Appendix A. Alternative Similarity and Common Tie Measurement

Alternative measures for content similarity between initiating providers and responding providers are calculated as follows.

$$Dice\ Similarity_{ijt} = \frac{2\sum_{k=1}^{16}(Category_{it}^{k} * Category_{jt}^{k})}{\sum_{k=1}^{16}(Category_{it}^{k})^2 + \sum_{k=1}^{16}(Category_{jt}^{k})^2}$$

and

$$Jaccard\ Similarity_{ijt} = \frac{\sum_{k=1}^{16}(Category_{it}^{k} * Category_{jt}^{k})}{\sum_{k=1}^{16}(Category_{it}^{k})^2 + \sum_{k=1}^{16}(Category_{jt}^{k})^2 - \sum_{k=1}^{16}(Category_{it}^{k} * Category_{jt}^{k})}$$

When analyzing the similarity between two providers' content categories, it is necessary to control for uncertainty in the observed categories of their videos. Following the concept of information entropy, which measures the expected amount of information revealed (Shannon and Weaver 1949), we measure the entropy of a provider's video categories as follows. The entropy of a provider's video categories increases as the provider produces videos in more categories.

$$Entropy_{it} = \sum_{k=1}^{16}(Category_{it}^{k} * Log\frac{1}{Category_{it}^{k}})$$

As for the measurement of topological overlap, according to Miritello (2013), topological overlap is also calculated based on common ties between the two nodes to reflect their tie strength. Similar to common ties, it measures the extent to which the two nodes share common connections to other nodes in the network. For a nondirectional network, Ravasz et al. (2002) suggest calculating topological overlap as

$$TOverlap_{ij} = \frac{n_{ji}}{min(k_i, k_j)}$$

where k_i and k_j are the degrees of the two nodes, and n_{ji} is the number of neighbors common to both of them. In other words, topological overlap measures the ratio of common ties to all the ties of the node with fewer connections. Therefore, ties between individuals who have many common friends (large overlap), are stronger than ties between people who have few common friends (small overlap) (Miritello 2013). A topological overlap of almost 1 between two nodes implies that they are connected to the same nodes, whereas a 0 value indicates that two nodes do not share any connections in common.

In our context, with indirect path ties as the measure of common friends, the topological overlap between i and j at time t can be calculated as follows:

$$TOverlap_{ijt} = \frac{IndirectPath_{ijt}}{min(k_{it}, k_{jt})}$$

where $IndirectPath_{ijt}$ is the number of indirect paths from j to i at time t, k_{it} is the number of i's incoming ties and k_{jt} is the number of j's outgoing ties at time t.

Appendix B. Correlation among Key Variables

Table B presents the correlations among the key variables. According to this table, $CommonTie_{ijt}$ is negatively correlated with $Similarity_{ijt}$, suggesting that the provider pairs with more common ties to other providers tend to have less similar content. Each of the two independent variables is correlated positively with the dependent variables. Given that all pairwise correlations are statistically significant, it is essential to include all the independent variables of interest in the model estimation simultaneously, as their coefficients cannot be properly interpreted otherwise.

Table B. Correlations among Key Variables

	$CommonTie_{ijt}$	$Similarity_{ijt}$	$Log\Delta Sub_{it}$	$Log\Delta View_{it}$
$CommonTie_{ijt}$	1			
$Similarity_{ijt}$	-0.096	1		
$Log\Delta Sub_{it}$	0.112	0.011	1	
$Log\Delta View_{it}$	0.082	0.022	0.569	1

Note. All pairwise correlations are significant at $p<0.001$

Appendix C. Preliminary Analyses on Reciprocation Probability

Figure C.1 plots the total number of reciprocations over time, and Figure C.2 plots the probability of reciprocation over time using the hazard function. Duration in both graphs measures days between when the initiator features a responding provider and the responding provider reciprocates. Both show that reciprocation probability decays over time. Most initiating providers received reciprocation from the responding providers within three days after they initiated the tie. Beyond two weeks after the initiation, reciprocation rarely occurs.

