

Online Appendix to “‘Be the Buyer’–Leveraging the Wisdom of the Crowd in E-Commerce Assortment Planning”

Leela Nageswaran^a, Yu (Rain) Kan^b, Uttara Ananthakrishnan^c

^aMichael G. Foster School of Business, University of Washington, Seattle, WA 98195, lnages@uw.edu

^bKelley School of Business, Indiana University, IN 47405, rainkan@iu.edu

^cHeinz College of Information Systems and Public Policy, Carnegie Mellon University, PA 15213, uttara@cmu.edu

Appendix A: Theoretical Model for Inventory Decisions

Our goal is to develop a model for inventory decisions in the absence and presence of crowdvoting. Consider a focal product’s inventory decision, say $I(\geq 0)$. Let N denote the total number of users and T denote the total number of rental periods. The per-unit cost of procuring the product is $c(> 0)$ and the platform’s fee is p for each user. Let $r(I)$ denote the percentage chance that a user who is interested in the focal product is able to rent it during the season. Given that the item can be rented up to $I \cdot T$ times, this yields $r(I) = \min\left\{1, \frac{IT}{D}\right\}$, where D is the total number of interested users. Crucially, D is uncertain, and we assume that a fraction $\Theta \sim U[0, \bar{\Theta}]$ of the users, N , are interested in renting the product. Since users are likely to stay on the platform only if the chance of renting items is high enough, we weight the platform’s revenue from each user by $r(I)$, such that the overall revenue is $pNr(I)$.¹⁹ This yields the following expected profit function for the platform in the absence of crowdvoting:

$$\pi(I) = -c \cdot I + pN \cdot E_D \min\left\{1, \frac{IT}{D}\right\}. \quad (11)$$

The platform chooses I to maximize $\pi(I)$. Here the tradeoff is between ordering costly inventory and ensuring a high enough service level, akin to the classical newsvendor model.²⁰

In the presence of crowdvoting, the platform makes an inventory decision, $I_g(\geq 0)$, aided by $v(\leq N)$ voters, out of whom $v_f(\leq v)$ have cast a “Favorite” vote for the item. We assume that the v_f voters are truth-telling and genuinely interested in renting the product. This assumption is verified empirically in Appendix F; specifically, voters who are aware of crowdvoted products available on the platform are more likely to rent the items they voted favorably on. For the $N - v$ users who did not vote, the platform experiences uncertainty as in the absence of crowdvoting regarding how many are interested in renting the product. This yields a total demand $D_g = v_f + (N - v)\Theta$, where $\Theta \sim U[0, \bar{\Theta}]$.²¹ The platform’s expected profit, $\pi_g(I_g)$, is then given by (11), replacing D by D_g .

¹⁹ This approach accounts for the fact that consumers’ decision to subscribe – and provide revenue p to the platform – hinges on the chance that they can rent their preferred items. An alternate approach would be to consider the revenue gained from each active user, with the number of active users increasing in the long-run average chance of renting preferred items; such a model provides qualitatively similar insights albeit without readily interpretable expressions.

²⁰ Prior work on rentals such as Slauch et al. (2016) focus on recirculation decisions considering stochastic rental duration and usage-based loss; in contrast, the rental durations at our partner company are quite consistent and usage-based loss is not as significant as the tradeoff modeled herein between service level and procurement cost.

²¹ The variance of the demand is $\text{var}(D_g) = \left(1 - \frac{v}{N}\right)^2 \text{var}(D)$ and its mean is $\text{mean}(D_g) = v_f - v \frac{\bar{\Theta}}{2} + \text{mean}(D)$. Thus, crowdvoting decreases the extent of demand uncertainty, and it increases the mean demand if $\frac{v_f}{v} > \frac{\bar{\Theta}}{2}$ – that is, the fraction of “Favorite” votes $\frac{v_f}{v}$ is high enough – and decreases the mean demand otherwise.

Our main proposition below reveals when crowdvoting increases the inventory level and the magnitude of such an increase. (The proof is in Appendix B.)

PROPOSITION A1. *The platform's optimal inventory level in the absence of crowdvoting is $I = \frac{\bar{\Theta}N}{T}e^{-\frac{c\bar{\Theta}}{pT}}$, and that in the presence of crowdvoting is $I_g = \frac{v_f+(N-v)\bar{\Theta}}{T}e^{-\frac{c(N-v)\bar{\Theta}}{pNT}}$ if $pNT\frac{1}{\bar{\Theta}}\log\left(\frac{v_f+(N-v)\bar{\Theta}}{v_f}\right)\frac{1}{N-v} > c$ and $I_g = 0$ otherwise. Moreover, provided that $pNT\frac{1}{\bar{\Theta}}\log\left(\frac{v_f+(N-v)\bar{\Theta}}{v_f}\right)\frac{1}{N-v} > c$, we have $I_g > I$ for sufficiently large number of "Favorite" votes, v_f , and $I_g - I$ is increasing in v_f .*

The first hypothesis in §3.2 follows from Proposition A1.

Appendix B: Proof of Proposition A1.

First consider the inventory decision in the absence of crowdvoting. We can simplify $\pi(I) = -cI + pN \int_0^{\bar{\Theta}} \min\left\{1, \frac{IT}{\bar{\Theta}N}\right\} dF(\Theta)$, where $F(x), x \geq 0$ denotes the cumulative density function of Θ , which is $\frac{x}{\bar{\Theta}}$ for $0 \leq x \leq \bar{\Theta}$, and one otherwise. Now

$$\pi(I) = -cI + pN \left[\int_0^{\frac{IT}{N}} dF(\Theta) + \int_{\frac{IT}{N}}^{\bar{\Theta}} \frac{IT}{\bar{\Theta}N} dF(\Theta) \right] = -cI + pIT \frac{1}{\bar{\Theta}} \left[1 + \left(\log \bar{\Theta} - \log\left(\frac{IT}{N}\right) \right) \right].$$

The first derivative with respect to (w.r.t.) I is $-c + pT\frac{1}{\bar{\Theta}} \left[1 + \left(\log \bar{\Theta} - \log\left(\frac{IT}{N}\right) \right) \right] - pIT\frac{1}{\bar{\Theta}} \frac{d}{dI} \log\left(\frac{IT}{N}\right) = -c + pT\frac{1}{\bar{\Theta}} \left[1 + \left(\log \bar{\Theta} - \log\left(\frac{IT}{N}\right) \right) \right] - pT\frac{1}{\bar{\Theta}} = -c + pT\frac{1}{\bar{\Theta}} \log\left(\frac{\bar{\Theta}N}{IT}\right)$. The second derivative w.r.t. I is negative. Thus, the optimal inventory level is $I = \frac{\bar{\Theta}N}{T}e^{-\frac{c\bar{\Theta}}{pT}}$.

