Web Appendix for "Modeling Dynamics in Crowdfunding"

A. Data

In this section, we introduce more details about our data. On average, each individual crowdfunder invests \$200 in 3 different projects over the course of 6 investments. The average investment amount in each investment is \$35. The average cost of a share is \$8.90. On average, the artists promised to share 44.7% of their stocks after goal completion. Out of 300 fundraising projects, only 22.3% projects completed their goal. It took 53 weeks on average for a successful project to reach its goal. In addition, the average goal completion level is 27.4%. More specifically, the average funding goal is \$27,818 and the average amount of funds raised is \$8,247.

B. Asymmetric Contemporaneous Social Interactions

In this section, we show the asymmetric relations between an individual crowdfunder and the crowd under our context. Specifically, we show that a focal crowdfunder's current decision does not have statistically significant effects on the others' current decision. This result implies that the total investment amount in project j by others at the current period (Y_{-ijt}) is not endogenous (no simultaneity) and therefore, the results of the exploratory analysis in the main article are consistent. Furthermore, this result justifies the adoption of the approximate aggregation approach, which reflects the asymmetric nature of contemporaneous social interactions between a focal crowdfunder and the crowd. We conduct two different tests to examine the effect of a focal crowdfunder's current decision on the crowd's current decision.

First, we conduct a Durbin–Wu–Hausman test (augmented regression test) for endogeneity (Davidson and MacKinnon 1995, Hausman 1978) to examine whether the total investment amount in project *j* by others at the current period (Y_{-ijt}) is endogenous in the original model (OLS 4 in the main article). In the first step, we regress Y_{-ijt} on potential excluded instruments and exogenous regressors. For potential excluded instruments which affect Y_{-ijt} but not Y_{ijt}, we use the total investment amount by others than *i* (-*i*) in other projects than *j* (-*j*) in week *t*-1, $Y_{-i,-j,t-1}$. This can be qualified as an excluded instrument, because the reference group (-*i*)'s current (Y_{-ijt}) and previous decisions ($Y_{-i,-j,t-1}$) are intercorrelated, and the focal crowdfunder's decision in project *j* (Y_{ijt}) might not be affected by the decision of others in projects other than *j* ($Y_{-i,-j,t-1}$). In the second step, we add the residuals obtained from the first stage regression into the original regression of the individual investment (OLS 4 in the main article). If the residual is relevant, then the total investment amount by others at the current period (Y_{-ijt}) is endogenous. Table 1 shows the results of the regression of the individual investment on the three factors and the residual (AR1). The residual is not significant, showing the possibility of no endogeneity. Furthermore, we compare the regression (AR1) with the original model (OLS 4 in the main article) by conducting a Wald test. We fail to reject the null hypothesis that the residual is irrelevant (F-statistics = 1.863, p-value = 0.828). This result implies that the original model (OLS 4) is not inconsistent and the total investment amount in project *j* by others at the current period (Y_{-ijt}) is not endogenous (no simultaneity).

Table 1. Linear Regression of the Individual Investment on the Residual and Three Factors

	AR 1		
	Est	SE	P-Value
Intercept	20.74	3.79	< 0.0001
Remaining amount left until goal at the next period (R_{jt+1})	-0.50	0.05	< 0.0001
Total investments by others at the current period (Y_{-ijt})	-0.39	3.15	0.901

Residual (from the first step)	4.30	3.16	0.173
Total investments by others at the next period (Y_{-ijt+1})	2.78	1.14	0.015
Individual fixed effects (δ_i)		Yes	
Project fixed effects (γ_j)		Yes	
Week fixed effects (w_t)		Yes	