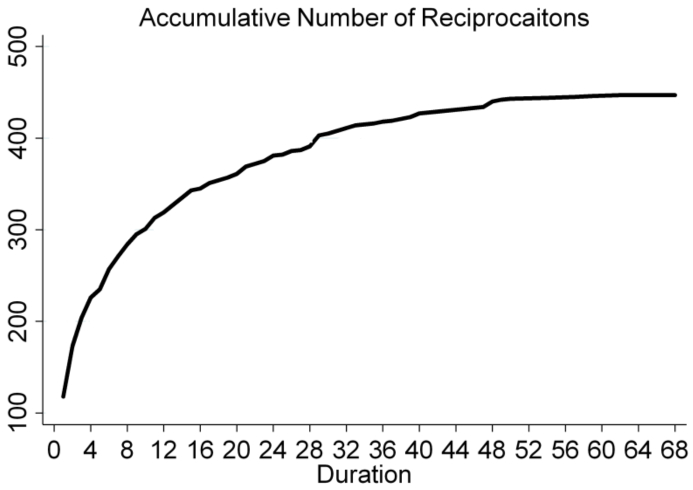

Figure C.1. Number of Reciprocations over Time

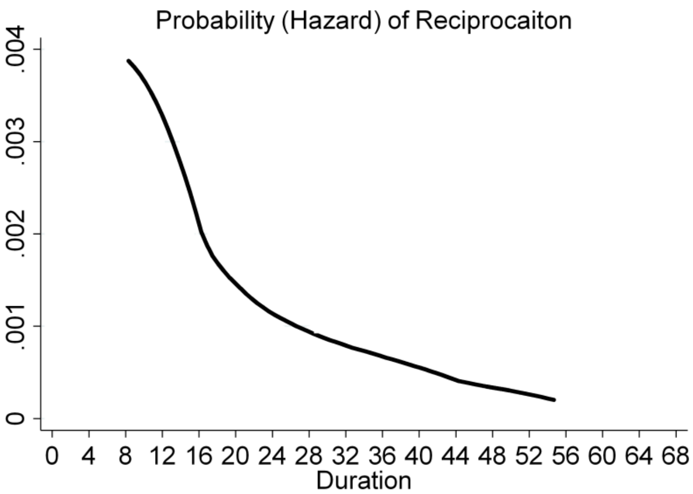

Figure C.2. Probability of Reciprocation over Time

Appendix D. Reciprocation Impact on the Overall Network

At the network level, we conduct a preliminary analysis to examine the impact of tie reciprocation on the status growth of the overall network using the following autoregressive with exogenous variables (ARX) time-series model:

$$\Delta NetworkSub_t = \beta_0 + \sum_{l=1}^{p} \emptyset_l * \Delta NetworkSub_{t-l} + \sum_{j=0}^{r} \gamma_j * \Delta NetworkReciprocations_{t-j} + \varepsilon_t \quad \text{(D.1)}$$

In this model, $\Delta NetworkSub_t$ is the increase in the log total number of subscribers for all the providers in our sample network for period t. $\Delta NetworkReciprocations_t$ is the number of ties reciprocated within the network for period t. The AR lags of the variables ($\Delta NetworkSub_{t-l}$ and $\Delta NetworkReciprocations_{t-j}$) are included to account for their histories (carryover effects) (Stephen and Toubia 2010, Luo et al. 2013). If tie reciprocation adds value to the network by making providers' channels more accessible, we should expect positive coefficients of $\Delta NetworkReciprocations_{t-j}$.

We first test the exogeneity of $\Delta NetworkReciprocations$ using a Granger causality test. The results show that $\Delta NetworkReciprocations$ is not Granger-caused by $\Delta NetworkSub$ (X^2=17.84, p =0.27) and that our ARX specification in Equation (D.1) is valid (Stephen and Toubia 2010). The best-fitting model (i.e., with the lowest Bayesian information criterion) is when p =1 and r = 0. The unit-root tests show that all the variables are stationary, and the estimated coefficient of $\Delta NetworkReciprocations_t$ is 0.357 (with $p<0.05$), meaning that tie reciprocations within the network have a significantly positive effect on the aggregate status growth of the providers within the network. In other words, more connected networks can improve the overall accessibility of their providers' content.

Appendix E. Model Assumptions and Data Cleaning

(1) Model Assumptions

Model E.1 (i.e., model (1) in the main body) is the baseline model that examines the impact of reciprocation on the initiating provider,

$$Log\Delta Sub_{it,FB} = W_{ijt}\beta_{FB} + Z_{it}\varphi + l_t + c_{ij,FB} + u_{ijt,FB} \quad \text{(E.1)}$$

Model E.2 (i.e., model (3) in the main body) is the baseline model that examines the initiating provider's reciprocation decision,

$$FB_{ijt} = 1\left[W'_{ijt}\gamma + \overline{W'}_{ij}\omega + k_t + \epsilon_{ijt} > 0\right] \quad \text{(E.2)}$$

$\overline{W'}_{ij} = T^{-1}\sum_{t=1}^{T} W'_{ijt}$ is the time averages (averaged across the panel) of the covariates (excluding the constant) added as additional explanatory variables to account for heterogeneity and potential correlation between W'_{ijt} and the original error term ε_{ijt} (Mundlak 1978, Murtazashvili and Wooldridge 2016).

