Deriving Value from Social Commerce Networks

Andrew T. Stephen & Olivier Toubia

Web Appendix

MARKETPLACE LEVEL ANALYSIS

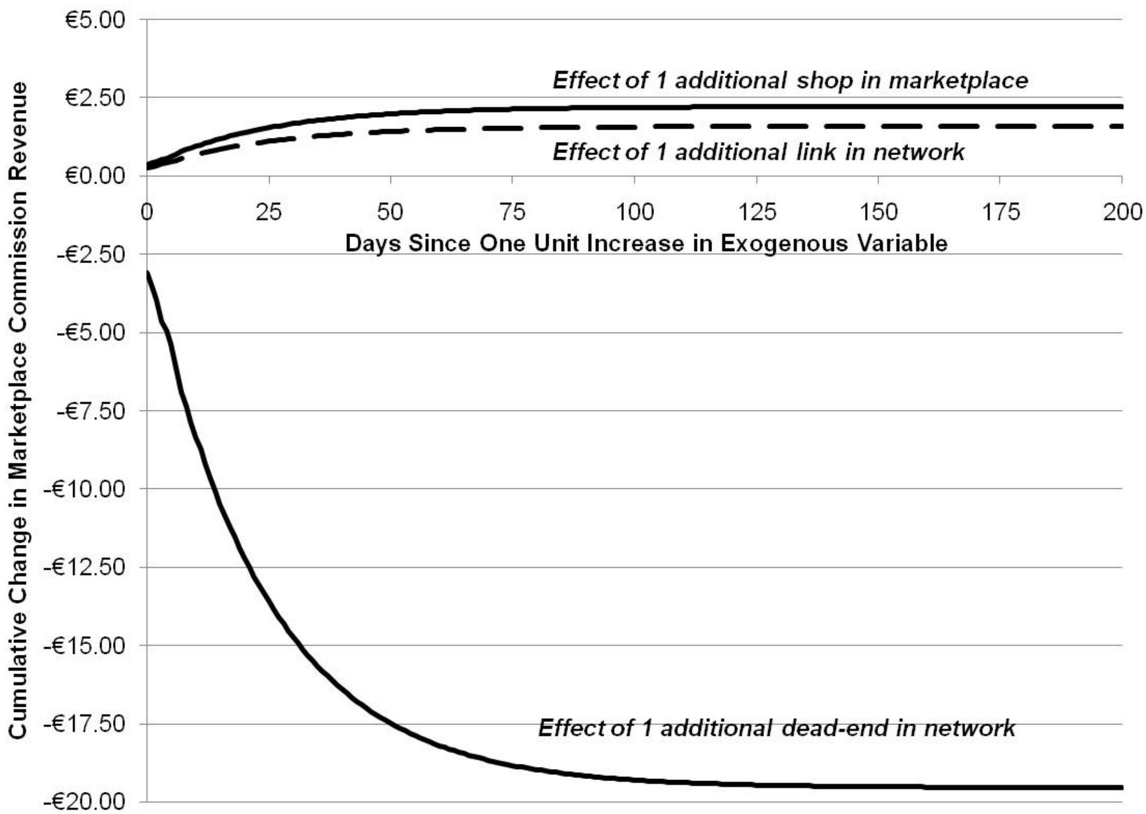
Impulse-Response Functions and Long-Run Impacts

Persistence modeling can be used to examine the long-run effects of the exogenous variables in Equation (2). This approach estimates how a "shock" in an "impulse" variable (e.g., adding a new link) affects a "response" variable over time (e.g., daily commission revenue). This approach has been used in the marketing literature to study the long-run impact of changes in marketing variables on performance variables such as sales (e.g., Dekimpe and Hanssens 2004, 2005) or service adoptions (e.g., Trusov et al. 2008). Specifically, we use impulse-response functions (IRFs). An IRF simulates the impact over time of a change in one variable (in our case one of the exogenous variables) on the full dynamic system (Bronnenberg, Mahajan, and Vanhonacker 2000). Readers are referred to Dekimpe and Hanssens (2004, 2005) for a more detailed explanation of impulse-response functions, and Pauwels (2004) for an example of their use in modeling the impact of marketing actions or changes in marketing variables on dynamic systems. More typical uses of IRFs in marketing (e.g., stock return models) attempt to quantify long-run effects of marketing mix variables that last for months or even years. In our case, given the rapid growth of this marketplace and its social network, our time frame is much shorter.

