Results Report

Authority and Delegation in Online Communities

Jacopo Bregolin
Toulouse School of Economics

October 2020

Abstract

Many online platforms rely on user-generated content and need to incentivize free effort. In this paper, I investigate if users provide more and better quality contributions when endowed with more autonomy over actions. Using a dynamic discrete choice model, I show that control rights have positive marginal value that is heterogeneous across different types of users, where types are ex-ante identifiable by the platform. I simulate counterfactuals with different delegation designs. Results show that the platform would lose an important share of production and quality of content in absence of delegation.

Introduction and setting

Many digital companies build online communities to benefit from user-generated content. Leading examples are social media platforms, like Facebook, or information aggregators like Wikipedia. Many of these companies rely on voluntary users also for the moderation of the community. Instead of hiring professional administrators, they allow users to modify the content published on their website so that it conforms to the platform rules.
Platforms anyway differ on the extent they give authority to users. Facebook lets users flag content, but it does not allow them to modify or eliminate the posts. In Wikipedia instead, all users can modify the articles. Users in Wikipedia have full authority, while Facebook users have none. Other platforms allocate full authority when users reach given performance thresholds. What trade-off platforms face when making this decision?

In this paper, I study the incentive effects of allocating authority on quality and quantity of contribution. I find that users value to gain authority, and that rewarding users' performance with authority incentivizes contribution. Nevertheless, the incentive effect is heterogeneous and the optimal policy depends on the composition of the community.

I use data from Stack Exchange, a family of websites where registered users ask questions and provide answers on different topics. The website is moderated by experienced users who have full autonomy over editing questions and answers. New users' edits need to be approved by the moderators. New users become moderators after reaching a performance threshold. This design induces two types of incentives. If users value acquiring autonomy, delegation incentivizes effort until users reach the threshold (*dynamic incentive*). If users value to contribute when endowed with more autonomy, delegation relaxes the participation constraint, as it relatively decreases the outside option (*static incentive*). Stronger *dynamic incentive* effect would suggest increasing the performance threshold, while stronger *static incentive* to decrease it. The paper studies the platform's trade-off by quantifying both incentive effects under different counterfactual performance thresholds. Figure 1 represents the trade-off.

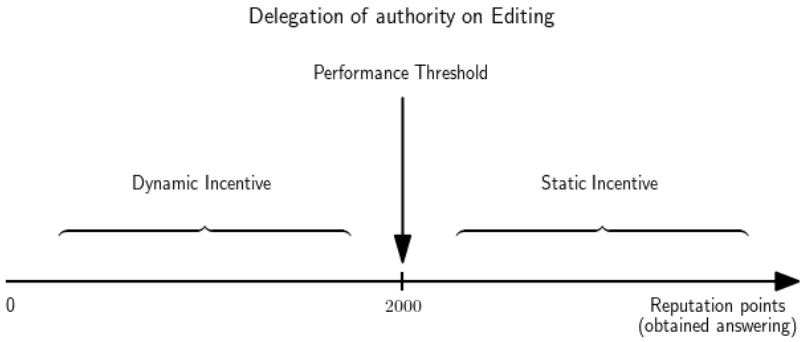

Figure 1: Delegation system in Stack Exchange. Users can answer questions or edit existing answers. Before they reach the performance threshold, edits are suggested and need third party approval, while, passed the threshold, users obtain authority on the implementation of edits. If users want to reach the threshold, the delegation system incentivizes contribution before users reach the threshold (*dynamic incentive*). If users value to participate when endowed with authority, the delegation system incentivizes contribution after users reach the threshold (*Static incentive*).

Data

I observe the full participation history of all users registered in *English Language Learners*, one of the websites of the Stack Exchange family. The analysis includes users who published at least one answer, and gained a positive number of reputation points. I categorize users in *types* based on the information they display in their profile page.[1] I recover 3 types: one group of users does not provide much information (*Anonymous*), a second group provide some information (*Identifiable*), and the third provide a lot of information, including links to external sources like Linkedin profiles (*Informative*). These differences are

[1] I implement a Multiple Correspondence Analysis combined with a K-Means clustering algorithm.

assumed to proxy for the broad motives that induce participation. Table 1 reports the total and average activity in answering and editing. It shows strong heterogeneity in behavior across groups.