We start by rewriting model (E.1) as follows:

$$Log\Delta Sub_{it} = W_{ijt}\beta_0 + FB_{ijt}W_{ijt}\delta_1 + Z_{it}\varphi + l_t + e_{ijt0} + FB_{ijt}v_{ijt1} \quad \text{(E.3)}$$

where $\delta_1 = \beta_1 - \beta_0$, $e_{ijt0} = c_{ij0} + u_{ijt0}$, $e_{ijt1} = c_{ij1} + u_{ijt1}$, and $v_{ijt1} = e_{ijt1} - e_{ijt0}$. β_0 captures the same baseline influence of the covariates on $Log\Delta Sub_{it}$ for both reciprocated and not-yet-reciprocated providers when controlling for various influences such as content quality, provider recognition, and genre popularity in the general diffusion process. δ_1 indicates the additional influence of the covariates on $Log\Delta Sub_{it}$ for reciprocated providers only, thus capturing the moderating effects of the key variables on reciprocation impact.

For identification purposes, we make the following three assumptions.

Assumption A: The unobserved heterogeneity in model (E.3) is linearly related to $\overline{W}_{ij}$ such that

$$e_{ijt0} = \overline{W}_{ij}\rho_0 + \sigma_{ijt0}, \quad v_{ijt1} = \overline{W}_{ij}\rho_1 + \sigma_{ijt1} \quad \text{(E.4)}$$

where $(\sigma_{ijt0}, \sigma_{ijt1})$ is assumed to be independent of W_{ij}.

Assumption B: The error term (ϵ_{ijt}) in model (E.2) is independent of W'_{ijt} and follows the standard normal distribution $\epsilon_{ijt} \sim Normal(0,1)$.

Assumption C: Error terms of two stages are assumed to be correlated, and for simplicity, we make linearity assumptions: $E(\sigma_{ijt0}|\epsilon_{ijt}) = \xi_0\epsilon_{ijt}$ and $E(\sigma_{ijt1}|\epsilon_{ijt}) = \xi_1\epsilon_{ijt}$.

Under Assumption A, E.3 can be written as

$$Log\Delta Sub_{it} = W_{ijt}\beta_0 + FB_{ijt}W_{ijt}\delta_1 + Z_{it}\varphi + l_t + \overline{W}_{ij}\rho_0 + FB_{ijt}\overline{W}_{ij}\rho_1 + \sigma_{ijt0} + FB_{ijt}\sigma_{ijt1} \quad \text{(E.5)}$$

Under Assumption B, we have

$$E(\epsilon_{ijt}|FB_{ijt}, W'_{ijt}) = \widehat{gr}_{ijt} = h(FB_{ijt}, W'_{ijt}\gamma + \overline{W}'_{ij}\varphi + k_t)$$
$$= FB_{ijt}\lambda(W'_{ijt}\gamma + \overline{W'}_{ij}\varphi + k_t) - (1 - FB_{ijt})\lambda(-W'_{ijt}\gamma - \overline{W'}_{ij}\varphi - k_t) \quad \text{(E.6)}$$

where $h(\cdot)$ is the generalized error function and $\lambda(\cdot) \equiv \phi(\cdot)/\Phi(\cdot)$ is the inverse Mills ratio

(Murtazashvili and Wooldridge 2016, Vella and Verbeek 1999). With Assumption C and by iterated expectations, $E(\sigma_{ijt0} + FB_{ijt}\sigma_{ijt1} | FB_{ijt}, W_{ij}) = \xi_0 h_{ijt} + \xi_1 FB_{ijt} h_{ijt}$. Finally, we can derive the following estimation equation for model (E.3):

$$Log\Delta Sub_{it} = W_{ijt}\beta_0 + FB_{ijt}W_{ijt}\delta_1 + \bar{W}_{ij}\rho_0 + FB_{ijt}\bar{W}_{ij}\rho_1 + \xi_0 h_{ijt} + \xi_1 FB_{ijt}h_{ijt} + Z_{it}\varphi + l_t + \tau_{ijt} \quad \text{(E.7)}$$

Therefore, under these assumptions, the two-stage estimation procedure can generate consistent and asymptotically normal estimators of $\beta_0, \delta_1, \gamma, \rho_0, \rho_1, \xi_0$, and ξ_1 (Murtazashvili and Wooldridge 2016).

(2) Data Cleaning

In addition to making the aforementioned assumptions, we undertake several data cleaning measures to ensure the identification of our model. First, we deleted provider pairs with reciprocation occurring on the same day as the tie initiation because we were unable to identify the initiating provider versus the responding provider for these pairs[1]. Second, to focus on tie reciprocation instead of initiation, we only used the observations on a provider pair after tie initiation, that is, after the initiating provider sent the outgoing tie to the responding provider. Third, some initiating providers deleted the outgoing ties to the responding providers after initiation by removing the responding provider from the featured provider's list[2]. For these pairs, we used the observations between tie initiation and the initiator's deletion of the tie for the estimation. Meanwhile, we observe that some responding providers who reciprocated previously also withdrew their reciprocation by later removing their outgoing ties to the initiators[3]. Therefore, we assume that the responding provider's reciprocation decision is not made on a one-time-only basis but on a per-period basis. The panel setting is consistently applied to both reciprocation probability and impact in our empirical model. Finally, our datasets include the reciprocated provider pairs with the same initiating provider. To differentiate the impacts of different responding providers on the same initiator, we add the outdegree and indegree of both providers, the responding provider's subscribers and views, and the subscribers and views of the initiator's indegrees and outdegrees as control variables in the estimation of reciprocation impact.