The cumulative IRFs (sometimes referred to as total short-run effects) based on the estimated coefficients from the ARX model in (2) are plotted in Figure A1. These plots show

how much commission revenue is generated by a punctual increase of one unit in each of the exogenous variables and how this "shock" impacts the system over time (the effect propagates through the lags of the dependent variable).

FIGURE A1
CUMULATIVE IMPULSE-RESPONSE FUNCTIONS



SHOP LEVEL ANALYSIS

Network Measures Used in the Analysis

The following network measures are used, as defined in Table 1 in the paper:

- *Indegree centrality*: number of incoming links received by a shop from other shops;
- *Outdegree centrality*: number of outgoing links from a shop to other shops;

- *Incoming proximity* (Faust and Wasserman 1992; de Nooy et al. 2005): this measures the reachability of shop *i* from other shops in the network. Incoming proximity is proportional to the proportion of shops in the network other than *i* that can reach *i* in a finite number of steps (shop *i*'s "indomain"), and inversely proportional to the mean geodesic distance (shortest path length) from these shops to *i*. Thus, incoming proximity is highest for shops that are accessible from a large number of shops in the network in only few steps. Note that this is a directed graph analog of Freeman's (1979) standard "closeness" metric for undirected graphs;
- *Outgoing proximity* (Faust and Wasserman 1992; de Nooy et al. 2005): this measures the reachability of other shops in the network from shop *i*. Outgoing proximity is proportional to the proportion of shops in the network other than *i* that can be reached from *i* in a finite number of steps (shop *i*'s "outdomain"), and inversely proportional to the mean geodesic distance (shortest path length) from shop *i* to these shops. Thus, outcoming proximity is highest for shops from which a large number of shops in the network are accessible in only few steps.
- *Incoming clustering coefficient* (Watts and Strogatz 1998; Zhou 2002): the incoming clustering coefficient of shop *i* is the degree of interconnectedness among the shops that link to shop *i*. Specifically, it is the proportion of possible links that exist among the shops in shop *i*'s incoming ego-network. The higher a shop's incoming clustering coefficient, the more densely interconnected its incoming ego-network is, i.e., that shop is connected to by other shops that are themselves highly interconnected (as opposed to being connected to by shops that are more dispersed throughout the network). Note that

clustering is not a centrality measure, but rather a measure of how dense each shop's ego-network is;

- *Outgoing clustering coefficient* (Watts and Strogatz 1998; Zhou 2002): the outgoing clustering coefficient of shop *i* is the degree of interconnectedness among the shops that shop *i* links to, and thus is the proportion of possible links that exist among the shops in shop *i*'s outgoing ego-network. The higher a shop's outgoing clustering coefficient, the more densely interconnected its outgoing ego-network is, i.e., that shop connects to other shops that are themselves highly interconnected;
- *Hub centrality* and *authority centrality* (Kleinberg 1999): these measures of centrality are directed graph analogs of eigenvector centrality (Bonacich 1987) for outgoing links (hub) and incoming links (authority). The basic concept of eigenvector centrality is that a shop is more prominent in the network if it is well-connected to other well-connected shops. A shop with a high hub score links to many shops with high authority scores, and a shop with a high authority score is linked to by many shops with high hub scores (these metrics are based on eigenvector decompositions of the network's adjacency matrix; see Kleinberg 1999 for derivations). Because these centrality measures are not directly related to accessibility, we do not expect them to have a significant impact on shop performance.