Next consider the inventory decision in the presence of crowdvoting. We can simplify $\pi_g(I_g) = -c \cdot I_g + pN \cdot E_{D_g} \min\left\{1, \frac{I_g T}{D_g}\right\} = -c \cdot I_g + pN \cdot E_{\Theta} \min\left\{1, \frac{I_g T}{v_f+(N-v)\bar{\Theta}}\right\}$, and furthermore, $\pi_g(I_g) = -cI_g + pN \int_0^{\bar{\Theta}} \min\left\{1, \frac{I_g T}{v_f+(N-v)\bar{\Theta}}\right\} dF(\Theta)$. This yields

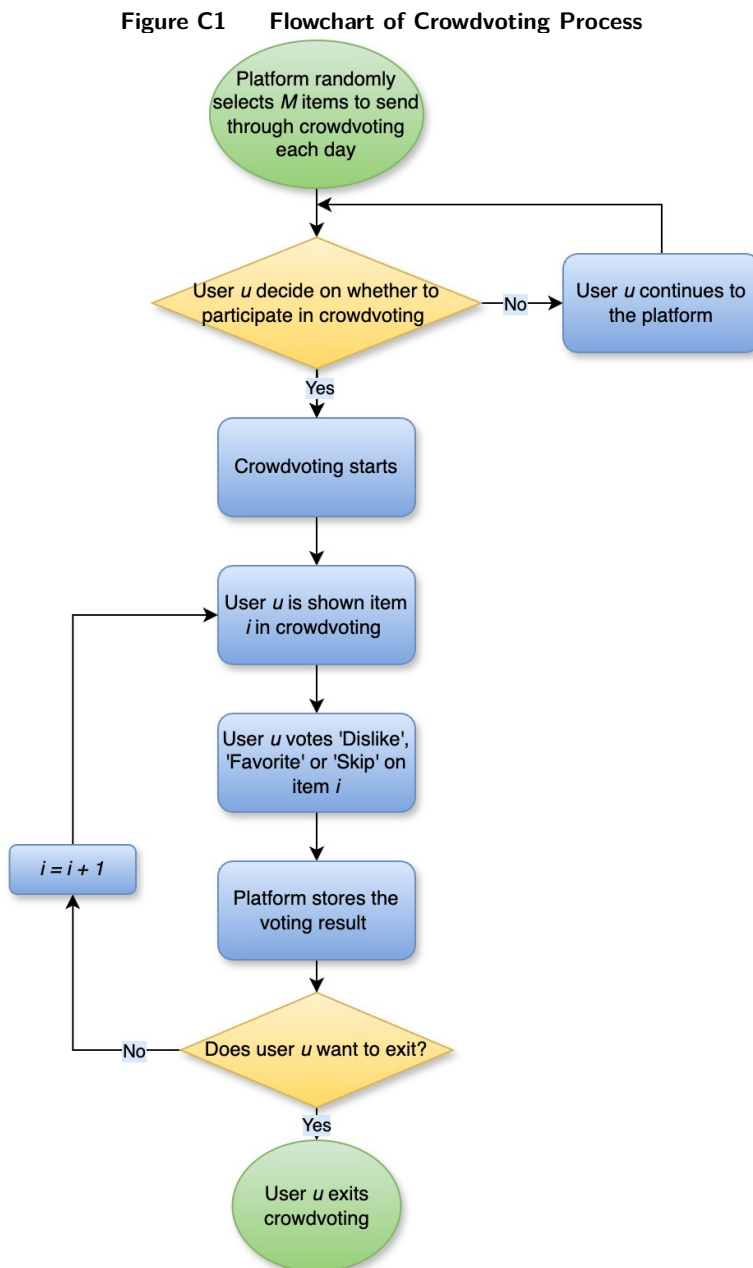
$$\pi_g(I_g) = \begin{cases} -cI_g + pN \left[\int_0^{\frac{I_g T - v_f}{N-v}} dF(\Theta) + \int_{\frac{I_g T - v_f}{N-v}}^{\bar{\Theta}} \frac{I_g T}{v_f+(N-v)\bar{\Theta}} dF(\Theta) \right], & \text{if } I_g > \frac{v_f}{T} \\ -cI_g + pN \left[\int_0^{\bar{\Theta}} \frac{I_g T}{v_f+(N-v)\bar{\Theta}} dF(\Theta) \right], & \text{otherwise.} \end{cases} \quad (12)$$

This implies that when $I_g \leq v_f/T$, we have $\pi_g(I_g) = -cI_g + pNI_g T \frac{1}{\bar{\Theta}} \left[\int_0^{\bar{\Theta}} \frac{1}{v_f+(N-v)\bar{\Theta}} d(\Theta) \right] = -cI_g + pNI_g T \frac{1}{\bar{\Theta}} \log\left(\frac{v_f+(N-v)\bar{\Theta}}{v_f}\right)\frac{1}{N-v}$. This function is increasing in I_g if $pNT\frac{1}{\bar{\Theta}}\log\left(\frac{v_f+(N-v)\bar{\Theta}}{v_f}\right)\frac{1}{N-v} > c$ and decreasing in I_g if $pNT\frac{1}{\bar{\Theta}}\log\left(\frac{v_f+(N-v)\bar{\Theta}}{v_f}\right)\frac{1}{N-v} < c$. And when $I_g > v_f/T$, we have $\pi_g(I_g) = -cI_g + pN \left[\int_0^{\frac{I_g T - v_f}{N-v}} dF(\Theta) + \int_{\frac{I_g T - v_f}{N-v}}^{\bar{\Theta}} \frac{I_g T}{v_f+(N-v)\bar{\Theta}} dF(\Theta) \right] = -cI_g + pN \frac{1}{(N-v)\bar{\Theta}} \left[I_g T - v_f + I_g T \log\left(\frac{v_f+(N-v)\bar{\Theta}}{I_g T}\right) \right]$. This function's first derivative w.r.t. I_g is $-c + pN \frac{1}{(N-v)\bar{\Theta}} T \log\left(\frac{v_f+(N-v)\bar{\Theta}}{I_g T}\right)$ and its second derivative w.r.t. I_g is negative. Thus, the unconstrained optimal solution solves $pN \frac{1}{(N-v)\bar{\Theta}} T \log\left(\frac{v_f+(N-v)\bar{\Theta}}{I_g T}\right) = c$, which can be simplified to $I_g = \frac{v_f+(N-v)\bar{\Theta}}{T}e^{-\frac{c(N-v)\bar{\Theta}}{pNT}}$. I_g is then decreasing in c . At $c = pNT\frac{1}{\bar{\Theta}}\log\left(\frac{v_f+(N-v)\bar{\Theta}}{v_f}\right)\frac{1}{N-v}$, this unconstrained solution evaluates to $I_g = \frac{v_f}{T}$. This implies that the optimal solution is $I_g = 0$ if $pNT\frac{1}{\bar{\Theta}}\log\left(\frac{v_f+(N-v)\bar{\Theta}}{v_f}\right)\frac{1}{N-v} < c$ and I_g otherwise.

Consider the case where $pNT\frac{1}{\bar{\Theta}}\log\left(\frac{v_f+(N-v)\bar{\Theta}}{v_f}\right)\frac{1}{N-v} > c$. Then $I_g - I = \frac{v_f+(N-v)\bar{\Theta}}{T}e^{-\frac{c(N-v)\bar{\Theta}}{pNT}} - \frac{\bar{\Theta}N}{T}e^{-\frac{c\bar{\Theta}}{pT}} = e^{-\frac{c\bar{\Theta}}{pT}} \left[\frac{v_f+(N-v)\bar{\Theta}}{T}e^{\frac{cv\bar{\Theta}}{pNT}} - \frac{\bar{\Theta}N}{T} \right]$. Then, $I_g > I$ for high enough v_f and increasing in v_f . \square

Appendix C: Additional Results: Section 3

Figure C1 shows a flowchart of steps involved in crowdvoting.

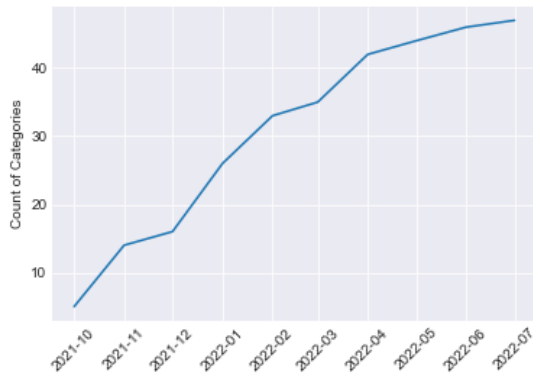


Note. Our partner chose $M \in \{10, 20\}$ during the study period.