[1] Two hundred and sixty reciprocations happened on the same day, which accounted for 5.46% of our data sample.

[2] In the remaining data sample, we observed that 845 (18.9%) initiating providers deleted the outgoing ties to the responding providers after initiation, of which 825 (18.4%) pairs were one-way relationships and 25 (0.56%) pairs were mutual ties.

[3] Eleven (0.25%) cases exist in our dataset.

Appendix F. Robustness Check for Multiple Reciprocations of the Same Initiators

Table F. Results of Robustness Check for Multiple Reciprocations of the Same Initiators

Stage 1: Reciprocation Probability	
	(1)
Dependent Variable / Independent Variable	FB_{ijt}
$CommonTie_{ijt}$	0.498 (0.042)***
$Similarity_{ijt}$	0.835 (0.363)*
$CommonTie_{ijt} * Similarity_{ijt}$	-0.532 (0.059)***
Other control variables	Y
Time-specific effects	Y
Control for category entropy of i and j	Y
Log likelihood	-5365
Stage 2: Reciprocation Impact	
Dependent Variable / Independent Variable	$Log\Delta Sub_{it}$
FB_{ijt}	4.379 (0.761)***
$FB_{ijt} * CommonTie_{ijt}$	-0.526 (0.071)***
$FB_{ijt} * Similarity_{ijt}$	-4.049 (1.525)**
$FB_{ijt} * CommonTie_{ijt} * Similarity_{ijt}$	0.500 (0.167)**
Other control variables	Y
Time-specific effects	Y
Control for category entropy of i and j	Y
Control for subscribers/views of i's outdegrees/indegrees	Y
R^2	71.0%
Observations	102,316
Provider pairs	4,369

Note. The estimations use the same model specification as in the main result estimation in Column 4 of Table 3. For brevity, only the coefficients of the key variables for our hypotheses are shown. ***p<0.001, **p<0.01, *p<0.05.

Appendix G. Robustness Checks for Correlated Provider Pairs

Table G presents the results of robustness checks for correlated provider pairs, where the standard errors are clustered by the initiating provider (Column (1)), the responding provider (Column (2)), and the responding provider in stage 1 and the initiating provider in stage 2 (Column (3)), respectively. All the results are consistent with our main results.

Table G. Robustness Check Results for Correlated Provider Pairs

Stage 1: Reciprocation Probability			
	(1) Cluster by i	(2) Cluster by j	(3) Cluster by j in stage 1 Cluster by i in stage 2
Dependent Variable / Independent Variable	FB_{ijt}		
$CommonTie_{ijt}$	0.290 (0.093)**	0.290 (0.099)**	0.290 (0.099)**
$Similarity_{ijt}$	1.153 (0.558)*	1.153 (0.560)*	1.153 (0.560)*
$CommonTie_{ijt} * Similarity_{ijt}$	-0.089 (0.109)	-0.089 (0.127)	-0.089 (0.127)
Other control variables	Y	Y	Y
Time-specific effects	Y	Y	Y
Control for category entropy of i and j	Y	Y	Y
Log likelihood	-28196	-28196	-28196
Stage 2: Reciprocation Impact			
Dependent Variable / Independent Variable	$Log\Delta Sub_{it}$		
FB_{ijt}	2.587 (0.224)***	2.587 (0.229)***	2.587 (0.231)***
$FB_{ijt} * CommonTie_{ijt}$	-0.222 (0.032)***	-0.222 (0.046)***	-0.222 (0.034)***
$FB_{ijt} * Similarity_{ijt}$	-3.297 (0.443)***	-3.297 (0.488)***	-3.297 (0.500)***
$FB_{ijt} * CommonTie_{ijt} * Similarity_{ijt}$	0.144 (0.049)**	0.144 (0.060)*	0.144 (0.056)**
Other control variables	Y	Y	Y
Time-specific effects	Y	Y	Y
Control for category entropy of i and j	Y	Y	Y
Control for subscribers/views of i's	Y	Y	Y
R^2	72.3%	72.3%	72.3%
Observations	112,818	112,818	112,818
Provider pairs	4,478	4,478	4,478

Note. All estimations use the same model specification as in the main result estimation in Column 4 of Table 3. Column (1) reports the standard errors clustered by the initiating provider, Column (2) reports the standard errors clustered by the responding provider, and Column (3) clusters the standard errors in the first stage by the responding provider and the second stage by the initiating provider. ***p<0.001, **p<0.01, *p<0.05.