Nested Model Comparisons

We first compared the full model in (3) to (6) in the paper to a nested model without network position effects; i.e., in (5) restricting $\boldsymbol{\beta}_3 = \boldsymbol{\beta}_5 = \boldsymbol{\beta}_6 = 0$. Following Newton and Raftery (1994) and Rossi, Allenby, and McCulloch (2005) we computed the log marginal densities for the two models using the harmonic means of the respective models' likelihoods across posterior draws (every tenth draw after burn-in, for computational reasons). The full model had a better fit ($-2\log\hat{p}(y \mid M_{product-only}) = 2.334 \times 10^{-5}$ versus $-2\log\hat{p}(y \mid M_{full}) = 2.308 \times 10^{-5}$; smaller is better), and a large log Bayes factor based on these log marginal densities ($\log BF_{full\,vs\,product-only}$ = 1289; note that $\log BF_{AvsB} > 5$ is "strong evidence" in favoring model A over model B), which provided very strong evidence in favor of the full model over the product-only restricted model. Note that because some of our priors were improper, special care should be taken when computing marginal densities and Bayes factors (Rossi and Allenby 2003). As a prior sensitivity analysis we used diffuse but proper priors and obtained very similar results (details available from the authors).

The mean correlation between actual and predicted commissions (over MCMC draws) for this restricted model was clearly inferior to the full model (.07 versus .22).

We then performed another nested model comparison, this time comparing the full model in (3) to (6) to a restricted model with $\boldsymbol{\alpha} = \mathbf{0}$, $\boldsymbol{\gamma} = \mathbf{0}$ and $\beta_1 = 0$, which was a simpler Tobit model without the latent ability variable and entering the network position and product assortment variables directly as regressors instead of their ability-adjusted residuals. This simpler model's fit and log Bayes factor were worse ($-2\log\hat{p}(y \mid M_{simple}) = 2.311 \times 10^{-5}$ versus $-2\log\hat{p}(y \mid M_{full})$ = 2.308×10^{-5}; $\log BF_{full\,vs\,simple} = 139$), and the correlation between actual and predicted commissions was also worse than for the full model (.15 versus .22). We therefore base our findings on the full model with the latent "ability" specification. Parameter estimates for the simpler Tobit model are reported in Table A1 for sake of comparison and as a robustness check. All effects that are significant in the full model are confirmed by this simpler specification, and the magnitudes of the effects are very comparable. A few effects, such as that of outgoing clustering coefficient and authority, are not significant in the full model but are significant in the simpler model. Since the former fits better than the latter we do not focus heavily on these points of difference. Indeed, basing our substantive findings on the full model with fewer significant effects is conservative.

TABLE A1

SIMPLER TOBIT MODEL: EFFECTS OF NETWORK POSITION AND PRODUCT ASSORTMENT ON SHOP-LEVEL COMMISSION REVENUES

Parameter	**Posterior Mean**	**Posterior Standard Error**	**95% Credible Interval**
Intercept (β_0)	.189***	.0001	.17, .21
Latent ability (β_1)	0[a]	n/a	n/a
Shop age (β_2)	.018	.0002	-.004, .04
Network effects			
Indegree centrality ($\beta_{3,1}$)	.914***	.0004	.84, .98
Outdegree centrality ($\beta_{3,2}$)	-.498***	.0005	-.59, -.41
Incoming clustering coefficient ($\beta_{3,3}$)	-.152***	.0002	-.19, -.12
Outgoing clustering coefficient ($\beta_{3,4}$)	.050***	.0002	.02, .09
Authority (in eigenvector centrality) ($\beta_{3,5}$)	-.084***	.0003	-.15, -.02
Hub (out eigenvector centrality) ($\beta_{3,6}$)	.038	.0005	-.02, .10
Inproximity (incloseness) centrality ($\beta_{3,7}$)	.178***	.0002	.14, .22
Outproximity (outcloseness) centrality ($\beta_{3,8}$)	-.071***	.0002	-.11, -.03
Product assortment effects			
Number of products ($\beta_{4,1}$)	.013	.0001	-.01, .04
Average product popularity ($\beta_{4,2}$)	-.015	.0002	-.04, .01
Quadratic and significant interaction effect[b]			
Indegree^2 ($\beta_{5,1}$)	-.017***	.00001	-.02, -.01
Outdegree^2 ($\beta_{5,2}$)	.007***	.00002	.003, .01
Authority × number of products ($\beta_{6,5}$)	.100**	.0004	.02, .18
Inproximity × number of products ($\beta_{6,7}$)	-.036**	.0002	-.07, -.01
Outproximity × number of products ($\beta_{6,8}$)	.021*	.0001	-.004, .05

