What motivates experts to contribute to public information goods? A field experiment at Wikipedia

Yan Chen* Rosta Farzan⁎ Robert Kraut#
Iman YeckehZaare* Ark Fangzhou Zhang*

*School of Information, University of Michigan
⁎School of Information Sciences, University of Pittsburgh
#School of Computer Science, Carnegie Mellon University

January 2018

User-Generated Content as Public Information Good

- User-generated content
  - Online reviews: Amazon, Yelp
  - Internet encyclopedia: Wikipedia
  - Online health support networks: ACS Cancer Support Network

User-Generated Content as Public Information Good

- User-generated content
  - Online reviews: Amazon, Yelp
  - Internet encyclopedia: Wikipedia
  - Online health support networks: ACS Cancer Support Network
- Public information goods
  - Non-rivalrous
  - Non-excludable (by choice)
  - Expertise matters: inputs are not perfect substitutes
    - quality
    - marginal cost
    - (affect)

Research Questions

Research Questions

- What motivates experts to contribute to public information goods?
  - Voluntary contribution to public goods
  - Free-riding problem

Research Questions

- What motivates experts to contribute to public information goods?
  - Voluntary contribution to public goods
  - Free-riding problem
- How motivating is social impact?
  - Number of recipients (Andreoni, 2006 & 2007)
  - 40% decrease after exogenous reduction in readership in Chinese Wikipedia (Zhang and Zhu, 2011)

Research Questions

- What motivates experts to contribute to public information goods?
  - Voluntary contribution to public goods
  - Free-riding problem
- How motivating is social impact?
  - Number of recipients (Andreoni, 2006 & 2007)
  - 40% decrease after exogenous reduction in readership in Chinese Wikipedia (Zhang and Zhu, 2011)
- How motivating are private benefits?

Introduction: Wikipedia

- Wikipedia: among top-5 most visited sites in the world
  - 5.4 million articles in the English Wikipedia
  - 50,000 high quality articles (March, 2017)
  - More than 500 million unique visitors each month

Introduction: Wikipedia

- Wikipedia: among top-5 most visited sites in the world
  - 5.4 million articles in the English Wikipedia
  - 50,000 high quality articles (March, 2017)
  - More than 500 million unique visitors each month
- Who contributes to Wikipedia?
  - Wikipedians: individual contributors
  - Active contributors mostly are non-experts (Lih 2009)
  - Experts seldom make contributions (YeckehZaare 2015)

Introduction: Wikipedia

- Wikipedia: among top-5 most visited sites in the world
  - 5.4 million articles in the English Wikipedia
  - 50,000 high quality articles (March, 2017)
  - More than 500 million unique visitors each month
- Who contributes to Wikipedia?
  - Wikipedians: individual contributors
  - Active contributors mostly are non-experts (Lih 2009)
  - Experts seldom make contributions (YeckehZaare 2015)
- Holes in Wikipedia
  - Science: imprecise, erroneous, incomplete
  - Women: sparse

Introduction: Wikipedia

- Wikipedia: among top-5 most visited sites in the world
  - 5.4 million articles in the English Wikipedia
  - 50,000 high quality articles (March, 2017)
  - More than 500 million unique visitors each month
- Who contributes to Wikipedia?
  - Wikipedians: individual contributors
  - Active contributors mostly are non-experts (Lih 2009)
  - Experts seldom make contributions (YeckehZaare 2015)
- Holes in Wikipedia
  - Science: imprecise, erroneous, incomplete
  - Women: sparse
- 2016: Wikipedia Year of Science

Introduction: Wikipedia

- Wikipedia: among top-5 most visited sites in the world
  - 5.4 million articles in the English Wikipedia
  - 50,000 high quality articles (March, 2017)
  - More than 500 million unique visitors each month
- Who contributes to Wikipedia?
  - Wikipedians: individual contributors
  - Active contributors mostly are non-experts (Lih 2009)
  - Experts seldom make contributions (YeckehZaare 2015)
- Holes in Wikipedia
  - Science: imprecise, erroneous, incomplete
  - Women: sparse
- 2016: Wikipedia Year of Science
- How do we motivate domain experts (scientists, etc.) to contribute?