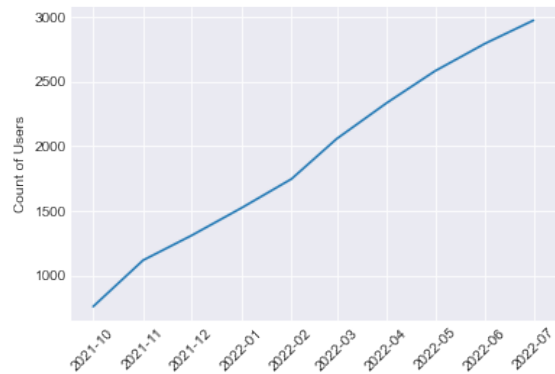
Table C1 shows the rollout schedule at a glance; specifically, the number of newly treated categories by month. Figure C2(a) and (b) show the rollout schedule at the category- and user-levels, respectively.

Table C1 Rollout Schedule: Newly Treated by Month

Month	Number of Categories that Newly Received Treatment
October 2021	5
November 2021	9
December 2021	2
January 2022	10
February 2022	7
March 2022	2
April 2022	7
May 2022	2
June 2022	2
July 2022	1

Figure C2 Rollout Schedule: By Category- and User-Level

(a) At Category Level



(b) At User Level

In Table C2, we summarize the characteristics of the items items purchased after crowdvoting was launched. In Table C3, we present statistics pertaining to category-level outcome variables. We provide the summary statistics at the user level in Table C4.

Table C2 Summary Statistics of Items

	Variable Definition	Mean	Std. Dev.	Min.	Max.
Outcome Variables:					
14DaysRental	Count of rentals within the first 14 days item becomes available on the platform	3.86	4.34	0	35
3MonthsRental	Count of rentals within the first 3 months item becomes available on the platform	9.40	10.1	0	87
RentalIndicator	Binary variable, indicating whether item is rented within the first 14 days it becomes available on the platform	0.83	0.38	0	1
InventoryLevel	Number of copies of each item	5.35	5.66	1	138
InventoryTurns	Count of rentals per copy within the first 6 months item becomes available on the platform	2.75	2.15	0	40
Controls:					
Holiday	Binary variable, indicating whether item is for holiday season	0.62	0.49	0	1
Summer	Binary variable, indicating whether item is for summer season	0.63	0.48	0	1
NeutralColor	Binary variable, indicating whether item is neutral color	0.47	0.50	0	1
ColdColor	Binary variable, indicating whether item is cold color	0.28	0.45	0	1
Relaxed	Binary variable, indicating whether item is relaxed style	0.61	0.49	0	1
Trendy	Binary variable, indicating whether item is trendy style	0.37	0.48	0	1
Dresses	Binary variable, indicating whether item belongs to dresses department	0.35	0.48	0	1
Tops	Binary variable, indicating whether item belongs to tops department	0.31	0.46	0	1
Cotton	Binary variable, indicating whether item is made of cotton	0.35	0.48	0	1
HasPockets	Binary variable, indicating whether item has pockets	0.30	0.46	0	1
LongSleeve	Binary variable, indicating whether item has long sleeves	0.46	0.50	0	1
RetailPrice	Retail price of item	104.3	77.25	13.0	765
TopBrand	Binary variable, indicating whether item is from a popular brand	0.42	0.49	0	1

Note: The statistics for RetailPrice are reported after rescaling due to the nondisclosure agreement.

Table C3 Summary Statistics of Categories

	Variable Definition	Mean	Std. Dev.	Min.	Max.
Outcome Variables:					
14DaysRental	Average count of rentals within the first 14 days item becomes available on the platform	3.01	2.61	0	20
3MonthsRental	Average count of rentals within the first 3 months item becomes available on the platform	7.49	5.93	0	41
Controls:					
ItemCnt	Number of unique styles of items purchased in the category in the month	4.49	8.35	0	98
BrandCnt	Number of unique brands purchased in the category in the month	3.03	4.92	0	46

Table C4 Summary Statistics of Users

	Variable Definition	Mean	Std. Dev.	Min.	Max.
Outcome Variables:					
Cancel	Binary variable, indicating whether the user cancels	0.62	0.49	0	1
ReviewCnt	Count of reviews posted by the user in the month	5.65	6.33	0	50
Treatment:					
CrowdvotingUser	Binary variable, indicating whether user votes	0.39	0.49	0	1
Votes	Count of votes by voting users	66.0	239.3	0	2301
BuyAware	Binary variable, indicating whether user clicks on the items purchased by the platform that they voted "Favorite"	0.23	0.42	0	1
Controls:					
ActiveDayCnt	Count of active days by the user in the month	5.05	5.51	0	31
Tenure	User's tenure in years	0.64	0.96	0	5.55
RentalCnt	Count of rentals by user	53.2	125.3	0	1247
RatingAvg	User's average rating across all rented items	3.84	0.68	1	5
Age	User's age in years	41.8	9.16	11.8	84.3
Central	Binary variable, indicating whether user is in central area	0.13	0.34	0	1
EastCoast	Binary variable, indicating whether user is in east coast area	0.29	0.45	0	1
WestCoast	Binary variable, indicating whether user is in west coast area	0.39	0.49	0	1
South	Binary variable, indicating whether user is in south area	0.19	0.39	0	1

Appendix D: Additional Results: Section 4

We present model-free evidence of crowdvoting’s impact on rental performance in Table D5.

Table D5 Impact of Crowdvoting on Rental Performance: Model-Free Evidence

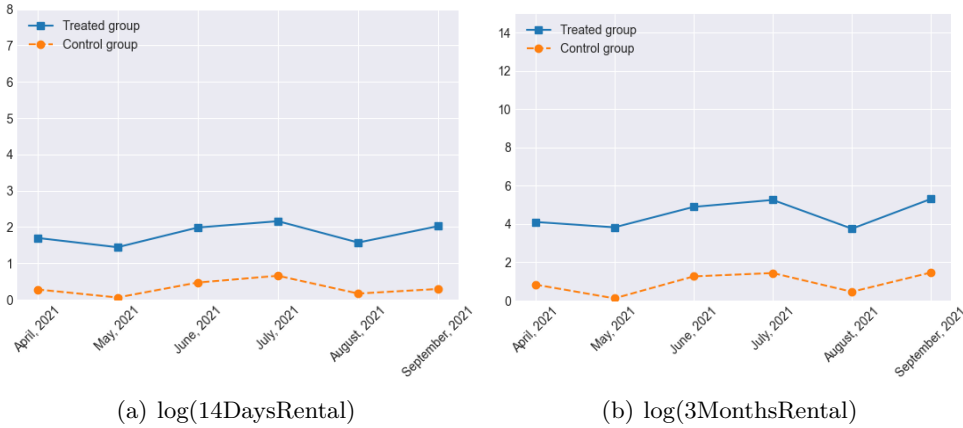
	(1)	(2)
	log(14DaysRental)	log(3MonthsRental)
PostCrowdvoting	0.81*** (0.05)	1.15*** (0.07)
Observations	827	827
R^2	0.254	0.245

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

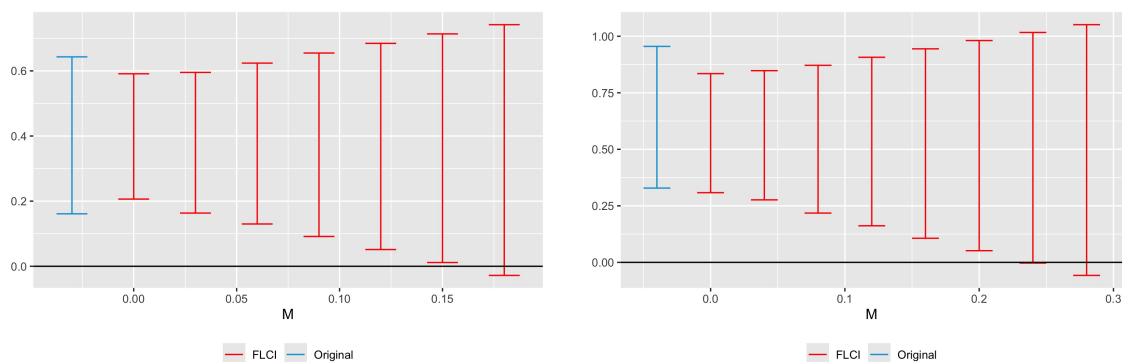
The unit of analysis is category (c) - month (t) level.