	Num	num answers		num suggested edits		num direct edits	
	Users	Total	Avg. per user	Total	Avg. per user	Total	Avg. per user
Data	9797	114926	11.73	1409	0.14	6759	0.69
Type *Anonymous*	5414	32511	6.00	309	0.06	465	0.09
Type *Identifiable*	3705	63500	17.14	836	0.23	2272	0.61
Type *Informative*	678	18915	27.90	264	0.39	4022	5.93

Table 1: Activity in the website for answering and editing. Suggested edits are edits made before the user obtained the threshold number of reputation points, while direct edits are made after, when users acquire full autonomy on editing.

Reduced form analysis

Using a regression discontinuity design, I test for the presence of *static incentives*. If users value more participation when endowed with authority, we should observe an increase in participation after they reach the performance threshold. I then estimate the following regression:

$$Y_{it} = \alpha_i + \gamma_t + \beta_\rho + a_m + b_c + \varepsilon_{it},$$

where i indexes users and t weeks. The outcome variable, Y_{it}, is the standardized number of edits(comments) published in week t by user i. The coefficients of interest are the parameters $\{\beta_\rho\}_{\forall\rho}$, capturing a fixed effect of missing ρ points to reach the threshold. I bin the number of points in 50-points intervals and re-label each bin relative to the bin at which users receive authority. In bin 0 users have just passed the threshold, in bin 1 they have 50 to 100 points more than the threshold, while in bin -1 they miss 1 to 50 points to reach the threshold. Estimates of the beta coefficients for the 6 bins before and after the threshold are plotted in figure 2. Finally, since users can also get authority via elections, I control if the user, in the given week, is an elected moderator (a_m) or a candidate in the election (a_c).

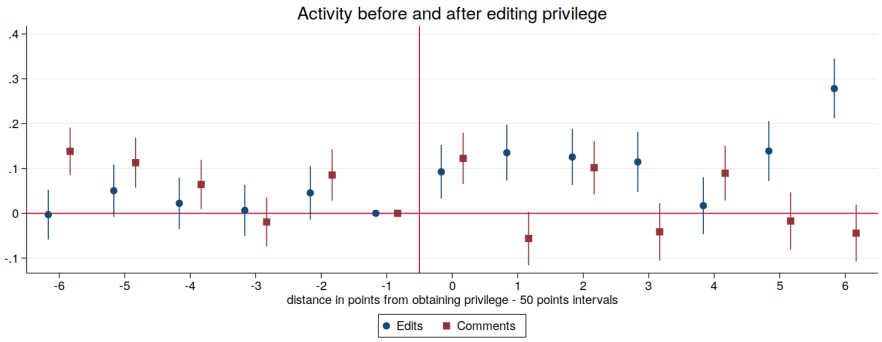

Figure 2: Contribution fixed effects of being n-distant from the performance threshold, where distance is measured in 50 points intervals. Circles are fixed effects on editing, while squares are fixed effects on commenting Vertical bars are confidence intervals.

Results show that once endowed with authority, users significantly increase the production of edits, while they do not in comments. Note that comments may have a similar purpose to edits, i.e. improve existing content, but the authority on commenting is unchanged. Figure 3 shows estimates of the same regression discontinuity model estimated separately for the different types of users. It shows that the effect is mainly driven by *Anonymous* users.

Structural Model

To quantify the incentive effects and simulate counterfactuals, I use a structural model of dynamic discrete choice. In each period, users decide their contribution in terms of number of answers, quality of answers, and number of edits. The utility function is modeled as follow:

$$U_{it} = \beta_0 R_{it} + \beta_1 CA_{it} + \beta_2 CE_{it} + \beta_3 cumT_{it} + Authority_{it} \times (\beta_4 + \beta_5 CA_{it} + \beta_6 CE_{it})$$

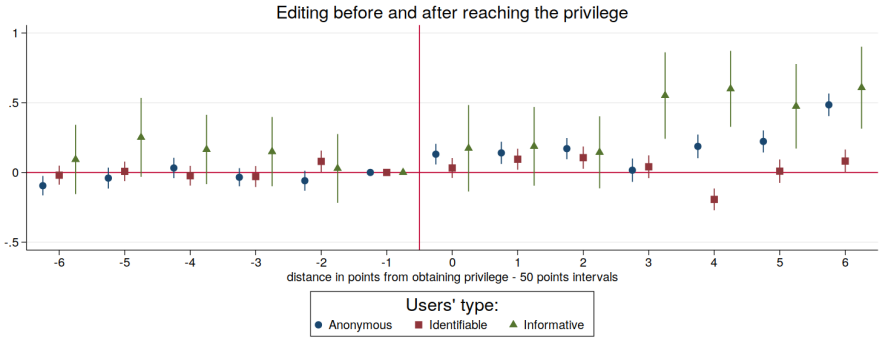

Figure 3: Editing fixed effects of being n-distant from the performance threshold, where distance is measured in 50 points intervals, for each type of user. Vertical bars are confidence intervals.