Example: Instrumental Variable

> *"..., the method of instrumental variables (IV) is used to estimate causal relationships when controlled experiments are not feasible ..."*

Instrumental variable

From Wikipedia, the free encyclopedia

In statistics, econometrics, epidemiology and related disciplines, the method of **instrumental variables** (**IV**) is used to estimate causal relationships when controlled experiments are not feasible or when a treatment is not successfully delivered to every unit in a randomized experiment.[1] Intuitively, IV is used when the correlation between the explanatory variable and the dependent variable does not plausibly reflect the causal relationship between the two. A valid instrument induces changes in the explanatory variable but has no independent effect on the dependent variable, allowing a researcher to uncover the causal effect of the explanatory variable on the dependent variable.

Instrumental variable methods allow for consistent estimation when the explanatory variables (covariates) are correlated with the error terms in a regression model. Such correlation may occur when changes in the dependent variable change the value of at least one of the covariates ("reverse" causation), when there are omitted variables that affect both the dependent and independent variables, or when the covariates are subject to measurement error. Explanatory variables which suffer from one or more of these issues in the context of a regression are sometimes referred to as endogenous. In this situation, ordinary least squares produces biased and inconsistent estimates.[2] However, if an *instrument* is available, consistent estimates may still be obtained. An instrument is a variable that does not itself belong in the explanatory equation but is correlated with the endogenous explanatory variables, conditional on the value of other covariates. In linear models, there are two main requirements for using IV:

- The instrument must be correlated with the endogenous explanatory variables, conditional on the other covariates. If this correlation is highly statistically significant, then the instrument is said to have a **strong first stage**. A weak correlation may provide misleading inferences about parameter estimates and standard errors.[3]
- The instrument cannot be correlated with the error term in the explanatory equation, conditional on the other covariates. In other words, the instrument cannot suffer from the same problem as the original predicting variable. If this condition is met, then the instrument is said to satisfy the **exclusion restriction**.

Contents [hide]

Literature

- Laboratory and field experiments on public goods
  - Ledyard (1995)
  - Vesterlund (2016): charitable giving

Literature

- Laboratory and field experiments on public goods
  - Ledyard (1995)
  - Vesterlund (2016): charitable giving
- What motivates Wikipedians (insiders)?
  - Reciprocity, social image (Algan et al. 2013)
  - Symbolic awards (Gallus 2016)
  - Better matching and lower cost (Cosley et al. 2007)

Literature

- Laboratory and field experiments on public goods
  - Ledyard (1995)
  - Vesterlund (2016): charitable giving
- What motivates Wikipedians (insiders)?
  - Reciprocity, social image (Algan et al. 2013)
  - Symbolic awards (Gallus 2016)
  - Better matching and lower cost (Cosley et al. 2007)
- What motivates domain experts (outsiders)?
  - Taraborelli, Mietchen, Alevizou and Gill (2011)

Experimental Design: 2×3 factorial design

- Social impact
  1. Average view: # of views of a typical WP article (426)
  2. High view: # of views of the recommended articles ($> 1,000$)

Experimental Design: 2×3 factorial design

- Social impact
  1. Average view: # of views of a typical WP article (426)
  2. High view: # of views of the recommended articles ($> 1,000$)
- Private benefits
  1. No Cite: no citation benefit mentioned
  2. Citation:
     - might cite your work
     - may include include some of your publications in their references
     - might refer to some of your research
  3. Citation & acknowledgement:
     - citation
     - acknowledge your contributions publicly

Experimental Design: 2×3 factorial design

	No Citation	Citation	Citation & Acknowledge
Average View	AvgView-NoCite ($n = 678$)	AvgView-Cite ($n = 669$)	AvgView-CiteAcknowledge ($n = 672$)
High View	HighView-NoCite ($n = 636$)	HighView-Cite ($n = 661$)	HighView-CiteAcknowledge ($n = 658$)

Total number of participants:

- Intent to treat: $n = 3,974$
- Treated group: $n = 3,288$

Domain experts in this experiment: Academic economists

- Participant information retrieved from RePEc: `https://ideas.repec.org`

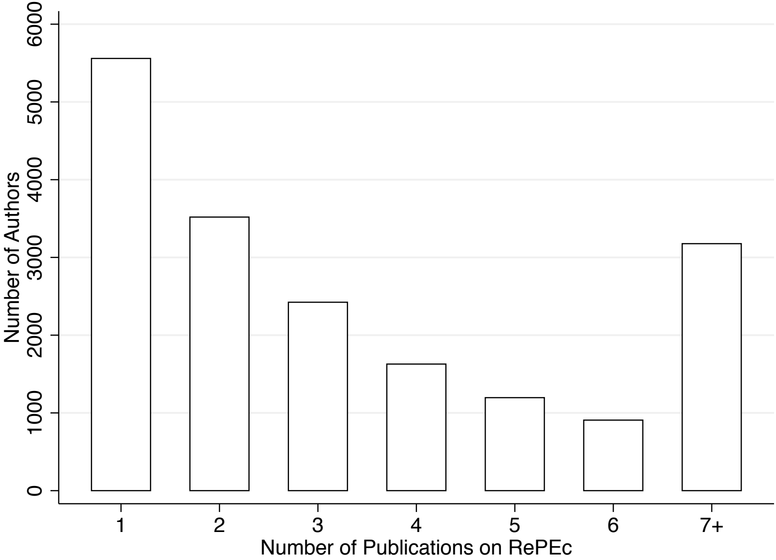
Domain experts in this experiment: Academic economists

- Participant information retrieved from RePEc: `https://ideas.repec.org`
- Why RePEc?
  - Data use policy: `http://repec.org/docs/RePEcDataUse.html`
  - Paper archive: matching experts with WP articles
  - Self-identified areas of specialization (identity)
  - RePEc ranking

Domain experts in this experiment: Academic economists

- Participant information retrieved from RePEc: `https://ideas.repec.org`
- Why RePEc?
  - Data use policy: `http://repec.org/docs/RePEcDataUse.html`
  - Paper archive: matching experts with WP articles
  - Self-identified areas of specialization (identity)
  - RePEc ranking
- Expert selection
  - Post at least six articles in English: 3,974
  - Accuracy of recommender system

Expert selection: Distribution of # of publications on RePEc



Wikipedia article selection

- Under namespace 0 (Main/Article)
- Not edit protected
- Not a “stub”
- At least 1,500 characters
- Viewed at least 1,000 times in the past 30 days (dynamically updated)

Implementation: Three-phase design

- Phase 1
  - Send personalized email invitations to experts
  - Treatments implemented
- Phase 2
  - Recommend relevant articles to interested experts
  - Articles selected to match experts' recent work
- Phase 3
  - Send thank-you email
  - Links to posted comments on Talk Page
  - Links to tutorial on editing Wikipedia articles

Phase 1: Personalized email

Dear Dr. Chen,

Would you be willing to spend 10 - 20 minutes providing feedback on a few Wikipedia articles related to behavioral and experimental economics? Wikipedia is among the most important information sources the general public uses to find out about a wide range of topics. A Wikipedia article is viewed on average 426 times each month. While many Wikipedia articles are useful, articles written by enthusiasts instead of experts can be inaccurate, incomplete, or out of date.

If you are willing to help, we will send you links to a few Wikipedia articles in your area of expertise. We will select only articles, with over 1,000 views in the past month, so that your feedback will benefit many Wikipedia readers.

These articles may include some of your publications in their references.

Please click one of the following links to continue:

Yes, please send me some Wikipedia articles to comment on.

No, I am not interested.

Thank you for your attention.

Sincerely,

Yan Chen, Daniel Kahneman Collegiate Professor of Information, University of Michigan

Robert Kraut, Herbert A. Simon Professor of Human-Computer Interaction, Carnegie Mellon University

Phase 2: Recommending relevant articles

Dear Dr. Bebchuk,

Thank you for your willingness to provide feedback on the quality of Wikipedia articles. The following articles are suggested by our algorithm as related to law & economics.

Please comment on the articles most relevant to your research. Your feedback can significantly improve these articles' accuracy and completeness, and the comments and the references that you provide will be incorporated therein. These articles might refer to some of your research. We would appreciate receiving your comments by Jan 14, 2017. Thank you very much for your help.