Figure D3 presents a model-free visual check for pre-treatment trends. From an inspection of Figure D3, we can verify that the short- and long-term rentals under treated and control groups indeed show similar trends over time.

Figure D3 Pre-Treatment Trends



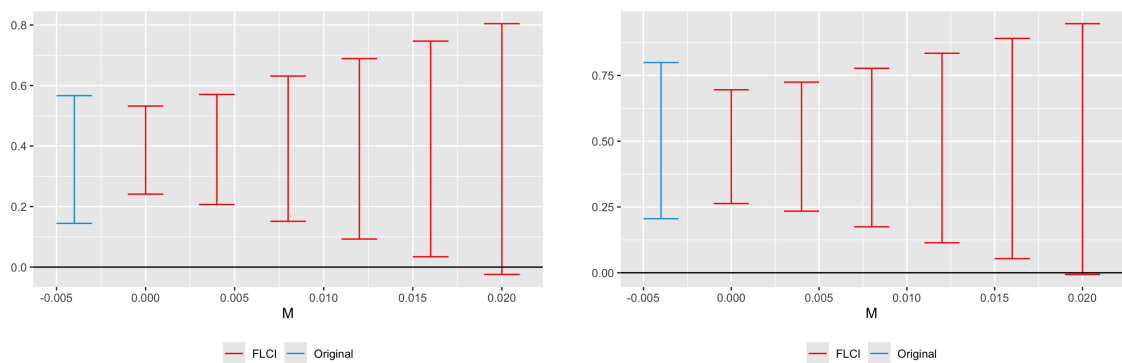
We provide the details for the sensitivity analysis using the Honest DID method proposed by [Rambachan and Roth \(2023\)](#). The approach provides ways for performing reliable causal inference in DID designs by allowing for some divergence between post-treatment trends and pre-treatment trends. Specifically, we generate confidence intervals for the re-estimated effect sizes from our main DID model by incrementally increasing the deviation from the linear trend by an arbitrary amount, M . In other words, we impose that the slope of the pre-trend can change by no more than M across consecutive periods, and imposing a smoothness restriction with $M = 0$ implies that the counterfactual difference in trends is exactly linear (larger values of M allow for more non-linearity) ([Rambachan and Roth 2024](#)). We then identify its breakdown value. We

Figure D4 Sensitivity Analysis using Honest DID Rambachan and Roth (2023): First Period Effect

(a) log(14DaysRental)

(b) log(3MonthsRental)

Note. We use 95% confidence interval. FLCI stands for fixed length confidence intervals.

Figure D5 Sensitivity Analysis using Honest DID Rambachan and Roth (2023): Overall Effect

(a) log(14DaysRental)

(b) log(3MonthsRental)

Note. We use 95% confidence interval. FLCI stands for fixed length confidence intervals.

present two figures: In Figure D4, we show the DID estimates for the first period after treatment, and in Figure D5, we show the DID estimates for the overall effect after treatment. In both figures, the bars represent the 95% confidence interval of the estimates. The confidence interval for the case of $M = 0$ shows that the estimates are significantly different from zero even when we include a treatment group-specific linear trend. The breakdown value is 0.18 and 0.24 for log(14DaysRental) and log(3MonthsRental), respectively, for the first period. These values indicate that the estimates remain significantly different from zero even when we allow for higher levels of deviations in the trends. Specifically, the first period effect under the two outcomes are significantly different from zero when we permit deviations from the pre-trend by as much as 0.18 and 0.24, respectively. Similar, the overall effect under the two outcomes are significantly different from zero when we permit deviations by as much as 0.02. These breakdown values are comparable to those reported in prior studies that have conducted similar smoothness restrictions based sensitivity analyses, namely up to 0.02 in Agarwal et al. (2024), Alpert et al. (2022), Ang (2021) and Miller et al. (2021).

Appendix E: Additional Results: Section 5

We present the results for the analyses described in §5.1.

Table E6 shows the results of Model (4). The coefficient of FirstTreatment is insignificant.

Table E6 Robustness Check: Treatment Timing and Rental Performance

	(1)	(2)
	log(14DaysRentalLag)	log(3MonthsRentalLag)
FirstTreatment	0.11 (0.13)	0.09 (0.16)
Year-Month FE	Yes	Yes
Category FE	Yes	Yes
Observations	579	579
R^2	0.477	0.521

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is category (c) - month (t) level.

Table E7 shows that none of the category features predict the timing of entry of a specific category into crowdvoting.

We present the details of our analysis for whether there is any difference in how certain the platform was as to the performance of items sent through crowdvoting relative to others. We run the following logistic regression:

$$\log \frac{P(\text{Crowdvoting}_i = 1)}{P(\text{Crowdvoting}_i = 0)} = \beta_0 + \beta_1 \text{SimilarItemCnt}_i + \lambda_t + \mu_c + \epsilon_{it}. \quad (13)$$

Here the outcome variable is Crowdvoting_i , which is 1 if the item i was sent into crowdvoting and 0 if it was picked based on staff input only. We collect all the observable features of the items in the firm's database. This includes color, occasion, style, material, and 20 other features. Then, we calculate the cosine similarity based on the 24 features, and we define similar items as those that previously existed on the platform and have a cosine similarity over 0.8. In this way, for each item that was purchased (through crowdvoting or through the staff opinion), SimilarItemCnt_i is calculated as the count of most similar items on the platform. We include the same fixed effects as in (3). If the platform has a large number of historical items similar to the focal item, the platform will have a higher level of certainty as to the performance of the focal item. Our goal is to demonstrate that the probability of entering crowdvoting is not associated with the level of certainty. Our results are reported in Table E8. If the firm had strategically entered items with more uncertainty into crowdvoting, the coefficient corresponding to SimilarItemCnt_i , β_1 , will be negative and significant. We see that β_1 is not significant.

We describe the details of our analyses following Altonji et al. (2005) next. In our setting, β_r is the coefficient from our main model, which considers only time and category-level fixed effects. For β_f , we include

Table E7 Robustness Check: Predicting Entry into Crowdvoting

	(1)
	TimeToCrowdvoting
MonthlyRental	0.00 (0.00)
RetailPrice	-0.03 (0.02)
TopBrand	-5.68 (5.10)
ColdColor	-0.23 (5.24)
SpringSeason	-12.76* (6.47)
Fitted	-1.52 (2.67)
Revealing	55.54* (30.75)
BusinessCasual	9.31* (5.35)
ForDate	-0.22 (5.16)
ForMaternity	-1.40 (2.77)
ShortSleeve	-0.81 (5.23)
AbstractPattern	7.99 (17.70)
FloralPattern	8.27 (7.90)
StripesPattern	1.44 (8.85)
WoolMaterial	-12.96* (6.88)
Observations	53
R^2	0.520

Standard errors are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is category (*c*) level.

additional controls by adding ItemCnt and BrandCnt, defined in Table C3. We compute the ratio for both our outcome variables; these are -162 and -114, respectively, for 14DaysRental and 3MonthsRental.