Appendix H. Including Reciprocations Occurring on the Same Day as Tie Initiation

We estimate the model with the same-day reciprocations included. For the same-day reciprocations, in column (1), we treat higher-status providers as the initiators, in column (2), we treat lower-status providers as the initiators, and in column (3), we randomly assign one provider in each pair to be the initiator. We find that our results are quantitively unchanged. (Here, for identification purposes, we assume the responding providers reciprocate the following day for the same day reciprocations.)

Table H. Two-Stage Estimation Results including Same-Day Reciprocations

Stage 1: Reciprocation Probability			
	(1)	(2)	(3)
Independent Variable \ Dependent Variable	FB_{ijt}	FB_{ijt}	FB_{ijt}
$CommonTie_{ijt}$	0.316 (0.016)***	0.356 (0.017)***	0.347 (0.016)***
$Similarity_{ijt}$	0.301 (0.129)*	0.754 (0.147)***	0.474 (0.137)***
$CommonTie_{ijt} * Similarity_{ijt}$	-0.010 (0.022)	0.003 (0.024)	0.002 (0.023)
Other control variables	Y	Y	Y
Time-specific effects	Y	Y	Y
Control for category entropy of i and j	Y	Y	Y
Log likelihood	-39461	-29553	-34751
Stage 2: Reciprocation Impact			
Independent Variable \ Dependent Variable	$Log\Delta Sub_{it}$	$Log\Delta Sub_{it}$	$Log\Delta Sub_{it}$
FB_{ijt}	1.963 (0.141)***	2.613 (0.227)***	2.319 (0.187)***
$FB_{ijt} * CommonTie_{ijt}$	-0.142 (0.038)***	-0.260 (0.048)***	-0.218 (0.041)***
$FB_{ijt} * Similarity_{ijt}$	-2.295 (0.347)***	-2.876 (0.403)***	-2.544 (0.472)***
$FB_{ijt} * CommonTie_{ijt} * Similarity_{ijt}$	0.133 (0.042)***	0.102 (0.048)*	0.145 (0.035)***
Other control variables	Y	Y	Y
Time-specific effects	Y	Y	Y
Control for category entropy of i and j	Y	Y	Y
Control for subscribers/views of i's outdegrees/indegrees	Y	Y	Y
R^2	73.6%	72.4%	73.0%
Observations	119,770	119,770	119,770
Provider pairs	4,738	4,738	4,738

Note. All estimations use the same model specification as in the main result estimation in Column 4 of Table 3.
***p<0.001, **p<0.01, *p<0.05.

Appendix I. The Mechanism of Triadic Closure

Adding $CommonFeaturing$ and $IndirectPath$ as two additional controls does not qualitatively change the coefficient estimates of $CommonTie_{ijt}$. Thus, our conclusions on hypotheses H1 and H3 still hold. According to the results in Table I, neither $CommonFeaturing$ or $CommonFeaturing_{ijt} * Similarity_{ijt}$ significantly affect the reciprocation probability by responding providers or reciprocation impact on the initiating provider. We also find that the effect of $ReversedIndirectPath$ and that of $ReversedIndirectPath * Similarity$ on reciprocation probability and reciprocation impact are always opposite the effect of $CommonTie$ and that of $CommonTie * Similarity$, consistent with the fact that $ReversedIndirectPath$ represents the paths through which viewers of initiating providers can discover responding providers. All of these results validate that the process of triadic closure in our context is common friends rather than homophily.

Table I. Robustness Checks for the Mechanism of Common Ties

Stage 1: Reciprocation Probability	
Dependent Variable / Independent Variable	FB_{ijt}
$CommonFeaturing_{ijt}$	-0.012 (0.020)
$ReversedIndirectPath_{ijt}$	-0.664 (0.088)***
$CommonTie_{ijt}$	0.304 (0.022)***
$Similarity_{ijt}$	1.087 (0.178)***
$CommonFeaturing_{ijt} * Similarity_{ijt}$	0.029 (0.027)
$ReversedIndirectPath_{ijt} * Similarity_{ijt}$	0.791 (0.104)***
$CommonTie_{ijt} * Similarity_{ijt}$	-0.107 (0.030)***
Stage 2: Reciprocation Impact	
Dependent Variable / Independent Variable	$Log\Delta Sub_{it}$
FB_{ijt}	2.181 (0.194)***
$FB_{ijt} * CommonFeaturing_{ijt}$	0.042 (0.099)
$FB_{ijt} * ReversedIndirectPath_{ijt}$	0.460 (0.163)**
$FB_{ijt} * CommonTie_{ijt}$	-0.235 (0.059)***
$FB_{ijt} * Similarity_{ijt}$	-3.359 (0.505)***
$FB_{ijt} * CommonFeaturing_{ijt} * Similarity_{ijt}$	0.100 (0.117)
$FB_{ijt} * ReversedIndirectPath_{ijt} * Similarity_{ijt}$	-0.388 (0.191)*
$FB_{ijt} * CommonTie_{ijt} * Similarity_{ijt}$	0.164 (0.074)*
Other control variables	Y
Time-specific effects	Y
Control for category entropy of i and j	Y
Control for subscribers/views of i's outdegrees/indegrees	Y
Stage-1 log likelihood	-27963
Stage-2 R^2	72.7%
Observations	112,818
Provider pairs	4,478