* The 90% credible interval does not contain zero (two-sided).
** The 95% credible interval does not contain zero (two-sided).
*** The 99% credible interval does not contain zero (two-sided).
[a] Fixed to zero.
[b] None of the other interaction effects are close to being significantly different from zero.

Notes: (1) These estimates are based on a non-hierarchical Tobit model without a latent ability variable. Estimation involved 20,000 MCMC draws (10,000 burn-in). The model mixed well and converged quickly. (2) The error standard deviation (σ) has posterior mean = 3.260, standard error = .0001, and 95% credible interval (3.24, 3.28). All network position- and product assortment-related variables were standardized (mean = 0, standard deviation = 1) before running the model. Thus, posterior means in this table can be compared as "standardized" coefficients.

Out-of-Sample Fit

Out-of-Sample Fit

The common approach of randomly splitting the data into an estimation sample and a hold-out validation sample is inappropriate here since our latent ability variable needs to be estimated for each shop, and predicting shop performance in a validation sample would require estimating latent ability for each shop in that sample, which would require using the validation data for estimation. Instead, we re-estimated the model with the network and product assortment variables measured at the end of month 5 (instead of the end of month 6) and the commission revenues measured during month 6 (instead of during month 7; i.e., “month 5/6” data instead of “month 6/7” data). As an indicator of robustness we found no qualitative differences between the results reported below from analysis of the month 6/7 data and the results from this month 5/6 data (details available from the authors). We then assessed out-of-sample fit by using these month 5/6 parameter estimates (including shop-level latent ability estimates) to predict the commission revenues earned during month 7 as a function of the network and assortment variables at the end of month 6 (i.e., using the month 5/6 data for calibration and the month 6/7 data for validation). The average predicted month 7 commission (mean taken across MCMC draws of the parameter estimates) was €0.1947, close to the actual mean month 7 commission of €0.1783 (and the actual mean lied well within the distribution of the predicted means from the MCMC draws). The MAD across draws was €0.077, which was slightly larger than for in-sample fit but still reasonable, and the mean correlation between actual and predicted commissions was .21.

Bayesian Estimation Procedure

Priors

- $Ability_i \sim N(-1, \eta^2)$.
- Diffuse on β_0, β_1, β_2, $\boldsymbol{\beta}_3$, $\boldsymbol{\beta}_4$, $\boldsymbol{\beta}_5$, $\boldsymbol{\beta}_6$, $\boldsymbol{\alpha}_0$, $\boldsymbol{\alpha}_1$, $\boldsymbol{\gamma}_0$ and $\boldsymbol{\gamma}_1$.
- $\sigma^2 \sim InverseGamma\left(\frac{r_0}{2}, \frac{s_0}{2}\right)$, where $r_0 = s_0 = 1$.
- $\eta^2 \sim InverseGamma\left(\frac{r_0}{2}, \frac{s_0}{2}\right)$, where $r_0 = s_0 = 1$.
- $\boldsymbol{\Lambda} \sim InverseWishart(n_0, n_0\boldsymbol{\Delta_0})$, where $n_0 = p + q + 3$, and $\boldsymbol{\Delta_0} = \mathbf{I}$, with p the number of network-related variables (7) and q the number of assortment-related variables (3).