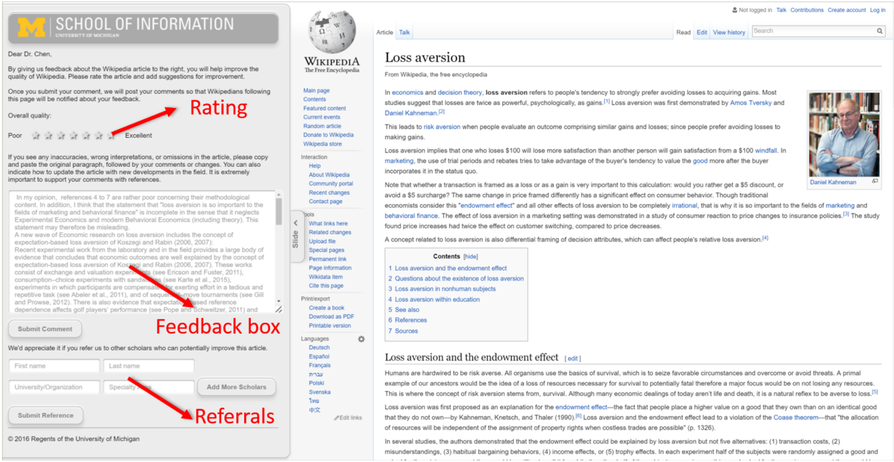
Wikipedia Article Title	Number of views in the past month	Link to review the article
Shareholder value	6,298	Click here
Corporate governance	38,351	Click here
Managerial economics	17,771	Click here
Economic nationalism	8,931	Click here
University of Delaware	17,123	Click here
Corporatocracy	10,479	Click here

Sincerely,

Yan Chen, Daniel Kahneman Collegiate Professor of Information, University of Michigan

Robert Kraut, Herbert A. Simon Professor of Human-Computer Interaction, Carnegie Mellon University

Phase 2: Interface design - lowering entry cost



- Lower entry barrier: no need to learn how to edit wiki
- Separate expert's comments from incorporation into WP article

Email Sending Procedure

- Emails sent 6:00 AM – 7:00 PM of expert's local time (based on location of primary affiliation)
- System tracks if each expert opens email
  - 84% opened first email (treated group)
- Responses:
  - Yes: phase 2 email sent immediately
  - No: dropped
  - No response after 2 weeks: 4 reminders
- Comments: manually verified before posting to article Talk page

What happened to these comments?

- ExpertIdeas Bot
  - Post comments on article talk page
  - Alert Wikipedia editors who watch this page
- Three scenarios
  - Best case: editors incorporate these comments
  - Intermediate case: editors comment on the comments
  - Worst case: nothing happens
- Students working with Wiki Ed to incorporate these comments
  - SI 563 (Game Theory)
  - 100% edits stayed after 4 months

Theory

- Public good: $y > 0$
- Number of consumers of this public good: $n \geq 0$
- Contribution level, a, from a choice set, $A \in [0, \bar{a})$
- Cost function, $c(a)$, is convex
- Social impact of public goods: $v(n)(y + ay)$
- Private benefit from contributions: $w(n)a$

$$\max_{a \in A} v(n)(y + ay) + w(n)a + \gamma(A - a) - \frac{c(a)}{s}. \tag{1}$$

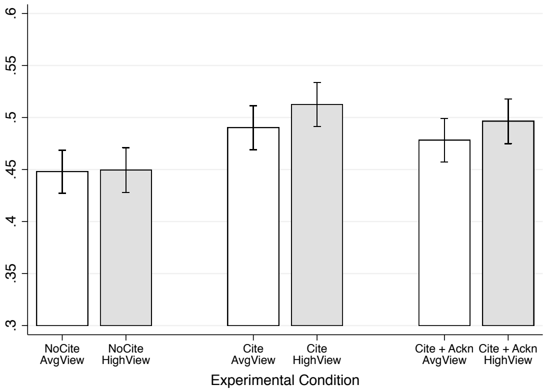
Assuming $c(a) = ca^2/2$, we obtain optimal contribution level:

$$a^* = [v(n)y + w(n) - \gamma]\frac{s}{c}, \tag{2}$$

Hypotheses

- Experts will be more interested to contribute when citation benefit is made salient: $\frac{\partial a^*}{\partial w} = \frac{s}{c} > 0$.
- Experts will be more interested to contribute with increasing # of views: $\frac{\partial a^*}{\partial n} = [v'(n)y + w'(n)]\frac{s}{c} > 0$.
- An expert with a higher reputation will contribute less: $\frac{\partial a^*}{\partial \gamma} = -\frac{s}{c} < 0$.
- Better matching between the content of the public information good and the agent's expertise leads to an increased level of contributions, i.e., $\frac{\partial a^*}{\partial s} = [v(n)y + w(n) - \gamma]/c \geq 0$ if and only if $v(n)y + w(n) \geq \gamma$.