Est: Estimate, SE: Standard Error

Second, we check whether a focal crowdfunder's decision has a statistically significant effect on the aggregate behavior of others, following Nair et al. (2010)'s approach. Since the focal crowdfunder's decision is endogenous, we run IV regressions. In the first step, we regress the individual investment (Y_{ijt}) on excluded instruments and exogenous variables. For excluded instruments which affect the individual investment in project j (Y_{ijt}) not the total investments in project j by others (Y_{-ijt}), we use the individual investment in other projects than j ($-j$) in week t-1 ($Y_{i,-j,t-1}$). The results from the first-stage regressions appear in Table 2. IV-FS 1, IV-FS 2, and IV-FS 3 include excluded instruments, exogeneous regressors, and all instruments, respectively. The signs of the parameters are consistent with the OLS results and intuitions. Also, the F-statistics from all regressions strongly reject the null hypothesis that the instruments have no explanatory power. We additionally run a Wald test to compare IV-FS 2 and IV-FS 3 (relevance of excluded instruments). We also strongly reject the null hypothesis that the excluded instrument is irrelevant (F-statistics = 301.69, p-value = < 0.0001). All these results show that the instruments are working correctly. In the second step of IV regressions, we run a regression of the total investments in project j by others at the current period (Y_{-ijt}) on exogeneous regressors and the fitted endogenous regressor obtained from the first stage (2SLS). The results of the IV regression appear in Table 3. We find that a focal crowdfunder's behavior does not have a statistically significant effect on the aggregate behavior of others. This result implies that the interactions between a focal crowdfunder and others under our context is asymmetric and therefore removes the simultaneity concerns. More importantly, this result justifies the adoption of the approximate aggregation approach, which reflects the asymmetric nature of contemporaneous social interactions between a focal crowdfunder and the crowd.

Table 2. First-Stage Regression of Individual Investment

	IV-FS 1		IV-FS 2		IV-FS 3	
	Est	SE	Est	SE	Est	SE
Intercept	3.68	0.08	20.45	2.96	20.53	2.96
Individual investment in other projects than *j* in week *t*-1 ($Y_{i,-j,t-1}$)	14.35	0.47			8.71	0.50
Remaining amount left until goal at the next period (R_{jt+1})			-0.49	0.02	-0.49	0.02
Total investments by others at the next period (Y_{-ijt+1})			2.63	0.15	2.63	0.15
Individual fixed effects			Yes		Yes	
Project fixed effects			Yes		Yes	
Week fixed effects			Yes		Yes	
F-Statistics	943.04		8.20		8.37	
(P-value)	(< 0.0001)		(< 0.0001)		(< 0.0001)	

Est: Estimate, SE: Standard Error

Table 3. IV Regression of the Total Investments by Others

	IV-Final		
	Est	SE	P-Value
Intercept	0.750	0.042	< 0.0001
Remaining amount left until goal at the next period (R_{jt+1})	-0.014	0.0004	< 0.0001
Individual Investment at the current period (fitted)	0.0001	0.001	0.911
Total investments by others at the next period (Y_{-ijt+1})	0.358	0.003	< 0.0001
Individual fixed effects (δ_i)		Yes	
Project fixed effects (γ_j)		Yes	
Week fixed effects (w_t)		Yes	

Est: Estimate, SE: Standard Error

C. Estimation procedure

We adopt the Bayesian IJC method to estimate the proposed model (Ching et al., 2012; Imai et al., 2009).

1) Draw a $\mathbf{\Omega}_i^{(n)}$

We sample individual specific $\mathbf{\Omega}_i$ using a random walk Metropolis–Hastings algorithm. The new draw $\mathbf{\Omega}_i^{(n)}$ is drawn from the normal distribution, $N\left(\mathbf{\Omega}_i^{(o)}, s_\Omega \mathbf{I}_k\right)$, where $\mathbf{\Omega}_i^{(o)}$ is an old draw of the parameter vector, s_Ω is a bandwidth which is set to have an acceptance rate of approximately 30%, $\mathbf{I}_k$ is the identity matrix, k is the length of $\mathbf{\Omega}_i$.

2) Compute $\tilde{E}\left[V_{ij}\left(S_{ijt+1};\Omega_{ij}^{(n)} \mid S_{ijt}, q_{ijt}\right)\right]$ and $\tilde{E}\left[V_{ij}\left(S_{ijt+1};\Omega_{ij}^{(o)} \mid S_{ijt}, q_{ijt}\right)\right]$ for all t ($A_{jt-1} < G_j$) and all j.