where i indexes the users, and t the week. The flow utility depends on the number of reputation points the user has accumulated up to period t (R_{it}), the net benefits/costs of answering and editing (CA_{it} and CE_{it} respectively), and the cumulative amount of *privileges* obtained, i.e. the sequential rewards that can be obtained accumulating points. It includes a dummy variable equal to 1 if the user reaches the required performance threshold for authority (*Authority*). Identification of the *dynamic incentive* relies on effort that users make when approaching the performance threshold: higher effort allows them to reach the threshold faster. Systematic higher effort approaching the threshold would identify a positive marginal utility of authority. Variation in the willingness to participate once endowed with authority identifies the *static incentive* effect, and it is captured by the interaction of *Authority* with the net-costs variables. The user, at each period, chooses the effort in answering and editing to maximize:

$$\mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1}(U_{it} + \varepsilon_{it})\right]$$

where *delta* is the discount factor and ε an idiosyncratic preference shock.

Estimates of the flow utility parameters are presented in table 2. It emerges that the *Anonymous* users have the highest costs of participation, but the highest marginal utility from the acquisition of authority. All users increase their willingness to participate in editing once endowed with authority, but the effect is again stronger for the *Anonymous* users, in line with the results of the reduced form analysis. While the *Informative* have also a similar value for gaining authority, this is not the case for the *Identifiable* users.

Variables	(no Heterogeneity)	(*Anonymous*)	(*Identifiable*)	(*Informative*)
R	0.0069***	0.0064***	0.0057***	0.0045***
	(0.0001)	(0.0005)	(0.0002)	(0.0004)
CA	-0.0001	-0.3563***	.00005	0.0007***
	(0.0008)	(0.0196)	(.0006)	(0.0002)
CE	-10.3311***	-7.9549***	-6.1724***	-5.7740***
	(0.4979)	(0.8927)	(0.4051)	(0.4757)
cumT	-0.7745***	-0.4177***	-0.7855***	-0.7681***
	(0.0206)	(0.0322)	(.028)	(0.0563)
Authority	1.3162***	1.5223***	0.1713	1.4709***
	(0.1203)	(0.3577)	(0.2535)	(0.5118)
CA x Authority	0.0609***	-0.0048***	-0.0018	-0.0008
	(0.0036)	(0.0016)	(0.0011)	(0.0014)
CE x Authority	12.2064***	0.6338***	0.2507***	0.2703***
	(0.5247)	(0.0593)	(0.0308)	(0.0274)
N. users	9,783	3,700	5,407	676
Sample size	991,657	471,837	407,098	112,722

$*\ p < 0.05$, $**\ p < 0.01$, $***\ p < 0.001$

Table 2: Estimates for the flow payoff parameters. Standard errors in parenthesis.

Performance required	Prod. Answers	Change	Anonymous	Identifiable	Informative
0 Points	12562.0		92	10967	1503
500 Points	13374.0	+6.46%	+13.04%	+1.6%	+41.52%
1000 Points	13300.0	+5.87%	+13.04%	+2.43%	+30.54%
NO Delegation	12886.0	+2.58%	+13.04%	+2.01%	+6.12%

Table 3: Total production of answers in each simulated scenario, preserving the proportion of types as in the true data.

Counterfactuals

Counterfactual simulation study how many answers users would publish under alternative delegation designs. In particular, I study delegation at different performance thresholds, including zero (full delegation), two intermediate levels, and infinity (no delegation). Results show that the response to the incentive is very heterogeneous. Figure 4 reports the average number of answers published by each user type during their participation history. The *Anonymous* users, who value authority the most, have the highest cost of participation. It follows that the marginal utility of authority is not high enough to incentive participation, and they make on average very little contribution. The *Identifiable* users instead have a lower cost of participation, but they are not responsive to the *dynamic incentive*, as their marginal utility of authority is not significantly different from zero. It follows that their participation history does not depend much on the delegation incentives, but rather on the accumulation of points and on intrinsic utility from answering. Finally, the users of the *Informative* type have a higher marginal utility of authority and low participation cost. This implies that they are particularly sensitive to the *dynamic incentive*. While they always produce a positive amount of answers, their contribution increases when delegation is conditional to the achievement of performance, more when the performance threshold is not too demanding.
Overall it emerges that the *dynamic incentive* dominates the *static incentive* in answering. This is because the additional value in answering after the threshold is reached is close to zero for all types, as captured by the estimates for the interaction variable $CA \times Authority$ in table 2.