We describe the details of our analyses following Oster (2019) next. The fundamental idea is that including relevant control variables (those likely to improve model R^2 , such as fixed effects and controls) would help mitigate the bias caused by omitted variables. This allows researchers to assess the magnitude of the influence of selection on unobservables needed to reject the treatment effect. We employ the same set of controls used previously in the test proposed by Altonji et al. (2005) as a baseline regression. The hypothetical regression

Table E8 Verification of No Systematic Difference between Crowdvoting Items and Others

	(1) Crowdvoting
SimilarItemCnt	-0.00 (0.00)
Controls	Yes
Year-Month FE	Yes
Category FE	Yes
Observations	5564

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is item (i) level.

with both observable and unobservable variables would have a higher R^2 , denoted as R_{max} , compared to the baseline regression, but [Oster \(2019\)](#) suggest that due to measurement errors, even this ideal regression would fail to fully explain the outcome variables. Hence, we follow the recommendation and set R_{max} to 1.3 times of the R^2 of the baseline regression. We then compute the δ statistic accordingly for our two outcome variables.

We present additional analyses and results for §5.2.

For our TWFE estimator, we employ Specification (1) in §4 and our coefficient of interest is β_1 . We also conduct a leads-and-lags analysis for the TWFE model ([Autor 2003](#)). We modify Specification (1), replacing the $\text{PostCrowdvoting}_{ct}$ indicator in (1) with a variable T_{cm} . T_{cm} denotes the interval in m months from when category c first gets treated, and we consider m from -6 to 6 . 0 is the month at which category c is treated, and -1 is the baseline. This yields:

$$\log(Y_{ct}) = \beta_0 + \sum_{m=-6}^{m=6} \beta_{1m} T_{cm} + \lambda_t + \mu_c + \epsilon_{ct}. \quad (14)$$

We estimate (14) and present the results in Figure E6. Specifically, Figure E6(a) shows the estimates for the 14DaysRental_{ct} outcome, and Figure E6(b) shows the same for the $3\text{MonthsRental}_{ct}$ outcome.

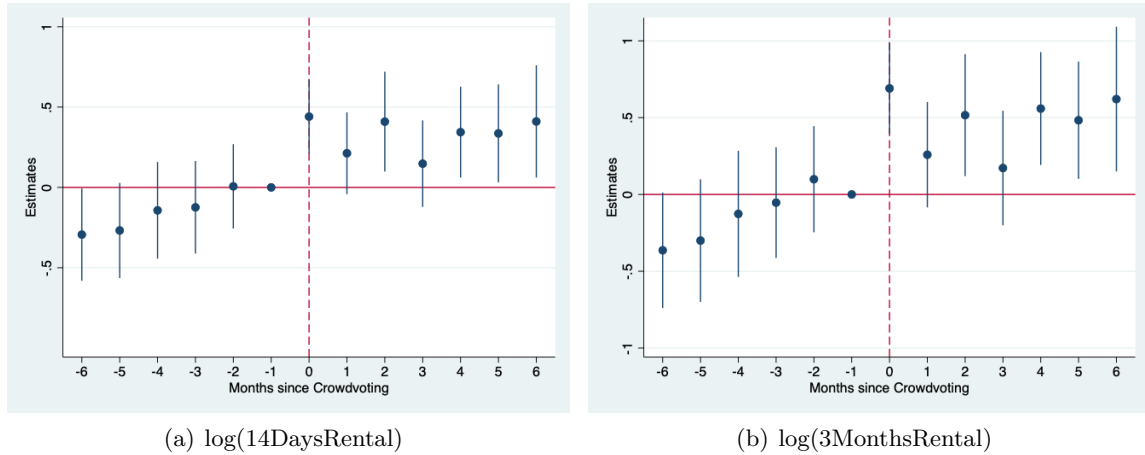
Figures E7-E8 correspond to the two additional alternate DID estimators.

Table E9 and Figure E9 show the robustness of our results when using the never-treated categories as an alternate set of controls in the [Callaway and Sant’Anna \(2021\)](#) estimator.

We next present detailed results for §5.3. Figure E10 shows the improvement in the balance between the treated and control groups after the matching for the PSM analysis.

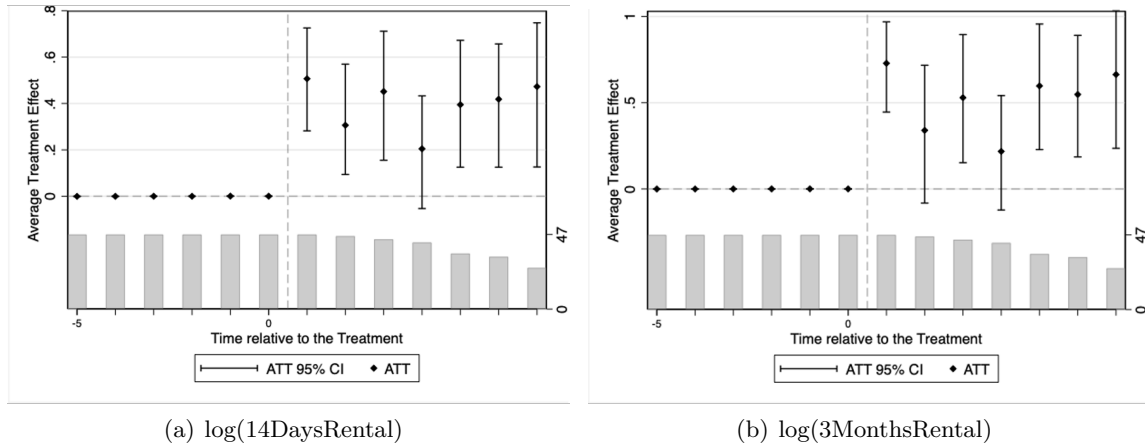
Table E10 presents our results for the impact of votes on rental performance. Among items bought through crowdvoting, those with a higher share of favorable votes have better rental outcomes. Over 90% of the items get rented within the first 14 days, that is, RentalIndicator_i is 1 for almost all the items. As a result, we do not observe a significant impact of votes on *whether* the item gets rented within 2 weeks of its launch (i.e., column (3)).

Figure E6 TWFE Leads-and-Lags Analysis



Note. We use 95% confidence interval.

Figure E7 Treatment Effect over Time using the Generalized Synthetic Controls Method



Note. The figures show the average treatment effect on treated units (ATT). We use 95% confidence interval.

