Stack Overflow badges and user behavior

An econometric perspective

Andrew Marder

February 20, 2015

Introduction

Stack Overflow is a question and answer community designed for programmers. It is the largest of 130 communities in the Stack Exchange network. Created in 2008, the knowledge organized by Stack Overflow has become a valuable resource for software developers. On January 20, Spoelsky (2015) announced that Stack Exchange had raised $40 million in venture capital funding. Stack Exchange gives users who ask questions access to expert technical help, users who answer questions build their reputation for technical expertise and use that reputation to find better jobs.

Although Stack Overflow's creation was rooted in computer science, the social sciences will provide important insights as the platform matures. There are a number of big picture questions to consider. How should Stack Overflow design its platform to maximize profits? Is the firm's profit-maximizing strategy efficient - does it maximize total surplus? How equitable is the profit-maximizing strategy - Stack Overflow creates value by improving job matches, how much of that value is captured by the platform, employers, and employees? What motivates users to contribute to Stack Overflow - how important are self-interested motives versus pro-social motives? What combination of reputation rules and badge awards maximize the quantity and quality of user contributions?

This paper takes a first step along the path of applying econometric analysis to publicly available Stack Overflow data. Specifically, I study how users behave around the time they are awarded badges. Taking a closer look at user actions, we find some badges are effective at motivating user contributions while others are not.

Rules of the game

Deterding et al. (2011) define "*gamification* as the use of game design elements in non-game contexts." Stack Overflow gamifies the process of asking and answering questions as follows. A user earns reputation points when another user votes on her posts (5 points when a question is voted up, 10 points when an answer is voted up, 15 points when an answer is accepted, and 2 points when an edit is approved). As a user earns reputation points she unlocks privileges on the site. For instance, a user must have at least 15 reputation points to vote up a question or answer.[1] Users are awarded badges for special achievements. One receives the *Informed* badge by reading the tour page.[2]

[1] A full list of privileges and necessary reputation points is available at http://stackoverflow.com/help/privileges.

[2] The Stack Overflow tour can be found at http://stackoverflow.com/tour, and all badges are listed on http://stackoverflow.com/help/badges.

How users behave when earning badges

Grant and Betts (2013) present empirical evidence that three of the badges awarded for various editing accomplishments are effective in encouraging users to make more edits in the two months preceding receipt of the badge compared to the two months after receiving the badge. This paper builds on their findings by:

1. Looking at the impact of badges on all types of user activity (posting questions, posting answers, and editing posts).
2. Comparing the impact of different badges. In addition to the three editing badges, this paper also looks at two badges awarded for asking questions.

Grant and Betts (2013) find that users who receive a badge for editing make more edits in the two-month window before receiving the badge compared to the two-month window after receiving the badge. I extend their work by exploring, on average, how many questions, answers, and edits a user posts around the time of receiving a badge. Let y_{it} be the number of edits user i makes on day t. Following the approach of Jacobson, LaLonde, and Sullivan (1993) define the dummy variable

$$D_{it}^k = \begin{cases} 1 & \text{if user } i \text{ earns the badge on day } t+k, \\ 0 & \text{otherwise.} \end{cases}$$

I regress the number of edits user i makes on day t on a user fixed effect α_i and a set of dummy variables indicating whether the user received the badge of interest on day $t+k$

$$\log(1+y_{it}) = \alpha_i + \sum_{k=-29}^{30} D_{it}^k \delta_k + \epsilon_{it}.$$

The model parameters are estimated using an ordinary least squares regression, and standard errors are clustered at the user level.

Let t_i^* denote the day user i recieves the badge. Figure 1 plots the expected number of actions taken on the k'th day since receiving the badge

$$f(k) = E\left[\log(1+y_{it}) \mid t = t_i^* + k\right].$$

Strunk & White Edited 80 posts Copy Editor Edited 500 posts (excluding own or deleted posts and tag edits) Archaelogist Edited 100 posts that were inactive for 6 months Curious Asked a well-received question on 5 separate days, and maintained a positive question record Inquisitive Asked a well-received question on 30 separate days, and maintained a positive question record

"Strunk & White : 6923" "Archaeologist : 622"
"Copy Editor : 1207"
"Inquisitive : 160"
"Curious : 1463"

The 95% confidence interval is tight around the line, standard errors were calculated using the delta method (Williams 2012). Figure 1 confirms the findings of Grant and Betts (2013), editing increases gradually before receiving the badge with a large jump in activity in the day immediately before earning the badge. We also see that editing drops quickly after receiving the badge and gradually declines over time. It's interesting to see how few questions were asked by the recipients of the editing badges in the two months around receiving the badge, and to see that the rate of answering questions tends to be constant throught the two month window.