To simulate the overall production of answers in the different scenarios, I sample users of each type according to the composition of the community in the real data. This correspond to 55% of *Anonymous* users, 38% of *Identifiable* users, and 7% of *Informative* users. Table 3 reports the total number of answers produced under each scenario and the contribution of each type of user. As suggested from the previous analysis, most of the difference in production is made by users of type *Informative*, who are the most reactive to the *dynamic incentives*. Figure 5 plots how the total production would be distributed along the lifetime of the platform.

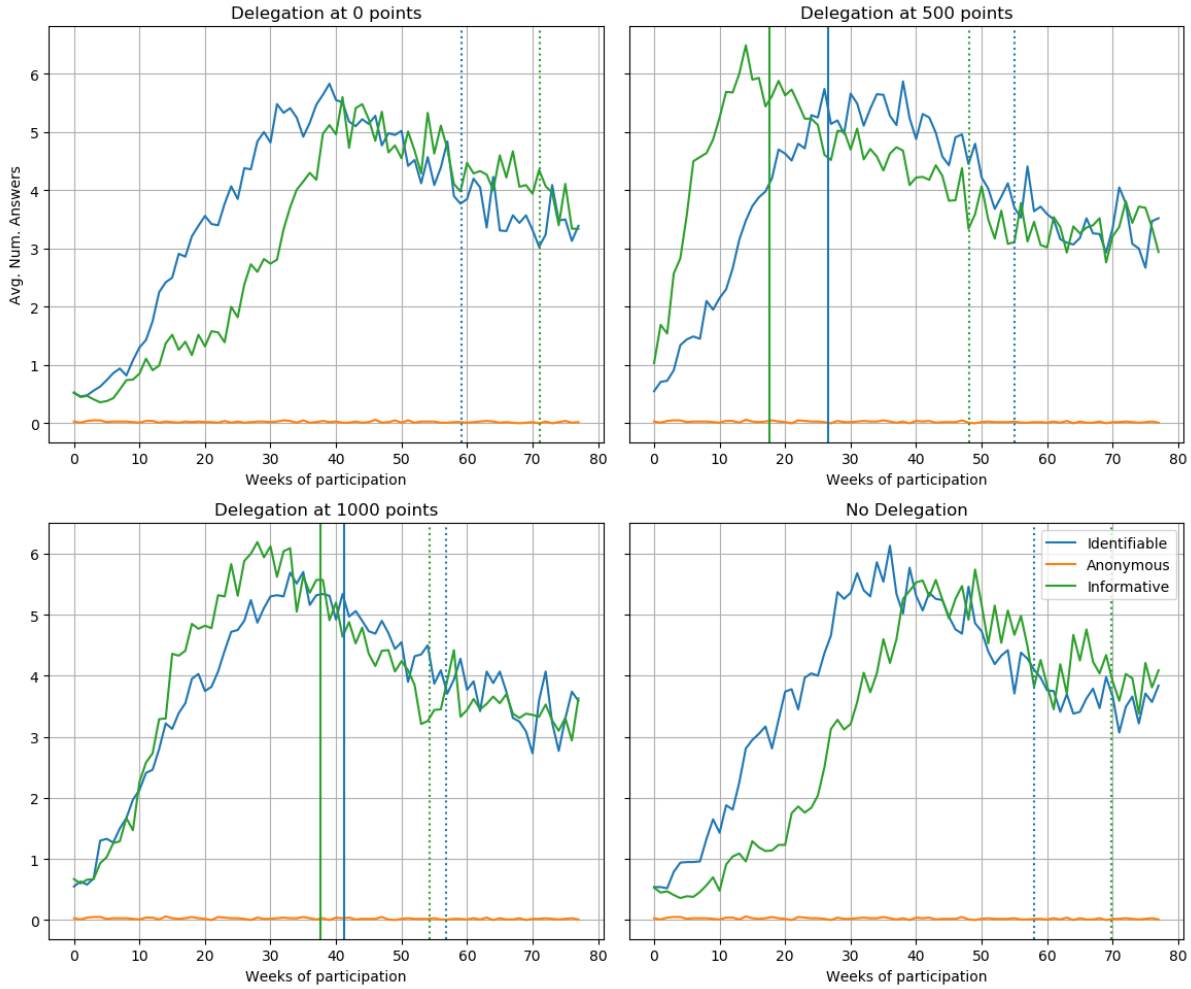

Figure 4: Simulated average number of answers produced by users of each type in 80 days of participation. Each plot corresponds to a different threshold of delegation, from full delegation to no delegation. Simulations were made for 100 users and 100 periods of participation, assuming that after 100 periods users exit the platform. Users cannot accumulate more than 1500 points. Vertical solid lines correspond to the average period in which the users reach the threshold, while vertical dotted lines correspond to the average period in which users reach the 1500 points.