Table E9 Impact of Crowd voting on Rental Performance: Using the Callaway and Sant’Anna (2021) Estimator (Never-Treated as Control Group)

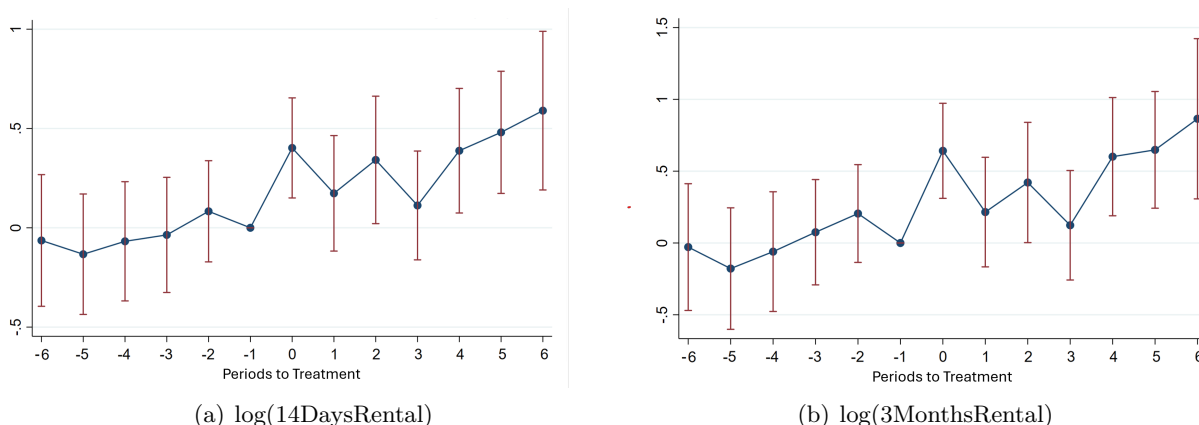
	(1)	(2)
	$\log(14\text{DaysRental})$	$\log(3\text{MonthsRental})$
Overall treatment effect	0.40***	0.57***
	(0.10)	(0.15)
Observations	827	827

Standard errors clustered at the category level are in parentheses.

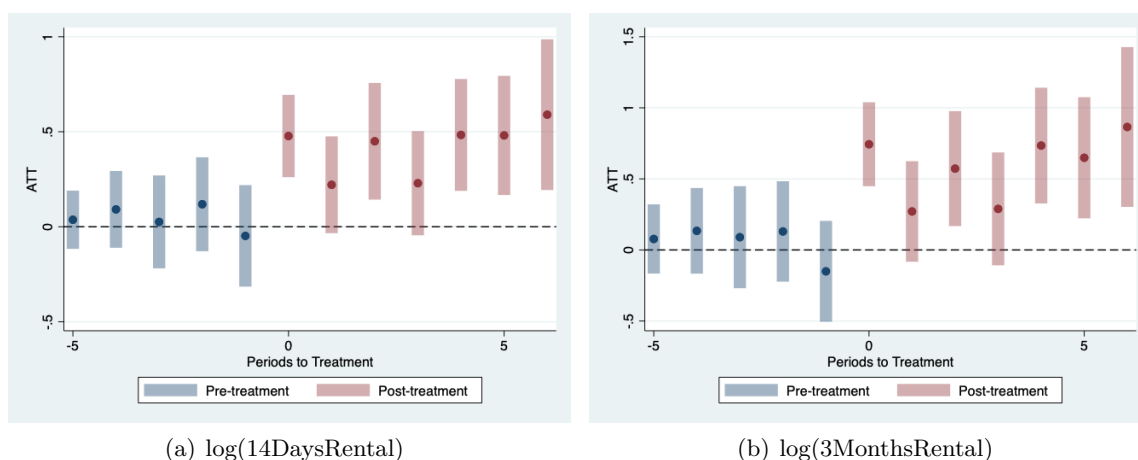
We use category and month fixed effects.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To study the truth-telling mechanism, we use the click-level data of all users on all items bought through crowdvoting, that is, all Crowd voting Buy items. Our outcome variable is $\text{RentedIndicator}_{it}$, which captures

Figure E8 Treatment Effects over Time Using the De Chaisemartin and d’Haultfoeuille (2020) Estimator

Note. We use 95% confidence interval.

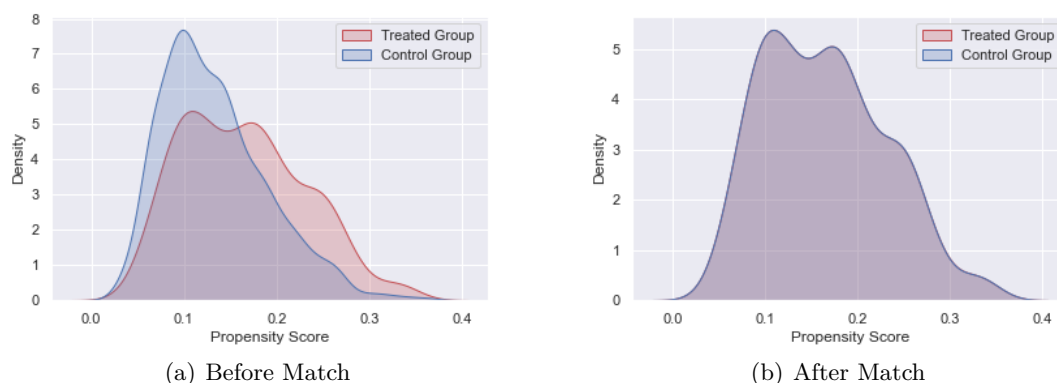
Figure E9 Treatment Effects over Time Using the Callaway and Sant’Anna (2021) Estimator (Never-Treated as Control Group)

Note. We use 95% confidence interval.

the conversion from clicks to renting. Specifically, we check if the user corresponding to click l rented the corresponding item after clicking on it once it was made available on the platform. The variables of interest are binary variables, Favorite_l and Dislike_l . Here, Favorite_l takes the value of 1 if the user voted “Favorite” on the item during the voting phase, and 0 otherwise. Similarly, Dislike_l takes the value of 1 if the user voted “Dislike” on the item during the voting phase, and 0 otherwise. These two variables explicitly capture customers’ preferences. We also add user- and item-level fixed effects through ρ_u and σ_i .

$$\text{RentedIndicator}_l = \beta_0 + \beta_1 \text{Favorite}_l + \beta_2 \text{Dislike}_l + \rho_u + \sigma_i + \epsilon_l. \quad (15)$$

Table E11 shows our results, and it verifies the claim that users’ votes are directionally aligned with their renting behavior. The positive and significant coefficient of Favorite_l suggests that upon becoming aware of crowdvoted products, voters are more likely to rent an item they voted favorably on. Conversely, the

Figure E10 PSM: Balance Improvement Between Treated and Control Groups

Note. The X-axis denotes the propensity score.

Table E10 Impact of Crowdvoting Votes on Rental Performance: Item Level Analysis

	(1)	(2)	(3)
	log(14DaysRental)	log(3MonthsRental)	RentalIndicator
HighFavoritePct	0.20** (0.08)	0.17* (0.09)	0.01 (0.02)
Controls	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes
Category FE	Yes	Yes	Yes
Observations	391	391	391
R^2	0.374	0.368	0.207

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is item (i) level.

Table E11 Verification of Truth Telling

	(1)
	RentedIndicator
Favorite	0.02** (0.01)
Dislike	-0.05*** (0.02)
User FE	Yes
ItemID FE	Yes
Observations	137867
R^2	0.163

Standard errors clustered at the user level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is click (l) level.

negative and significant coefficient of Dislike_t suggests that voters are less likely to rent an item they voted unfavorably on.

Lastly, we present additional analyses and results for §5.4. First, we present our study of the timing of treatment for complementary categories in our dataset. For each category c , we use the complementary category, \bar{c} , defined in §5.4. Recall that $|c| = 47$ (the total number of treated categories in our dataset). We then focused on the treatment timing of category pairs (c, \bar{c}) for those categories c that got treated. We noticed that only in 1 of these category pairs, c was treated before \bar{c} . (All other pairs either received treatment in the same month or c was treated after \bar{c} , so they pose minimal concern.) Even so, this pair of categories received treatment in consecutive months (October and November 2021). This leads us to conclude that the possibility of spillovers is highly unlikely in our data and provides reassurance that our control group categories could serve as valid controls. We also ran the main model after removing this category pair from our data. The results are presented in Table E12 and show that our results remain robust.

