. The log likelihood function is as follows:

$$\ln L(\theta) = \ln \prod_{t=1}^{T} \prod_{j=1}^{J} \prod_{i=1}^{N} \left[\Pr(D_{ijt} = 1) \cdot \Pr(U_{ijt} \geq 0)\right]^{Y_{ijt}} \left[1 - \Pr(D_{ijt} = 1) \cdot \Pr(U_{ijt} \geq 0)\right]^{1-Y_{ijt}}$$

$$= \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{i=1}^{N} \left[Y_{ijt} \ln \frac{\Pr(D_{ijt}=1) \cdot \Pr(U_{ijt} \geq 0)}{1-\Pr(D_{ijt}=1) \cdot \Pr(U_{ijt} \geq 0)} + \ln\left(1 - \Pr(D_{ijt} = 1) \cdot \Pr(U_{ijt} \geq 0)\right)\right],$$

where $\Pr(D_{ijt} = 1) = 1 - (1 - p_j)^{t-1} (1 - q_j)^{\sum_{m=1}^{t-1} R_{ijm}}$, and

$$U_{ijt} = \alpha_j + \gamma_1 R_{ij,t-1} + \gamma_2 R_{ij,t-1}^2 + \varphi_1 S_{ij,t-1} + \varphi_2 S_{ij,t-1}^2 + \beta_j V_{ij,t-1} + \delta_1 R_{ij,t-1} V_{ij,t-1} +$$

$$\delta_2 R_{ij,t-1}^2 V_{ij,t-1} + \rho_1 S_{ij,t-1} V_{ij,t-1} + \rho_2 S_{ij,t-1}^2 V_{ij,t-1} + \Theta \boldsymbol{M_i} + \tau G_{jt} + Month_t + \varepsilon_{ijt}.$$

We choose the parameter values that can maximize the log likelihood function. A major advantage of the structural approach is that it allows for interesting counterfactual analysis that is simply not possible with reduced-form regressions by recovering fundamental structural parameters. If observational learning can provide consumers with useful quality information when they purchase experience goods, we would like to know the value of observational learning and evaluate the effectiveness of different seeding strategies. In order to compute such values, we need to be able to adjust optimal consumer behavior when observational learning is amplified by different seeding strategies.

In our context, restaurants must consider two critical factors that can affect the success of seeding strategies: (1) The initial set of targeted consumers (the portion of targeted consumers *x*); and (2) seeding effort, *e*. Note that for (1), restaurants can collaborate with location-based networks and target well-connected consumers in terms of degree centrality, and for (2), seeding effort is measured by the number of additional check-ins made by consumers belonging to the initial targeted set. If restaurants provide a higher reward in the form of check-in deals (high seeding efforts), they can induce a larger number of consumers' check-ins. For example, a new restaurant might attract more check-ins by the deal "free lunch special on your 4th visit" than the deal "free Jamaica iced tea on your 4th visit."

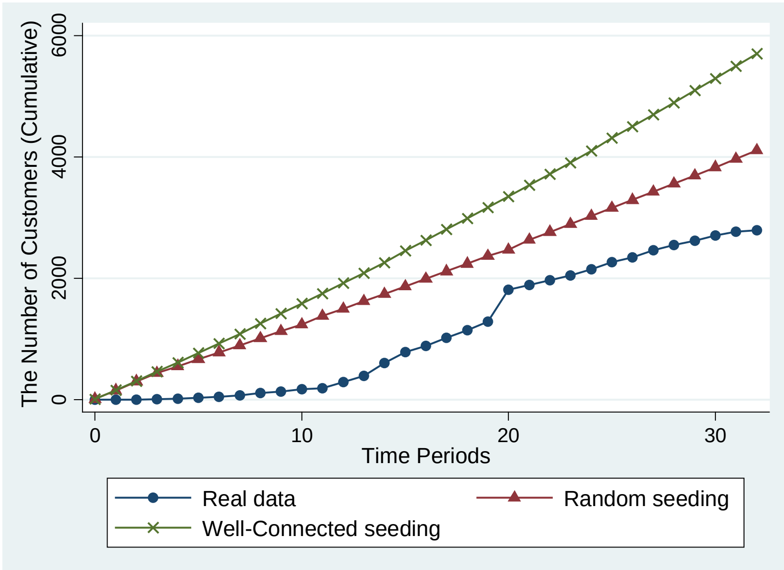

Figure K.1 The Effect of Seeding Strategies

Figure K.1 is an illustrating example and shows the effect of the seeding strategies on the number of customer check-ins of a restaurant. In Figure K.1, we compare two different seeding strategies: the seeding strategy targeting well-connected customers with parameter x = 6% and e = 5 (Strategy 1) and the random seeding strategy (Strategy 2). Strategy 1 means that we target six percent (x = 6%) of well-connected consumers (in terms of degree centrality) and induce each of them to make five (e = 5) more check-ins in period 0. The random seeding strategy means that six percent of consumers are randomly selected as the initial targeted set and each of them is induced to make five additional check-ins in period 0. We find that the seeding strategy 1 can increase the cumulative number of consumer check-ins by more than 100% at the end of our sample period. Figure K.1 also shows that targeting well-connected customers is more effective than the random seeding strategy.

References

1. Cameron, A. C., and Trivedi, P. K. (2009). Microeconometrics Using Stata. College Station, TX: Stata press.
2. Ke, T. T., & Yang, C. Z. (2016). Peer Effect of iPhone Adoptions on Social Networks. MIT Sloan Research Paper No. 5177-16. Available at SSRN: https://ssrn.com/abstract=2837263.
3. Ma, H., Yang, H., Lyu, M. R., & King, I. (2008). Sorec: Social Recommendation Using Probabilistic Matrix Factorization. In Proceedings of the 17th ACM conference on Information and knowledge management (pp. 931-940). ACM.
4. Shi, Z., and Whinston, A. B. (2013). Network Structure and Observational Learning: Evidence from a Location-Based Social Network. *Journal of Management Information Systems*, 30(2), 185-212.