Note. All estimations use the same model specification as in the main result estimation in Column 4 of Table 3. ***p<0.001, **p<0.01, *p<0.05.

Appendix J. Robustness Checks for Alternative Measurement of Content Similarity, Common Ties, and Reciprocation

In the main analysis, the content similarity between providers is measured as cosine similarity. Alternatively, we can use dice similarity or Jaccard similarity. Columns (2) to (4) of Table J.1 present the results using these alternative similarity measures. All the results are qualitatively consistent with our main results and support our research hypotheses.

Instead of using common ties in the main analysis, we use the measurement of topological overlap (i.e., $TOverlap_{ijt}$) as a robustness check. The results shown in Column (1) of Table J.1 are generally consistent with those in the main analyses.

Table J.1. Robustness Checks for Common Ties and Content Similarity Measures

Stage 1: Reciprocation Probability			
	(1) Topological Overlap	(2) Dice Similarity	(3) Jaccard Similarity
Dependent Variable / Independent Variable	FB_{ijt}		
$TOverlap_{ijt}$	1.420 (0.143)***		
$CommonTie_{ijt}$		0.290 (0.018)***	0.288 (0.017)***
$Similarity_{ijt}$	1.303 (0.174)***	1.213 (0.181)***	1.235 (0.189)***
$TOverlap_{ijt} * Similarity_{ijt}$	-2.038 (0.184)***		
$CommonTie_{ijt} * Similarity_{ijt}$		-0.091 (0.026)***	-0.094 (0.026)***
Other control variables	Y	Y	Y
Time-specific effects	Y	Y	Y
Control for category entropy of i and j	Y	Y	Y
Log likelihood	-31694	-28190	-28186
Stage 2: Reciprocation Impact			
Dependent Variable / Independent Variable	$Log\Delta Sub_{it}$		
FB_{ijt}	3.624 (0.281)***	2.585 (0.234)***	2.520 (0.267)***
$FB_{ijt} * TOverlap_{ijt}$	-2.761 (0.256)***		
$FB_{ijt} * CommonTie_{ijt}$		-0.213 (0.050)***	-0.187 (0.043)***
$FB_{ijt} * Similarity_{ijt}$	-3.499 (0.449)***	-3.384 (0.511)***	-3.413 (0.397)***
$FB_{ijt} * TOverlap_{ijt} * Similarity_{ijt}$	2.970 (0.314)***		
$FB_{ijt} * CommonTie_{ijt} * Similarity_{ijt}$		0.132 (0.065)*	0.096 (0.058)+
Other control variables	Y	Y	Y
Time-specific effects	Y	Y	Y
Control for category entropy of i and j	Y	Y	Y
Control for subscribers/views of i's	Y	Y	Y
R^2	72.3%	72.3%	72.3%
Observations	112,818	112,818	112,818
Provider pairs	4,478	4,478	4,478

Note. All estimations use the same model specification as in the main result estimation in Column 4 of Table 3, with the only differences being in the common tie and similarity measure. ***p<0.001, **p<0.01, *p<0.05, +p<0.1.

We then use FB_weight_{jit} to replace FB. The results are shown in Table J.2, and we find that when two providers have more common ties or when they produce more similar content, responding providers tend to place initiating providers in a more prominent position. The more prominently the position initiating providers are ranked in the responding providers' page, the more benefits the initiating providers receive. Additionally, common ties and content similarity negatively moderate the reciprocation effect. The interaction effect between common ties and content similarity on the reciprocation impact is significantly negative. All these findings are consistent with our main analyses.