Phase 1: Treatment effects on positive response



- NoCite & AvgView (baseline: 45%): high compared to APS campaign
- High View by itself: positive but insignificant effect
- Citation & High View: the highest positive response rate

Treatment effects: Average marginal effects of multinomial logistic regression on participation

	Positive	No response	Negative
HighView	0.002	0.021	-0.022
	(0.030)	(0.026)	(0.027)
Citation	0.040	0.022	-0.064**
	(0.030)	(0.026)	(0.027)
Acknowledgment	0.030	0.019	-0.050*
((Interaction terms snipped))	(0.029)	(0.026)	(0.027)
HighView+HighView×Citation	0.022	-0.002	-0.020
	(0.030)	(0.026)	(0.025)
Citation+HighView×Citation	0.063**	-0.001	-0.062**
	(0.030)	(0.027)	(0.026)
HighView+HighView×Acknowledgement	0.018	0.017	-0.036
	(0.030)	(0.027)	(0.026)
Acknowledgement+HighView×Acknowledgement	0.047	0.016	-0.063**
	(0.030)	(0.027)	(0.027)

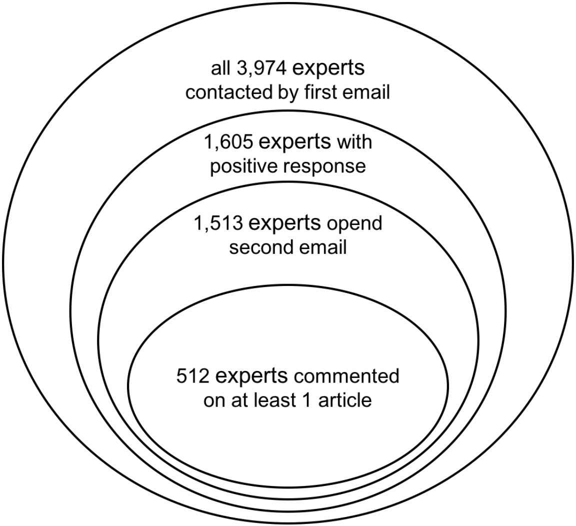
1. Citation at HighView increases positive response by 6 p.p.;
2. Citation decreases negative response by 6 p.p. at both views;
3. Acknowledgement at HighView decreases negative response by 6 p.p.

Reputation and social distance

	Positive Response	No Response	Negative Response	Positive Response	No Response	Negative Response
HighView	0.002	0.021	-0.022	0.004	0.019	-0.023
	(0.030)	(0.026)	(0.027)	(0.030)	(0.026)	(0.027)
Citation	0.040	0.022	-0.064**	0.038	0.026	-0.064**
	(0.030)	(0.026)	(0.027)	(0.030)	(0.026)	(0.026)
Acknowledgment	0.030	0.019	-0.050*	0.020	0.024	-0.045*
((Interaction terms snipped))	(0.029)	(0.026)	(0.027)	(0.030)	(0.026)	(0.027)
Author Abstract Views				0.033	-0.417**	0.384***
				(0.188)	(0.192)	(0.145)
English Affiliation				-0.017	-0.043***	0.060***
				(0.018)	(0.015)	0.015
Behavioral & experimental econ.				0.210***	-0.075***	-0.134***
				(0.034)	(0.028)	(0.025)

1. **Reputation**: A 1,000-view increase in the number of author abstract views is associated with a 0.83 p.p. increase in the likelihood of a negative response. ABV normalized to [0, 1] from [51, 46,057].
2. **Social distance**: Behavioral and experimental economists are 21 (13.5) p.p. more (less) likely to respond positively (negatively) than others.