The Bayesian IJC method approximates the expected value function by using previously stored pseudo-expected value functions $W_{ij}^l\left(S_{ij}^l;\Omega_{ij}^l\right)$ from $H_{ij}^r = \left\{A_j^l, a_j^l, \Omega_{ij}^l, W_{ij}^l\left(., A_j^l, a_j^l;\Omega_{ij}^l\right)\right\}_{l=r-N}^{r-1}$, where r is the current iteration and N is the number of past value functions to store. Since the cumulative total shares (A) and period total shares (a) have a large support compared to the cumulative individual shares (C), we regard the cumulative total shares and period total shares as a continuous variable. Thus, we store the pseudo-expected value functions at all $c_{ij} \in \left\{1, \cdots, C_j^{\max}\right\}$. $C_j^{\max}$ is the maximum of cumulative individual shares observed from the data. The pseudo-expected value functions can be obtained as

$$\tilde{E}\left[V_{ij}\left(S_{ijt+1};\Omega_{ij} \mid S_{ijt}, q_{ijt}\right)\right] = \left[\sum_{l=r-N}^{r-1} W_{ij}^l\left(C_{ijt-1} + q_{ijt}, A_j^l, a_j^l,;\Omega_{ij}^l\right) \cdot \frac{K_{h^\Omega}\left(\Omega_{ij}^l, \Omega_{ij}\right) K_{h^A}\left(A_j^l, A_{jt-1} + a_{jt}\right) f\left(a_j^l \mid a_{jt};\phi_j\right)}{\sum_{m=r-N}^{r-1} K_{h^\Omega}\left(\Omega_{ij}^m, \Omega_{ij}\right) K_{h^A}\left(A_j^m, A_{jt-1} + a_{jt}\right) f\left(a_j^m \mid a_{jt};\phi_j\right)}\right],$$

where K_{h^Ω} and K_{h^A} is the Gaussian kernel with bandwidth 5 and 1000, respectively, and f is the normal density of the period total shares in Equation (2) from the main article.

3) Compute likelihoods

We compute the likelihood function in Equation (7) from the main article using the pseudo-expected value functions obtain from the previous step. Then, we accept the new draw with the probability,

$$\min\left(\frac{L\left(\boldsymbol{\Omega}_i^{(n)}\right)\cdot \exp\left[-\frac{1}{2}\left(\boldsymbol{\Omega}_i^{(n)}-\bar{\boldsymbol{\Omega}}\right)' V_\Omega^{-1}\left(\boldsymbol{\Omega}_i^{(n)}-\bar{\boldsymbol{\Omega}}\right)\right]}{L\left(\boldsymbol{\Omega}_i^{(o)}\right)\cdot \exp\left[-\frac{1}{2}\left(\boldsymbol{\Omega}_i^{(o)}-\bar{\boldsymbol{\Omega}}\right)' V_\Omega^{-1}\left(\boldsymbol{\Omega}_i^{(o)}-\bar{\boldsymbol{\Omega}}\right)\right]},1\right).$$

4) Update and store $W_{ij}^r\left(.,A_j^r,a_j^r;\Omega_{ij}^r\right)$ for all j

To update the pseudo-expected value function, we simulate a draw of A_j^r and a_j^r from the uniform distribution over the support of $\left(0,1.1\cdot G_j\right)$ and $\left(0,a_j^{\max}\right)$, respectively. $a_j^{\max}$ is the maximum of period total shares observed from the data. We obtain the alternative specific ($\tilde{q}_{ij}$) pseudo-expected value function as

$$\begin{aligned} W_{ij}^{r,\tilde{q}_{ij}}\left(s_{ij};\Omega_{ij}^r\right) &= u_{ij}^P\left(\tilde{q}_{ij}\middle|s_{ij}\right)+\rho\tilde{E}\left[V_{ij}\left(s_{ij}';\Omega_{ij}^{(n)}\mid s_{ij},\tilde{q}_{ij}\right)\right],\ \text{if } A_j^r < G_j \\ &= \frac{c_{ij}}{G_j}M_jR_j,\ \text{if } A_j^r \ge G_j, \end{aligned}$$

for all s_{ij} and $\tilde{q}_{ij}$, where $s_{ij}=\left(c_{ij},A_j^r,a_j^r\right), c_{ij}\in\left\{1,\cdots,C_j^{\max}\right\}, \tilde{q}_{ij}\in\mathbf{Q}$.

