Understanding Voluntary Knowledge Provision and Content Contribution through a Social Media-Based Prediction Market: A Field Experiment

Online Appendix A: The Institutional Context of the Social Media-Based Prediction Market

The operation of this prediction platform is a typical prediction market with a binary option. Given the prediction market price, prediction market participants trade by simply clicking on either the "YEA" or "NAY" button on the prediction page. We use the following example to show how the prediction market works. For each prediction event, the starting price is 50 points. If more people disagree than agree, then the price of the prediction event will fall. For instance, the prediction event "Bitcoin will trade in a range between $200 and $1350 this year" had fallen from a starting point of 50 to 31.82. If a participant agrees with this prediction now, she can click on the "YEA" button and "buy" it using 31.82 points (play money, not real money). If the prediction turns out to be correct, she will win 68.18 points (100 - 31.82) when the prediction event closes. If the prediction is not correct, she will lose 31.82 points. In this case, the market prices can indicate what the crowd thinks the probability of the event is.

When a participant participates in a market, she will long/short based on her belief. For example, if she clicks on the "YEA" button, it implies that she wants to hold a positive position and buy the contract; she clicks on the "NAY" button, it implies that she wants to hold a negative position and sell the contract. However, due to the prediction market mechanism set by the company, a participant cannot buy or short more than one contracts even if she has a very strong positive/negative belief.

Our social media-based prediction market is a public prediction market with thousands of active prediction market participants. The site typically runs about 40 – 100 prediction markets at the same time. The following table shows a list of prediction topics.

Table A.1 A List of Prediction Content

VMware CEO Patrick Gelsinger will be the next CEO of $MSFT #ballmer.
Oil will average $100 + per barrel in 2014.
The S&P500 will plunge 30% by the end of 2014.
US crude oil production will expand for the sixth straight year in 2014.
Payroll data will come in lower than the 150,000 new jobs expected for Feb.
The Honda Fit will displace the Civic as the best-selling small car in North America before the end of 2014.
The Bank of England will raise interest rates in 2014.
Wendy Davis will run for Governor of Texas next year and win.
The US will boycott Sochi2014 over Russia's anti-gay policies.
Apple will authorize a 4 for 1 share split
George W. Bush will be the next Commissioner of MLB.
GOP regains control of the Senate in 2014 elections
Narendra Modi will not be prime minister of India after the elections later this year.
Obamacare will be repealed in advance of the 2014 elections
bitcoin will lose 99% of its value as it falls below $10
Foursquare will be acquired in 2014
Abenomics will not succeed in lifting Japan's core inflation rate to 2 per cent by the end of 2014.
An immigration deal in 2014 is unlikely
More incumbent Congress people will lose their seats this year than any time in the past 50 yrs
Congress will pass a federal minimum wage increase #minimumwage
Matthew McConaughey will win #Oscar for best actor
Norway will lead in medals at Sochi2014.
Anti-EU fringe parties will capture more than 25% of the vote in European parliamentary elections.
The plan for a new high-speed rail link between London, Leeds and Manchester will not be scrapped.
China's growth rate will not fall below 7 per cent in 2014

Online Appendix B: Fixed Effects Model

Even if the participants are randomly selected into a control/treatment group in our experiment, we may still worry about unobserved individual heterogeneity. In order to control for it, we introduce the following fixed effects model:

$$Y_{it} = \alpha_i + \beta_0 + \beta_1 t + \beta_2 (t \cdot AS_i) + \beta_3 (t \cdot ASOE_i) + \varepsilon_{it},$$

where α_i is the unobserved individual fixed effect. In a fixed effects model, any time-invariant variables (such as AS_i and $ASOE_i$) will be differenced out, so in our regression equation, we include only t, $t \cdot AS_i$, and $t \cdot ASOE_i$. The coefficients on the interaction terms capture the effects of audience size and online endorsements on individual prediction accuracy. In column 4 of Table 4, we find that these two coefficients are significantly positive, which are consistent with our previous results: the impacts of audience size and social endorsements are statistically significant.

Online Appendix C: Suggestive Evidence on User Effort

In this appendix, we provide additional suggestive evidence that the participants put more efforts because of social image.

It is worth nothing that we are not able to directly measure participants' effort using our data, so our evidence is mainly "suggestive" and indirect. In our data, we can observe the time at which each participant accepted our invitation, and the time at which each participant made predictions. Therefore, we can compute the average time each participant spent on prediction tasks. If the social treatments AS_i and $ASOE_i$ significantly increase the average time each participant spent on prediction tasks, then it tentatively suggests that the participants put more efforts because of social image. Therefore, we estimate the following regression model:

$$Time_{it} = \beta_0 + \beta_1 AS_i + \beta_2 ASOE_i + \beta_3 t + \beta_4 (t \cdot AS_i) + \beta_5 (t \cdot ASOE_i) + \varepsilon_{it}, \qquad \text{(C.1)}$$

where $Time_{it}$ is the average time each participant spent on prediction tasks (the unit is a day). The estimation results are presented in Table C.1. We find that the coefficients on $t \cdot AS_i$ and $t \cdot ASOE_i$ are significantly positive, suggesting that the social treatments AS_i and $ASOE_i$ increase the average time each participant spent on prediction tasks. The empirical evidence indirectly shows that the participants put more efforts on prediction tasks due to the impact of social image.

Table C.1. The Impact on the Average Time Each Participant Spent on Prediction Tasks

	(1)	(2)	(3)
VARIABLES	OLS	Robust Variance	Cluster Bootstrapping
t	-0.628**	-0.628**	-0.628**
	[-2.212]	[-2.034]	[-2.105]
Treatment AS	0.353	0.353	0.353
	[0.726]	[0.545]	[0.533]
Treatment AS + OE	-0.247	-0.247	-0.247
	[-0.854]	[-0.663]	[-0.592]
Treatment AS * *t*	1.252***	1.252***	1.252***
	[2.833]	[2.658]	[2.667]
(Treatment AS+OE) * *t*	2.035***	2.035***	2.035***
	[3.104]	[2.854]	[2.932]

t-statistics in brackets: *** p<0.01, ** p<0.05, * p<0.1