Samples through Phases 1 and 2



Phase 2: Contribution Quantity

- 1,513 (94%) opened phase-2 email
- 512 (34%) commented on at least one WP article
- 1,190 comments received by November 30, 2016
- Large variance in quantity (word count)
  - Some wrote one-line comments
  - Some rewrote the entire article

Table: Participants' responses in Phase 2, by experimental conditions

	AvgView NoCite (1)	AvgView Cite (2)	AvgView CiteAckn. (3)	HighView NoCite (4)	HighView Cite (5)	HighView CiteAckn. (6)
Comment at least 1 article	0.331	0.314	0.335	0.363	0.316	0.376
	(0.471)	(0.465)	(0.473)	(0.482)	(0.466)	(0.485)
Number of articles commented	0.884	0.783	0.708	0.843	0.665	0.849
	(1.658)	(1.492)	(1.295)	(1.451)	(1.263)	(1.432)
Average word count	44	41	65	96	43	60
	(177)	(160)	(219)	(600)	(131)	(160)
Observations	242	258	257	223	275	258

Example 1: Traveler's Dilemma

- Original:
  "When the game is played experimentally, most participants select a value close to $100."

Example 1: Traveler's Dilemma

- Original:
  "When the game is played experimentally, most participants select a value close to \$100."
- Proposed change:
  "When the game is played experimentally, most participants select a value higher than the Nash equilibrium and closer to \$100. More precisely, the Nash equilibrium strategy solution proved to be a bad predictor of people's behaviour in a TD with small bonus/malus and a rather good predictor if the bonus/malus parameter was big."

Example 1: Traveler's Dilemma

- Original:
  "When the game is played experimentally, most participants select a value close to \$100."
- Proposed change:
  "When the game is played experimentally, most participants select a value higher than the Nash equilibrium and closer to \$100. More precisely, the Nash equilibrium strategy solution proved to be a bad predictor of people's behaviour in a TD with small bonus/malus and a rather good predictor if the bonus/malus parameter was big."
- Expert: Piergiuseppe Morone, Professor of Economic Policy at University of Rome

Example 1: Traveler's Dilemma

- Original:
  "When the game is played experimentally, most participants select a value close to \$100."
- Proposed change:
  "When the game is played experimentally, most participants select a value higher than the Nash equilibrium and closer to \$100. More precisely, the Nash equilibrium strategy solution proved to be a bad predictor of people's behaviour in a TD with small bonus/malus and a rather good predictor if the bonus/malus parameter was big."
- Expert: Piergiuseppe Morone, Professor of Economic Policy at University of Rome
- Expertise inferred from this paper:
  Morone, A., P. Morone and A.R. Germani. "Individual and group behaviour in the traveler's dilemma: An experimental study." *JEBO*, 2014.

Example 2: Repeated Game

- Article
  https://en.wikipedia.org/wiki/Repeated_game
- Talk Page
  https://en.wikipedia.org/wiki/Talk:Repeated_game
- Consent obtained from Oleg Korenok and Karl Schlag

Cosine similarity

- Cosine similarity of two documents measure the similarity between them in terms of overlapping vocabulary
    1. Doc 1: Expert's abstract, a
    2. Doc 2: Wikipedia article, b

Cosine similarity

- Cosine similarity of two documents measure the similarity between them in terms of overlapping vocabulary
  1. Doc 1: Expert's abstract, a
  2. Doc 2: Wikipedia article, b
- Construct two vectors, **A** and **B**
  - enter both text files into a tokenizer, which divides text into a sequence of tokens, which roughly correspond to "words"
  - results processed by a stemmer, which reduces inflected or derived words to their word stem, base or root form
  - results passed to a tf-idf vectorizer (term frequency–inverse document frequency)

Cosine similarity

- Cosine similarity of two documents measure the similarity between them in terms of overlapping vocabulary
  1. Doc 1: Expert's abstract, a
  2. Doc 2: Wikipedia article, b
- Construct two vectors, **A** and **B**
  - enter both text files into a tokenizer, which divides text into a sequence of tokens, which roughly correspond to "words"
  - results processed by a stemmer, which reduces inflected or derived words to their word stem, base or root form
  - results passed to a tf-idf vectorizer (term frequency–inverse document frequency)
- Calculate cosine similarity between **A** and **B**:

$$\cos(\theta) = \frac{\mathbf{A}^T \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