Markov Chain Monte Carlo Simulation Steps

Step 1

$$L(Ability_i \mid Performance_i, Performance_i^*, \boldsymbol{\beta}, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \sigma^2, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1) \sim Normal(m_i, V_i) \text{, where}$$

$$m_i = V_i \cdot \left(\frac{Performance_i^* - \beta_0 - \beta_2 Age_i - \sum_{j=1}^{J} \beta_{3,j}\delta_j - \sum_{k=1}^{K} \beta_{4,k}\zeta_k - \sum_{l=1}^{L} \beta_{5,l}\delta_{i,j}^2 - \sum_{j',k'} \beta_{6,j'k'}\delta_{i,j'}\zeta_{i,k'}}{\beta_1} \cdot \frac{\beta_1^2}{\sigma^2} + [\boldsymbol{\alpha}_1, \boldsymbol{\gamma}_1]\boldsymbol{\Lambda}^{-1}[(Network_i - \boldsymbol{\gamma}_0)'; (Assortment_i - \boldsymbol{\alpha}_0)'] + \frac{1}{\eta^2} \right)$$

$$V_i = \left(\frac{\beta_1^2}{\sigma^2} + [\boldsymbol{\alpha}_1, \boldsymbol{\gamma}_1]\boldsymbol{\Lambda}^{-1}[\boldsymbol{\alpha}_1, \boldsymbol{\gamma}_1]' + \frac{1}{\eta^2} \right)^{-1}$$

Step 2

$$L(Performance_i^* \mid Performance_i, Ability_i, \boldsymbol{\beta}, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \sigma^2, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1) \sim$$
$$TruncatedNormal(\beta_0 + \beta_1 Ability_i + \boldsymbol{\beta}_2\boldsymbol{\delta}_i + \boldsymbol{\beta}_3\boldsymbol{\zeta}_i, \sigma^2)$$

if $Performance_i < 0$. Otherwise $Performance_i^* = Performance_i$.

Step 3

$$L(\eta^2 \mid Performance_i, Performance_i^*, Ability_i, \boldsymbol{\beta}, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \sigma^2, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1) \sim$$
$$InverseGamma\left(\frac{r_0 + \sum_{i=1}^{N}(Ability_i + 1)^2}{2}, \frac{s_0 + N}{2}\right)$$

Step 4

$$L(\sigma^2 \mid Performance_i, Performance_i^*, Ability_i, \boldsymbol{\beta}, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \sigma^2, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1) \sim$$
$$InverseGamma\left(\frac{r_0 + \sum_{i=1}^{N}(Performance_i^* - (\beta_0 + \beta_1 Ability_i + \beta_2 Age_i + \boldsymbol{\beta}_3\boldsymbol{\delta}_i + \boldsymbol{\beta}_4\boldsymbol{\zeta}_i + \boldsymbol{\beta}_5\boldsymbol{\delta}_i^2 + \boldsymbol{\beta}_6\boldsymbol{\delta}_i\boldsymbol{\zeta}_i))^2}{2}, \frac{s_0 + N}{2}\right)$$

Step 5

$$L(\boldsymbol{\Lambda} \mid Performance_i, Performance_i^*, Ability_i, \boldsymbol{\beta}, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \sigma^2, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1) \sim$$
$$InverseWishart(n_0 + N, n_0\boldsymbol{\Delta}_0 + \sum_{i=1}^{N}[\boldsymbol{\delta}_i'; \boldsymbol{\zeta}_i']'[\boldsymbol{\delta}_i, \boldsymbol{\zeta}_i])$$

Step 6

$$L([\beta_0; \beta_1; \beta_2; \boldsymbol{\beta}_3; \boldsymbol{\beta}_4; \boldsymbol{\beta}_5; \boldsymbol{\beta}_6] \mid Performance_i, Performance_i^*, Ability_i, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \sigma^2, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1) \sim$$
$$Normal((\mathbf{X'X})^{-1}\mathbf{X'y}^*, \sigma^2(\mathbf{X'X})^{-1})$$

where $\mathbf{X}_i = [1, Ability_i, Age_i, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \boldsymbol{\delta}_i^2, \boldsymbol{\delta}_i\boldsymbol{\zeta}_i]$