Then, we update and store the pseudo-expected value function and as

$$W_{ij}^{r}\left(c_{ij}, A_{j}^{r}, a_{j}^{r}; \Omega_{ij}^{r}\right) = \ln\left[\sum_{\tilde{q}_{ijt}\in\mathbf{Q}} \exp\left(W_{ij}^{r,\tilde{q}_{ij}}\right)\right] \text{ and } \Omega_{ij}^{r} = \Omega_{ij}^{(n)} \text{ for all } c_{ij} \in \left\{1,\cdots,C_{j}^{\max}\right\}.$$

5) Repeat steps 1 to 4 for all individuals $i \in \{1,\cdots,I\}$

6) Generate $\bar{\mathbf{\Omega}} \big| \{\mathbf{\Omega}_i\}, V_\Omega$

$$\bar{\mathbf{\Omega}} \sim MVN_k\left(\bar{\bar{\mathbf{\Omega}}}, \left(\left(V_\Omega/I\right)^{-1} + \left(100\mathbf{I}_k\right)^{-1}\right)^{-1}\right),$$

where $\bar{\bar{\mathbf{\Omega}}} = \left(\left(V_\Omega/I\right)^{-1} + \left(100\mathbf{I}_k\right)^{-1}\right)^{-1}\left(V_\Omega^{-1}\sum_{i=1}^{I}\mathbf{\Omega}_i + \left(100\mathbf{I}_k\right)^{-1}(0)\right)$.

7) Generate $V_\Omega \big| \{\mathbf{\Omega}_i\}, \bar{\mathbf{\Omega}}$

$$V_\Omega \sim IW\left(v_\Omega + I, \sum_{i=1}^{I}\left(\mathbf{\Omega}_i - \bar{\mathbf{\Omega}}\right)'\left(\mathbf{\Omega}_i - \bar{\mathbf{\Omega}}\right) + \bar{S}_\Omega\right), \text{ where } v_\Omega = k+3, \bar{S}_\Omega = v_\Omega \mathbf{I}_k.$$

D. Simulating Rational Expectations Equilibrium

We adopt Ahn et al. (2015)'s approach to simulate the rational expectations equilibrium. Note that we obtain the rational expectations equilibrium when simulating investment demands (e.g., prediction, counterfactuals, etc.).

1) Choose project j and set the structural parameters Ω_{ij}.

2) Guess arbitrary $\phi_j^{(0)}$ and $\phi_j^{(1)}$ for the transition of the period total sales, a_{jt}.

3) Obtain a_{jt} for all t (from $t = 0$ to $t = T$)

   a. Choose an arbitrary a_{jt}^{old}

b. Simulate individual demand $Q_{ijt}^{new} \mid a_{jt}^{old}$ for all *i* using Equation (6) in the main article.

i. To compute $\tilde{E}\left[V_{ij}\left(S_{ijt+1} \middle| S_{ijt}\right)\right]$, we generate 1,000 pseudo-expected value functions recursively running Step 4) in Web Appendix B.

c. Compute a_{jt}^{new} by adding up $Q_{ijt}^{new} \mid a_{jt}^{old}$ for all *i*. If $\left|a_{jt}^{old} - a_{jt}^{new}\right| < \delta^{a}$, then stop. Otherwise, set $a_{jt}^{old} = a_{jt}^{new}$ and run step 3-a) through step 3-c) until convergence.

4) Obtain $\tilde{\phi}_{j}^{(0)}$ and $\tilde{\phi}_{j}^{(1)}$ by estimating $a_{jt} = \tilde{\phi}_{j}^{(0)} + \tilde{\phi}_{j}^{(1)} a_{jt-1} + \eta_{jt}$.