Contribution quantity: Compound Poisson Linear Model

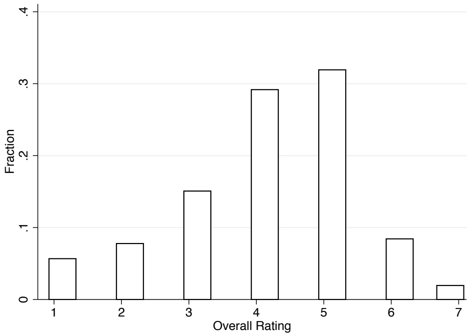
Dependent Variable	log(Word Count)		
HighView	0.165	0.109	0.068
	(0.275)	(0.282)	(0.281)
Citation	0.017	-0.025	-0.058
	(0.270)	(0.277)	(0.275)
Acknowledgement	0.152	0.154	0.086
	(0.267)	(0.273)	(0.275)
Cosine Similarity		2.138***	2.219***
		(0.459)	(0.461)
log(Page Length)		-0.012	-0.017
		(0.079)	(0.080)
Author Abstract View			0.729
			(1.545)
English Affiliation			0.148
			(0.156)
Behavioral & Experimental Econ.			0.619**
			(0.235)
N	8,825	8,659	8,559

Contribution quantity

- **Cosine similarity**: The more similar an article is to an expert's published abstract, the longer the corresponding comment is. More specifically, a one-unit increase in cosine similarity leads to 9 times increase in the length of the expert's comments.
- **Social distance**: Behavioral and experimental economists contribute 16% more than experts in other fields.
- **Cosine similarity** has a similar significant effect on overall contribution quality.

Contribution quality

- Each comment independently rated by 3 trained coders
  - Doctoral students in Information Economics
  - Masters students in Economics and Information Economics
  - Junior and senior undergraduate economics majors
- Assignment based on courses taken
- Use median rating for analysis
- Distribution of median "overall quality"



Quality of comments: Self citation (logit)

Dependent Variable	Self-citation		
HighView	0.008	0.019	0.014
	(0.059)	(0.062)	(0.060)
Citation	-0.011	-0.015	-0.010
	(0.057)	(0.058)	(0.058)
Acknowledgement	0.093	0.095	0.103*
	(0.060)	(0.061)	(0.062)
Cosine Similarity		0.001	-0.011
		(0.186)	(0.186)
log(Page Length)		-0.012	-0.017
		(0.031)	(0.031)
Author Abstract View			-0.763*
			(0.462)
English Affiliation			0.057*
			(0.156)
Behavioral & Experimental Econ.			-0.021
			(0.049)
HighView+HighView×Acknowledgement	-0.187***	-0.207***	-0.206***
	(0.052)	(0.052)	(0.253)
Acknowledgement+HighView×Acknowledgement	-0.102**	-0.130**	-0.118**
	(0.051)	(0.053)	(0.052)

- Compared to AvgView-Acknowledgement, HighView-Acknowledgement discourages self-citation.

Concluding Remarks

- Eliciting interests from experts
  - Citation benefit at High View increases participation;
  - Public acknowledgement at High View decreases negative response.
  - Longer social distance and higher reputation decrease participation

Concluding Remarks

- Eliciting interests from experts
  - Citation benefit at High View increases participation;
  - Public acknowledgement at High View decreases negative response.
  - Longer social distance and higher reputation decrease participation
- Eliciting contribution quantity
  Similarity of an article to an expert's publication encourages contribution from experts

Concluding Remarks

- Eliciting interests from experts
  - Citation benefit at High View increases participation;
  - Public acknowledgement at High View decreases negative response.
  - Longer social distance and higher reputation decrease participation
- Eliciting contribution quantity
  Similarity of an article to an expert's publication encourages contribution from experts
- Eliciting high quality
  Acknowledgement elicits higher quality comments.

Concluding Remarks

- Eliciting interests from experts
  - Citation benefit at High View increases participation;
  - Public acknowledgement at High View decreases negative response.
  - Longer social distance and higher reputation decrease participation
- Eliciting contribution quantity
  Similarity of an article to an expert's publication encourages contribution from experts
- Eliciting high quality
  Acknowledgement elicits higher quality comments.
- Lessons learned:
  - Ask;
  - Who asks - social distance;
  - What do you ask: recommender system and expertise matching

Concluding Remarks

- Eliciting interests from experts
  - Citation benefit at High View increases participation;
  - Public acknowledgement at High View decreases negative response.
  - Longer social distance and higher reputation decrease participation
- Eliciting contribution quantity
  Similarity of an article to an expert's publication encourages contribution from experts
- Eliciting high quality
  Acknowledgement elicits higher quality comments.
- Lessons learned:
  - Ask;
  - Who asks - social distance;
  - What do you ask: recommender system and expertise matching
- Generalizable to other expert communities? arXiv