Table J.2. Robustness Checks for Reciprocation Decisions (FB)

Stage 1: Reciprocation Probability		
Dependent Variable / Independent Variable	FB_weight_{jit}	
$CommonTie_{ijt}$	0.028 (0.001)***	0.032 (0.001)***
$Similarity_{ijt}$	0.091 (0.011)***	0.103 (0.011)***
$CommonTie_{ijt} * Similarity_{ijt}$		-0.007 (0.002)***
Other control variables	Y	Y
Time-specific effects	Y	Y
Control for category entropy of i and j	Y	Y
R^2	12.1%	12.1%
Stage 2: Reciprocation Impact		
Dependent Variable / Independent Variable	$Log\Delta Sub_{it}$	
FB_weight_{jit}	5.595 (0.608)***	5.849 (0.961)***
$FB\_weight_{jit} * CommonTie_{ijt}$	-0.154 (0.075)*	-0.286 (0.089)***
$FB\_weight_{jit} * Similarity_{ijt}$	-5.198 (0.911)***	-5.629 (0.851)***
$FB\_weight_{jit} * CommonTie_{ijt} * Similarity_{ijt}$		0.189 (0.028)***
Other control variables	Y	Y
Time-specific effects	Y	Y
Control for category entropy of i and j	Y	Y
Control for subscribers/views of i's outdegrees/indegrees	Y	Y
R^2	72.1%	72.1%
Observations	112,818	112,818
Provider pairs	4,478	4,478

Note. All estimations use the same model specification as in the main result estimation in Column 2 and Column 4 of Table 3, with the only differences in FB measure. ***p<0.001, **p<0.01, *p<0.05.

Appendix K. Alternative Model Specification and Estimation Results

We consider reciprocation as a one-time decision, such that

$$FB_{ij} = 1[W_{ij}\kappa_1 + \zeta_{1ij} > 0] \tag{K.1}$$

where $W_{ij} = [1, X'_i, X_j, Y_{ij}]$, and X'_i, X_j, and Y_{ij} are as defined in equation (3). But without the time index, they take the values when the responding provider j reciprocates for the reciprocated pairs ($FB_{ij} = 1$) and the values on the 14th day after the initiating provider i initiates for the non-reciprocated pairs ($FB_{ij} = 0$). The maximum duration for the reciprocation decision is set at 14 days because based on our data, reciprocation rarely occurs beyond 14 days after the initiation. Accordingly, the cross-sectional specification for reciprocation impact is

$$Log\Delta Sub_{iT} = W_{ij}FB_{ij}\kappa_2 + Z_i\varphi + \zeta_{2ij} \tag{K.2}$$

where $Log\Delta Sub_{iT}$ is the new subscribers received within T days after reciprocation (the 14th day since initiation) for reciprocated (non-reciprocated) pairs. Similar to W_{ij}, Z_i is as defined in equation (1) and takes the values on the day of reciprocation for reciprocated pairs and on the 14th day after initiation for non-reciprocated pairs. The reciprocation effect is identified by comparing the initiating providers of the reciprocated pairs to those of the non-reciprocated pairs. We assume that

$$\begin{bmatrix}\zeta_{1ij}\\ \zeta_{2ij}\end{bmatrix} \sim N[0,\Sigma], \Sigma = \begin{pmatrix}\partial_1^2 & \iota\partial_1\partial_2\\ \iota\partial_1\partial_2 & \partial_2^2\end{pmatrix}$$

where $\partial_1^2 = 1$ under the probit assumption, and ι captures the correlation between reciprocation probability and impact.

We then use a maximum likelihood estimation (MLE) to jointly estimate equations (K.1) and (K.2). To construct the likelihood function, we need to consider two scenarios, specified as follows:

• Scenario 1: FB_{ij} = 0, which indicates that the initiating provider was not reciprocated by the responding provider. The likelihood corresponding to this scenario is

$$\begin{aligned} L_{1ij} &= P\left(W_{ij}\kappa_1 + \zeta_{1ij} \le 0, Log\Delta Sub_{iT}\right)\\ &= P\left(\zeta_{1ij} \le -W_{ij}\kappa_1 \middle| \zeta_{2ij}\right) * f(Log\Delta Sub_{iT})\\ &= \Phi\left(\frac{-W_{ij}\kappa_1 - \frac{\partial_1}{\partial_2}\iota\left(Log\Delta Sub_{iT} - W_{ij}FB_{ij}\kappa_2 - Z_i\varphi\right)}{\sqrt{(1-\iota^2)\partial_1^2}}\right) * \phi\left(\frac{Log\Delta Sub_{iT} - W_{ij}FB_{ij}\kappa_2 - Z_i\varphi}{\partial_2}\right) \end{aligned} \tag{K.3}$$

where Φ is the standard normal distribution function, and ϕ is the standard normal density function.