Step 7

For all j:

$$L([\gamma_{0,j}; \gamma_{1,j}] \mid Performance_i, Performance_i^*, Ability_i, \boldsymbol{\beta}, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \sigma^2, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \gamma_{0,(l \neq j)}, \gamma_{1,(l \neq j)}) \sim$$
$$Normal\left((\mathbf{X'X})^{-1}\mathbf{X'W}, \boldsymbol{\Lambda}_{j,j}(\mathbf{X'X})^{-1}\right)$$

where $\mathbf{X}_i = [1, Ability_i]$ and $\boldsymbol{W}_i = [\boldsymbol{Network}_{i,j}]$, and similarly for all k:

$$L([\alpha_{0,k}; \alpha_{1,k}] \mid Performance_i, Performance_i^* Ability_i, \boldsymbol{\beta}, \boldsymbol{\delta}_i, \boldsymbol{\zeta}_i, \sigma^2, \alpha_{0,(l \neq k)}, \alpha_{1,(l \neq k)}, \gamma_0, \gamma_1) \sim$$
$$Normal\left((\mathbf{X'X})^{-1}\mathbf{X'W}, \boldsymbol{\Lambda}_{J+k,J+k}(\mathbf{X'X})^{-1}\right)$$

Note that we did not update $[\gamma_{0j}; \gamma_{1j}]$ and $[\alpha_{0j}; \alpha_{1j}]$ for all j's and k's simultaneously for tractability reasons.

REFERENCES

Bonacich, Phillip (1987), "Power and Centrality: A Family of Measures," *American Journal of Sociology*, 92 (5), 1170-1182.

Bronnenberg, Bart J. and Wilfred Vanhonacker (2000), "The Emergence of Market Structure in New Repeat-Purchase Categories: A Dynamic Approach and an Empirical Application," *Journal of Marketing Research*, 37 (1), 16-31.

Dekimpe, Marnik G. and Dominque M. Hanssens (2004), "Persistence Modeling for Assessing Marketing Strategy Performance." In Christine Moorman and Donald R. Lehmann (eds), *Assessing Marketing Strategy Performance*, Cambridge, MA: Marketing Science Institute.

——— and ——— (2005), "The Persistence of Marketing Effects on Sales," *Marketing Science*, 14 (1), 1-21.

Faust, Katherine and Stanley Wasserman (1992), "Centrality and Prestige: A Review and Synthesis," *Journal of Quantitative Anthropology*, 4 (1), 23-78.

Freeman, Linton C. (1979), "Centrality in Social Networks: Conceptual Clarification," *Social Networks*, 1, 215-239.

Kleinberg, Jon M. (1999), "Authoritative Sources in a Hyperlinked Environment," *Journal of the Association for Computing Machinery*, 46 (5), 604-632.

Newton, Michael A. and Adrian E. Raftery (1994), "Approximate Bayesian Inference with the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society Series B*, 56 (1), 3-48.

de Nooy, Wouter, Andrej Mrvar, and Vladimir Batagelj (2005), *Exploratory Social Network Analysis with Pajek*, Cambridge, UK: Cambridge University Press.

Pauwels, Koen (2004), "How Dynamic Consumer Response, Competitor Response, Company Support, and Company Inertia Shape Long-Term Marketing Effectiveness," *Marketing Science*, 23 (4), 596-610.

Rossi, Peter E., and Greg M. Allenby (2003), "Bayesian Statistics and Marketing," *Marketing Science*, 22 (3), 304-328.

Rossi, Peter E., Greg M. Allenby, and Rob McCulloch (2005), *Bayesian Statistics and Marketing*, New York, NY: Wiley.

Watts, Duncan J. and Steven H. Strogatz (1998), "Collective Dynamics of Small-World Networks," *Nature*, 393 (6684), 440-442.

Zhou, Haijun (2002), "Scaling Exponents and Clustering Coefficients of a Growing Random Network," *Physical Review E*, 016125.