Table E12 Impact of Crowdvoting on Rental Performance: Excluding Categories with Later Treated

	Complement Categories	
	(1)	(2)
	$\log(14\text{DaysRental})$	$\log(3\text{MonthsRental})$
PostCrowdvoting	0.33*** (0.11)	0.47*** (0.16)
Year-Month FE	Yes	Yes
Category FE	Yes	Yes
Observations	801	801

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is category (c) - month (t) level.

Second, we present our modified specification to identify any spillover effect, in line with the approach in Miguel and Kremer (2004). We use the following:

$$\log(Y_{ct}) = \beta_0 + \beta_1 \text{PostCrowdvoting}_{ct} + \beta_2 N_{ct} + \beta_3 N_{ct}^T + \lambda_t + \mu_c + \epsilon_{ct}, \quad (16)$$

where c stands for category and t stands for month. N_{ct} and N_{ct}^T are defined in §5.4. Our results are presented in Table E13.

For the item-level spillover analysis, we use the following specification:

$$\log(\text{RentalCnt}_{it}) = \beta_0 + \beta_1 \text{PostCrowdvoting}_{it} + \lambda_t + \mu_c + \epsilon_{it}. \quad (17)$$

Our variable of interest is the binary variable PostCrowdvoting , which is set to 1 for the month after crowdvoted item i becomes available on the platform. The results are shown in Table E14, and we observe that the coefficient of PostCrowdvoting is not significant.

Table E13 Impact of Crowdvoting on Rental Performance: Spillovers

	(1)	(2)
	log(14DaysRental)	log(3MonthsRental)
PostCrowdvoting	0.41*** (0.08)	0.52*** (0.11)
N_{ct}	0.01*** (0.00)	0.01*** (0.00)
N_{ct}^T	-0.01 (0.01)	-0.01 (0.01)
Year-Month FE	Yes	Yes
Category FE	Yes	Yes
Observations	827	827
R^2	0.515	0.557

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is category (c) - month (t) level.

Table E14 Impact of Crowdvoting on Rental Performance: Spillovers at Item Level

	(1)
	log(RentalCnt)
PostCrowdvoting	-0.02 (0.01)
Year-Month FE	Yes
Category FE	Yes
Observations	10516
R^2	0.049

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is item (i) - month (t) level.

Appendix F: Additional Results: Section 6

We present the results for the analyses described in §6.1. Table F15 shows the impact on inventory turns.

Table F15 Impact on Inventory Turns

	(1)
	InventoryTurns
CrowdvotingBuy	0.19** (0.08)
Controls	Yes
Year-Month FE	Yes
Category FE	Yes
Observations	2828
R^2	0.091

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is item (i) level.

We next report our analysis comparing purchase price. We modify (3) and use the logarithm of the price paid by the platform to purchase the item as the outcome variable.

Table F16 Verification of No Systematic Price Difference between Crowdvoting Items and Others

	(1)
	$\log(\text{PurchasePrice})$
CrowdvotingBuy	0.06 (0.04)
Controls	Yes
Year-Month FE	Yes
Category FE	Yes
Observations	2828
R^2	0.178

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is item (i) level.

$$\log(\text{PurchasePrice}_i) = \beta_0 + \beta_1 \text{CrowdvotingBuy}_i + \beta_2 \text{Controls}_i + \lambda_t + \mu_c + \epsilon_i. \quad (18)$$

Table F16 shows our results, and it verifies the claim in §6.1 that there is no systematic difference in the per-unit purchase price of items bought post crowdvoting compared to items bought based solely on staff opinion.

We next report our analysis comparing users' rental count. Our outcome variable is number of items rented by user u in month t . The variable of interest is $\text{CrowdvotingStart}_u$, which indicates the start date

of crowdvoting. We control for users' tenure and number of active days in month t , and add user-level fixed effects μ_u .

$$\text{MonthlyRentalCnt}_{ut} = \beta_0 + \beta_1 \text{CrowdvotingStart}_u + \beta_2 \text{Tenure}_u + \beta_3 \text{ActiveDayCnt}_{ut} + \mu_u + \epsilon_u. \quad (19)$$

Table F17 shows our results and verifies the claim in §6.1 that the number of rentals per user is unchanged.

Table F17 **Verification of No Systematic Difference in Users' Monthly Rental Count Before and After Crowd voting**

	(1)
	MonthlyRentalCnt
CrowdvotingStart	0.17 (0.12)
Tenure	-2.45*** (0.19)
ActiveDayCnt	0.54*** (0.02)
User FE	Yes
Observations	18276
R^2	0.703

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is user (u) - month (t) level.

Table F18 reports our findings comparing items' brand type.

Table F18 **Impact on Brand Mix**

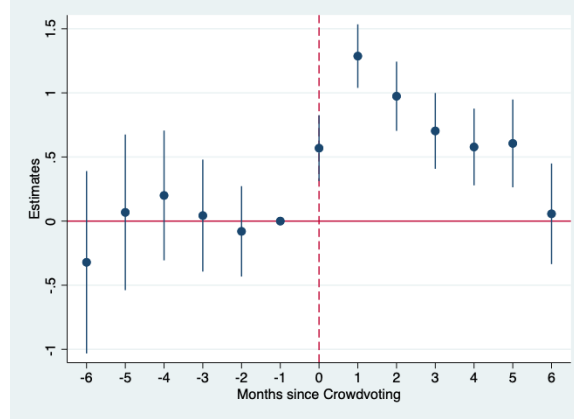
	(1)
	CrowdvotingBuy
NewBrand	0.91*** (0.27)
Controls	Yes
Year-Month FE	Yes
Category FE	Yes
Observations	2770

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is item (i) level.

We report on the analyses considered in §6.2. For our DID framework in §6.2 to be valid, there should not be a significant difference between the control and the treated groups, that is, there should be no selection on engagement between users who participate in crowdvoting earlier compared to the users who participate

Figure F11 Impact on Voter Engagement: Leads-and-Lags Analysis

Note. We use 95% confidence interval.

at a later time. We conduct a leads-and-lags analysis to verify this assumption. We modify (8) and replace the variable $\text{PostParticipation}_{ut}$ by T_{um} as follows:

$$\text{ReviewCnt}_{ut} = \beta_0 + \sum_{m=-6}^{m=6} \beta_{1m} T_{um} + \beta_2 \text{ActiveDayCnt}_{ut} + \lambda_t + \mu_u + \epsilon_{ut}. \quad (20)$$

Here T_{um} refers to the interval in m months from when user u receives treatment, that is, begins to participate in crowdvoting. As before, 0 is the month at which the user u is treated and we fix time period corresponding to -1 as the baseline in our estimation. We show the results in Figure F11. We observe that there is no significant difference in the reviewing behavior among users who start participating in crowdvoting earlier and those who participate in crowdvoting at a later stage, thus allowing us to conclude that we fail to falsify the parallel trends assumption.