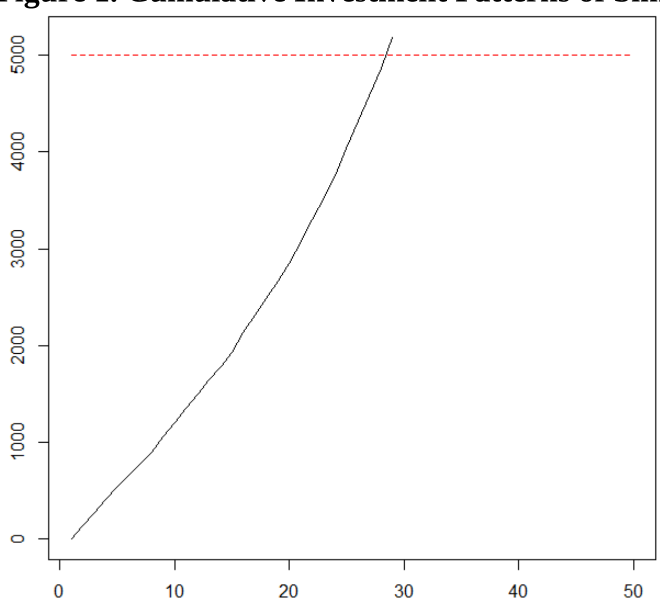
5) If $\max\left(\left|\phi_{jt} - \tilde{\phi}_{jt}\right|\right) < \delta^{\phi}$, then stop. Otherwise, set $\phi_{jt}^{(0)} = \tilde{\phi}_{jt}^{(0)}$ and $\phi_{jt}^{(1)} = \tilde{\phi}_{jt}^{(1)}$, and run step 3) through step 5) until convergence.

E. Simulation Study

We conduct a simulation study to see whether the true model parameters are recovered well. In the simulation study, we allow for consumer heterogeneity by using a random-effect specification for 100 individual crowdfunders (I = 100). 50 investment decisions of each individual crowdfunder for one crowdfunding project are generated (T = 50, J = 1). We set the similar settings with our empirical analysis (ρ = 0.9, G = 5,000, M = 10,000, R = 0.5, H = 0.1). Individual-level investment decisions are simulated by following steps suggested in Web Appendix D. However, we compute the expected value function, $E\left[V_{ij}\left(S_{ijt+1} \middle| S_{ijt}\right)\right]$ by using the nested fixed point algorithm (Rust, 1987) instead of the Bayesian IJC method (Ching et al., 2012; Imai et al., 2009) in data generating process in order to examine whether the Bayesian IJC method can recover the true expected value functions generated by the nested fixed point algorithm. In the data generation process, we obtain the expected value functions by recursively applying the Bellman operator in Equation (5) from the main article until convergence. In our estimation, we obtain the pseudo-expected value functions by following the steps suggested in Web Appendix C. We keep every 10th draw among 15,000 MCMC draws and use the last 5,000 MCMC draws to report the posterior distribution of model parameters. Also, 1,000 past pseudo-expected value functions are used to approximate the expected value function (N = 1000). We set the bandwidth for $K_{h^{\Omega}}$ and $K_{h^{A}}$ as 1 and 100, respectively.

The project is assumed to manifest prominent delaying behavior and social interactions. Figure 1 shows the cumulative total shares of the project by aggregating the individual-level investment decisions simulated. We can see that the individual crowdfunders postpone their investment in the early stage and then rushes into as the goal gets closer. Finally, it achieves the goal in period 30. Table 4 shows the simulation results. We find that all the true parameters are well recovered.

Figure 1. Cumulative Investment Patterns of Simulated Data



Solid & Black: Cumulative Total Shares
Dashed & Red: Goal

Table 4. Simulation Results

Parameter	True Value	Posterior Mean	Posterior Std. Dev.
$\bar{\tilde{\alpha}}$	1.50	1.41	0.24
$\bar{\tilde{\pi}}$	4.50	4.69	0.13
$\bar{\tilde{v}}$	0.50	0.77	0.13
$\bar{\tilde{\theta}}$	0.00	-0.21	0.18
$\bar{\tilde{k}}$	-5.00	-5.76	0.12

[References]

Ahn, D.-Y., Duan, J.A., and Mela, C.F. (2015). Managing user-generated content: A dynamic rational expectations equilibrium approach. Mark. Sci. *35*, 284–303.

Ching, A.T., Imai, S., Ishihara, M., and Jain, N. (2012). A practitioner's guide to Bayesian estimation of discrete choice dynamic programming models. Quant. Mark. Econ. *10*, 151–196.

Imai, S., Jain, N., and Ching, A. (2009). Bayesian Estimation of Dynamic Discrete Choice Models. Econometrica *77*, 1865–1899.

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. Econometrica *55*, 999–1033.