• Scenario 2: FB_{ij} = 1, which indicates that the initiating provider was reciprocated by the responding provider. The likelihood corresponding to this scenario is

$$\begin{aligned} L_{2ij} &= P\left(W_{ij}\kappa_1 + \zeta_{1ij} > 0, Log\Delta Sub_{iT}\right)\\ &= P\left(W_{ij}\kappa_1 + \zeta_{1ij} > 0 \middle| \zeta_{2ij}\right) * f(Log\Delta Sub_{iT}) \end{aligned}$$

$$= \left(1 - \mathrm{P}\left(W_{ij}\kappa_1 + \zeta_{1ij} \leq 0 \middle| \zeta_{2ij}\right)\right) * \mathrm{f}(Log\Delta Sub_{iT})$$

$$= \left(1 - \mathrm{P}\left(\zeta_{1ij} \leq -W_{ij}\kappa_1 \middle| \zeta_{2ij}\right)\right) * \mathrm{f}(Log\Delta Sub_{iT})$$

$$= \left[1 - \Phi\left(\frac{-W_{ij}\kappa_1 - \frac{\partial_1}{\partial_2}\iota\left(Log\Delta Sub_{iT} - W_{ij}FB_{ij}\kappa_2 - Z_i\varphi\right)}{\sqrt{(1-\iota^2)\partial_1^2}}\right)\right] * \phi\left(\frac{Log\Delta Sub_{iT} - W_{ij}FB_{ij}\kappa_2 - Z_i\varphi}{\partial_2}\right) \tag{K.4}$$

Therefore, the full log likelihood is given by

$$lnL_{ij} = ln\left(\prod_{ij=1}^{N} L_{1ij}^{(1-FB_{ij})} * L_{2ij}^{FB_{ij}}\right) = \sum_{ij=1}^{N}\left[\left(1 - FB_{ij}\right) * lnL_{1ij} + FB_{ij} * lnL_{2ij}\right] \tag{K.5}$$

The estimation results using this alternative model specification are shown in Table K. With the alternative model specification, the proposed hypotheses still hold.

Table K. Robustness Check for Alternative Model Specification

Stage 1: Reciprocation Probability		
	(1)	(2)
Dependent Variable / Independent Variable	FB_{ij}	
$CommonTie_{ij}$	0.188 (0.022)***	0.190 (0.022)***
$Similarity_{ij}$	0.205 (0.092)*	0.178 (0.095)+
$CommonTie_{ij} * Similarity_{ij}$	-0.029 (0.033)	-0.011 (0.033)
Other control variables	Y	Y
Time-specific effects	Y	Y
Control for category entropy of i and j	Y	Y
Log likelihood	-4580	-4747
Stage 2: Reciprocation Impact		
Dependent Variable / Independent Variable	$Log\Delta Sub_{i,7}$	$Log\Delta Sub_{i,14}$
FB_{ij}	1.634 (0.558)**	2.163 (0.621)***
$FB_{ij} * CommonTie_{ij}$	-0.356 (0.036)***	-0.350 (0.039)***
$FB_{ij} * Similarity_{ij}$	-0.577 (0.257)*	-0.592 (0.282)*
$FB_{ij} * CommonTie_{ij} * Similarity_{ij}$	0.103 (0.060)+	0.122 (0.065)+
Other control variables	Y	Y
Time-specific effects	Y	Y
Control for category entropy of i and j	Y	Y
Control for subscribers/views of i's outdegrees/indegrees	Y	Y
Provider pairs[4]	3,857	2,970
Variance–Covariance Matrix		
∂_1^2	Fixed as 1	Fixed as 1
∂_2^2	2.004 (0.036)	2.119 (0.039)
ι	-0.913 (0.013)	-0.892 (0.015)

Note. Standard errors are shown in parentheses. ***p<0.001, **p<0.01, *p<0.05, +p<0.1

[4] Some initiating providers were reciprocated by multiple responding providers. The reciprocation effect from one responding provider j_1 on initiating provider i cannot be identified if there was reciprocation from another responding provider j_2 within the T days following j_1's reciprocation. In such cases, $i - j_1$ was excluded.

References

Luo X, Zhang J, Duan W (2013) Social media and firm equity value. *Inf. Syst. Res.* 24(1):146-163.

Miritello G (2013) *Temporal Patterns of Communication in Social Networks* (Springer, Berlin).

Mundlak Y (1978) On the pooling of time series and cross-section data. *Econometrica* 46(1):69-85.

Murtazashvili I, Wooldridge JM (2016) A control function approach to estimating switching regression models with endogenous explanatory variables and endogenous switching. *J. Economet.*190(2):252-266.

Ravasz E, Somera A, Mongru D, Oltvai Z, Barab´asi A (2002) Hierarchical organization of modularity in metabolic networks. *Science* 297(5586):1551-1555.

Shannon CE, Weaver W (1949) *The Mathematical Theory of Communication* (University of Illinois Press, Urbana).

Stephen AT, Toubia O (2010) Deriving value from Social Commerce networks. *J. Market. Res.* 47(2):215-228.

Vella F, Verbeek M (1999) Two-step estimation of panel data models with censored endogenous variables and selection bias. *J. Economet.* 90(2):239-263.