Figure 2 plots user activity around receiving badges for asking questions. User activity around question badges differs in interesting ways to badges for edits:

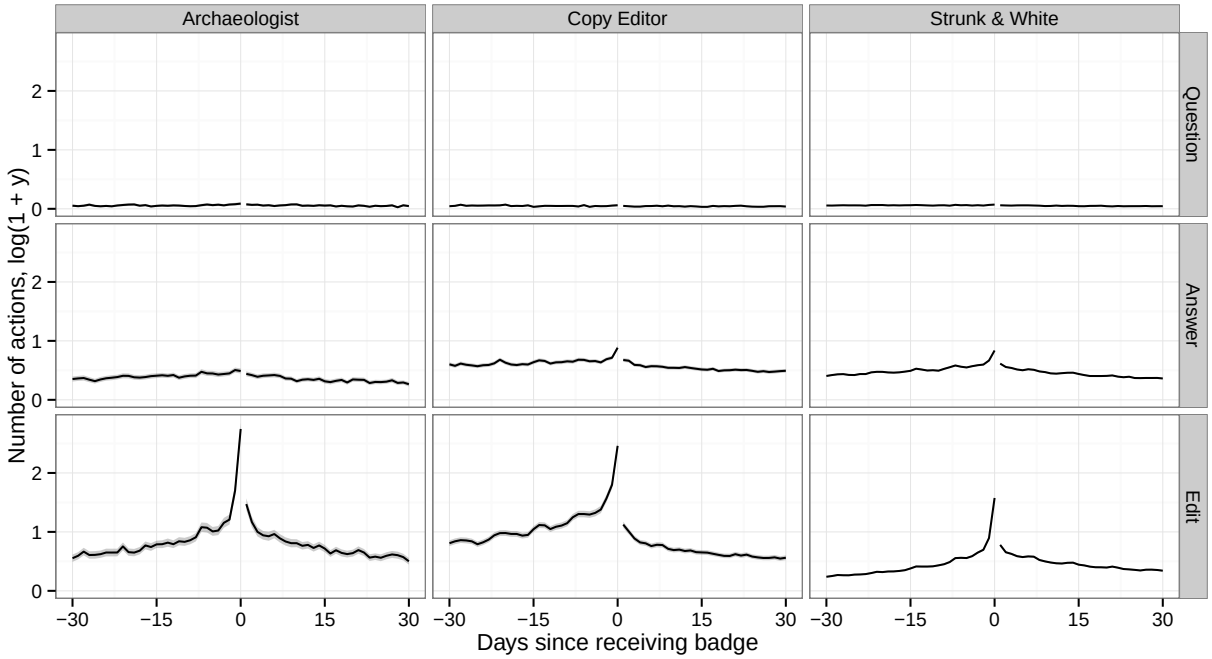

Figure 1: User activity over time - badges for edits

1. All types of actions tend to increase in the thirty days leading up to earning the badge. In Figure 1, only edits increase. In Figure 2, the number of questions, answers, and edits posted all increased.

2. User activity stops almost immediately after receiving the badge.

Conclusion

In line with Spence's (1973) model of job market signaling, Stack Overflow provides a platform for job searchers to signal their ability by answering difficult technical questions publicly. In the process, job searchers contribute to the provision of a valuable public good.

When interpreting the empirical results of this paper, please consider Holland and Rubin's motto "no causation without manipulation" (Holland 1986). There is no manipulation of the explanatory variable in this study, consequently we have not identified the causal effect of badges. To estimate the causal impact of badges on user activity we need to find a source of exogenous variation (Miller 2013).

This paper confirms the empirical observation of Grant and Betts (2013), on average users who receive a badge for editing make more edits in the 30 days prior to receiving the badge compared to the 30 days after receiving the badge. In addition, we show that the average number of questions and answers posted do not change around the receipt of an editing badge. Finally, we show that users who received badges for asking questions behaved differently. In particular, we found that the number of answers and edits posted followed a similar time trend as the number of questions asked. This suggests there may be an interesting spillover effect, encouraging users to ask more questions may have a secondary effect of encouraging general activity on the site.