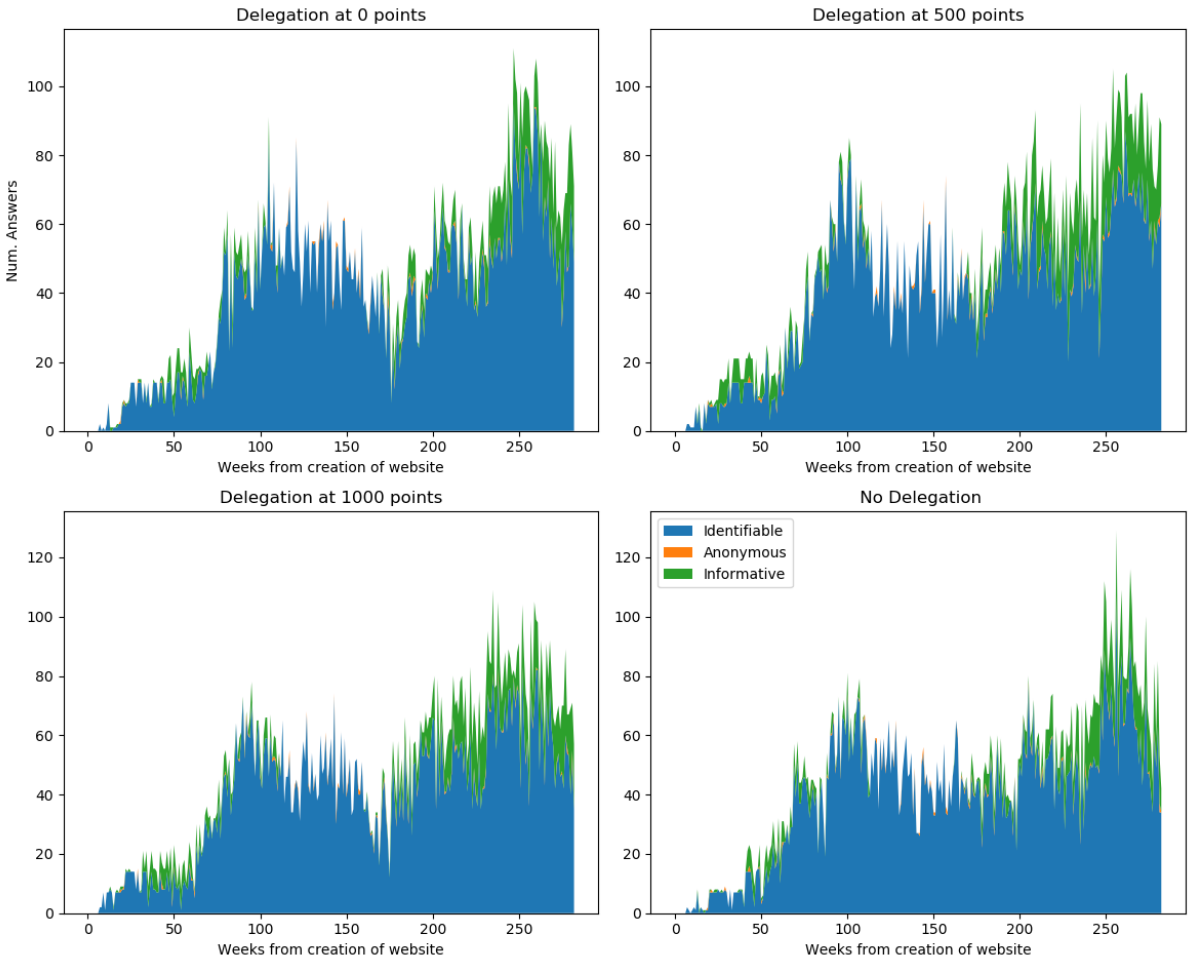

Figure 5: Total simulated production of answers along the lifetime of the website, per type of users.