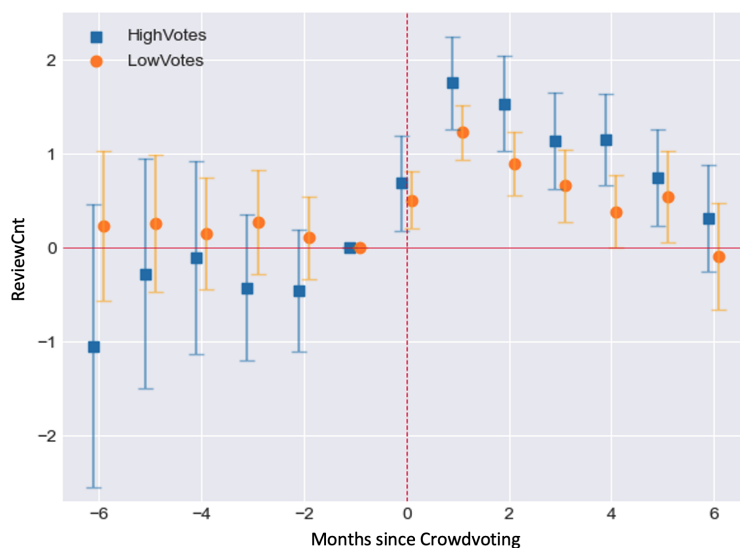
Figure F12 shows the impact on engagement for two classes of voters: HighVotes denotes those voters who cast more than the median number of votes and LowVotes denotes those voters who cast fewer than the median number of votes. We see that there is an increase in the reviewing behavior of voters who are more active than others.

We report on the analyses considered in §6.4. We start with the rating analysis. The variable of interest is CrowdvotingBuy_i , which indicates whether the item is purchased through crowdvoting. We control for 13 item-level features, as in Specification (3) in our paper and include user- and category-level fixed effects. This yields the following specification:

$$\text{HighRating}_r = \beta_0 + \beta_1 \text{CrowdvotingBuy}_i + \beta_2 \text{Controls}_i + \lambda_u + \mu_c + \epsilon_r. \quad (21)$$

Table F20 reports results for the counterfactual analysis considering the rental performance of similar items.

Table F21 reports our estimates from re-running the main specifications considering those categories containing items that never had an item in crowdvoting as the control group.

Figure F12 Voter Engagement over Time: the Impact of Votes**Table F19 Impact of Crowdvoting on Star Rating: Rental Level Analysis**

	(1)
	HighRating
CrowdvotingBuy	0.02**
	(0.01)
Controls	Yes
User FE	Yes
Category FE	Yes
Observations	90639
R^2	0.242

Standard errors clustered at the user level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is rental (r) level.

We next present our analysis into the billboard effect. For all voting users and items bought after going through crowdvoting, we define $\text{RentedIndicator}_{ui}$ to indicate whether we observe a rental of item i from user u . We then see if their voting behavior – whether user u voted “Favorite” on item i denoted by Favorite_{ui} and whether user u voted “Dislike” on item i denoted by Dislike_{ui} – predicts their chance of renting items (while controlling for user- and item-level fixed effects). The specification is below:

$$\text{RentedIndicator}_{ui} = \beta_0 + \beta_1 \text{Favorite}_{ui} + \beta_2 \text{Dislike}_{ui} + \lambda_i + \mu_u + \epsilon_{ui}. \quad (22)$$

Our results are shown in Table F22.

We next report the results by re-running our main specifications after excluding all the rentals from voting users who voted “Favorite” on purchased items. Since other users did not have this potential influence, any effect on rentals can now be attributed to crowdvoting alone. We report our results in Table F23. The coefficients of PostCrowdvoting are positive and significant.

Table F20 Counterfactual Analysis on the Role of Vote Count on Rentals

	(1)	(2)
	log(14DaysRental)	log(3MonthsRental)
FavoriteVotesCnt	0.0011** (0.04)	0.0012** (0.06)
TotalVotesCnt	-0.04 (0.02)	-0.03 (0.03)
log(RetailPrice)	-0.32*** (0.02)	-0.49*** (0.03)
Observations	2652	2652
R^2	0.076	0.095

Standard errors are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The unit of analysis is item (i) level.

Table F21 Impact of Crowdvoting on Rental Performance: Robustness to Not Voted as Control Group

	(1)	(2)
	log(14DaysRental)	log(3MonthsRental)
PostCrowdvoting	0.29** (0.12)	0.40** (0.16)
Year-Month FE	Yes	Yes
Category FE	Yes	Yes
Observations	699	699

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is category (c) - month (t) level.

Table F22 Impact of Crowdvoting on Rental Performance: Billboard Effect

	(1)
	RentedIndicator
Favorite	0.01*** (0.00)
Dislike	-0.01*** (0.00)
User FE	Yes
Item FE	Yes
Observations	1177435
R^2	0.017

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is user (u) - item (i) level.

Table F23 Impact of Crowdvoting on Rental Performance: Excluding Voting Users' Rentals

	(1)	(2)
	log(14DaysRental)	log(3MonthsRental)
PostCrowdvoting	0.24**	0.37**
	(0.11)	(0.15)
Year-Month FE	Yes	Yes
Category FE	Yes	Yes
Observations	827	827

Standard errors clustered at the category level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is category (c) - month (t) level.

Appendix G: Additional Results: Section 7

We present evidence of voter fatigue. We collected the days that each voter in our dataset voted. For each user u , we define VotingGap_{uv} as the number of days between the current vote v and the subsequent vote, and we define $\text{TimePostCrowdvoting}_u$ as the number of days since the user u cast their first vote. By regressing the gap between the voting days on the number of days since they started participating in crowdvoting through the specification in (23), we can provide evidence of fatigue (if any). We add user-level fixed effects μ_u in (23) to account for unobserved time-invariant factors among users.

$$\text{VotingGap}_{uv} = \beta_0 + \beta_1 \text{TimePostCrowdvoting}_v + \mu_u + \epsilon_{uv}. \quad (23)$$

Our results are shown in Table G24 and reveal that there is fatigue: The coefficient of $\text{TimePostCrowdvoting}_u$ is positive and significant showing that the number of days between votes progressively increases as the time since first vote increases. In other words, voters do decrease the frequency of active voting days over time.

Table G24 Evidence of Voter Fatigue

	(1) VotingGap
TimePostCrowdvoting	0.02*** (0.00)
User FE	Yes
Observations	36478
R^2	0.438

Standard errors clustered at the user level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is user (u) - vote (v) level.

To evaluate whether ignoring voters' stated preferences leads to dissatisfaction, we have conducted the following new analysis. As in the main paper, we consider the ReviewCnt_{ut} as the dependent variable: a lower value then serves as a proxy for voters' disengagement. Our main variable of interest is the number of times the platform ignores voters' stated preferences; for each voting user u and a given month, we add up two quantities: (1) the number of items for which the user voted "Favorite" but the platform did not buy the item and (2) the number of items for which the user voted "Dislike" but the platform bought the item anyway. Define $\text{CumulativeIgnoredVote}_{ut}$ as the cumulative sum of this variable up until month t . We also control for the total number of votes, $\text{CumulativeVote}_{ut}$, cast by each voting user u up until month t . If the coefficient of $\text{CumulativeIgnoredVote}_{ut}$ were to be negative and significant, we could interpret it as evidence of the aforementioned dissatisfaction. The model is specified below:

$$\text{ReviewCnt}_{ut} = \beta_0 + \beta_1 \text{CumulativeVote}_{ut} + \beta_2 \text{CumulativeIgnoredVote}_{ut} + \lambda_t + \mu_u + \epsilon_{ut}. \quad (24)$$

Our results are presented in Table G25 and show that ignoring the voters' preferences over time could indeed lead to disengagement.

Table G25 Impact on User Engagement: Evidence of Voters' Dissatisfaction

	(1)
	ReviewCnt
CumulativeVote	0.0033** (0.0015)
CumulativeIgnoredVote	-0.0062* (0.0034)
Year-Month FE	Yes
User FE	Yes
Observations	31298
R^2	0.662

Standard errors clustered at the user level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The unit of analysis is user (u) - month (t) level.