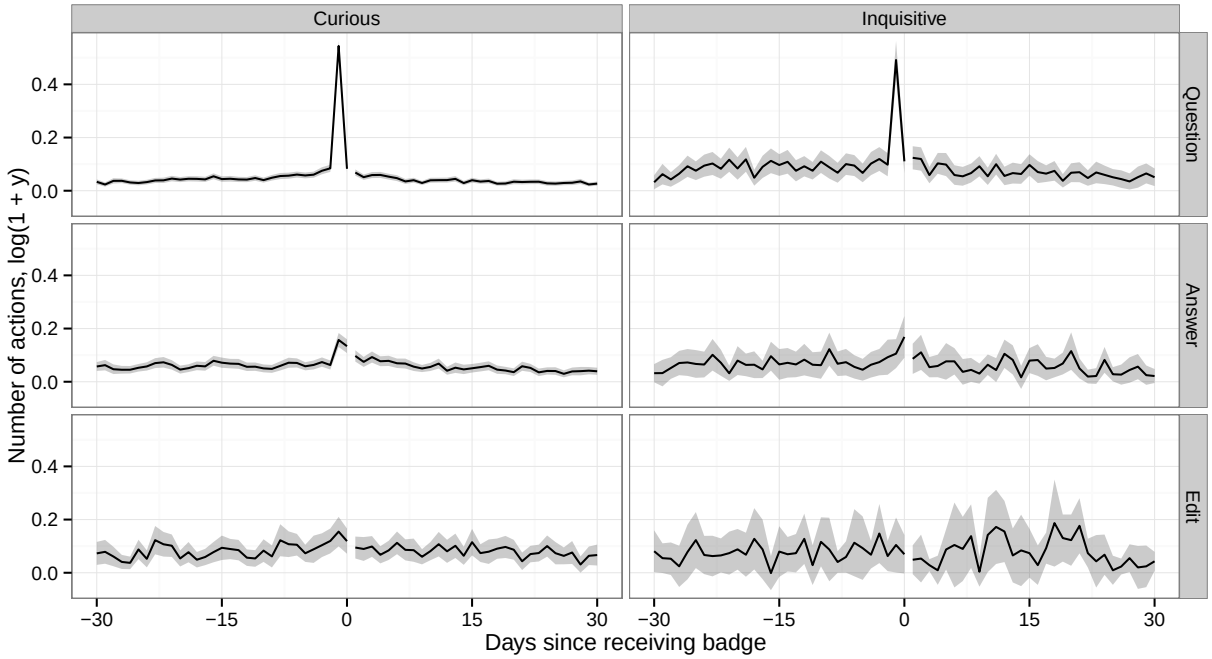

Figure 2: User activity over time - badges for questions

References

Antin, Judd, and Elizabeth Churchill. 2011. "Badges in social media: A social psychological perspective." In *CHI 2011 Gamification Workshop Proceedings*, 1–4. http://uxscientist.com/public/docs/uxsci_2.pdf.

Deterding, Sebastian, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. "From game design elements to gamefulness: Defining 'Gamification'." In *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11*, 9–11. doi:10.1145/2181037.2181040.

Grant, Scott, and Buddy Betts. 2013. "Encouraging user behaviour with achievements: An empirical study." In *IEEE International Working Conference on Mining Software Repositories*, 65–68. doi:10.1109/MSR.2013.6624007.

Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–60. doi:10.1080/01621459.1986.10478354.

Jacobson, Louis, Robert LaLonde, and Daniel Sullivan. 1993. "Earnings losses of displaced workers." *The American Economic Review*, 685–709. http://www.jstor.org/stable/2117574.

Miller, Ashley. 2013. "Principal turnover and student achievement." *Economics of Education Review* 36 (October): 60–72. doi:10.1016/j.econedurev.2013.05.004.

Spence, Michael. 1973. "Job Market Signaling." *The Quarterly Journal of Economics* 87: 355–74. doi:10.2307/1882010.

Spoelsky, Joel. 2015. "Andreessen Horowitz Invests in Stack Exchange." http://blog.stackoverflow.com/2015/01/andreessen-horowitz-invests-in-stack-exchange/.

Williams, Richard. 2012. "Using the margins command to estimate and interpret adjusted predictions and marginal effects." *Stata Journal* 12 (2). http://econpapers.repec.org/article/tsjstataj/v_3a12_3ay_3a2012_3ai_3a2_3ap_3a308-331.htm.

Ying, Annie T. T. 2015. “Mining Challenge 2015: Comparing and combining different information sources on the Stack Overflow data set.” In *The 12th Working Conference on Mining Software Repositories*, to appear.