Stack Overflow and Gamification - Something from Nothing?

Andrew Marder

February 6, 2015

What is Stack Overflow?

Deterding et al. (2011) define "*gamification* as the use of game design elements in non-game contexts." Stack Overflow is a question and answer website designed for programmers that employs a few game elements:

1. 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).

2. 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. A list of all privileges is available here.

3. Users are awarded badges for special achievements. One receives the "Informed" badge by reading the entire tour page. A list of badges is available here.

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. Looking at three different badges, each aimed at promoting a different type of user activity.

Data

There are almost 3.5 million registered users on Stack Overflow. Less than one percent of those users have been awarded the Strunk & White, Copy Editor, or Archaeologist badges.

How do users behave when earning badges?

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

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

The 95% confidence interval is tight around the line, standard errors were calculated using the delta method (Williams 2012). Confirming the conclusion of Grant and Betts (2013), we see that badge recipients drastically increase activity before receiving the Copy Editor badge making 24.6 edits in the 24 hours immediately before receiving the badge and dropping down to 2.9 edits in the 24 hours immediately after receiving the badge.
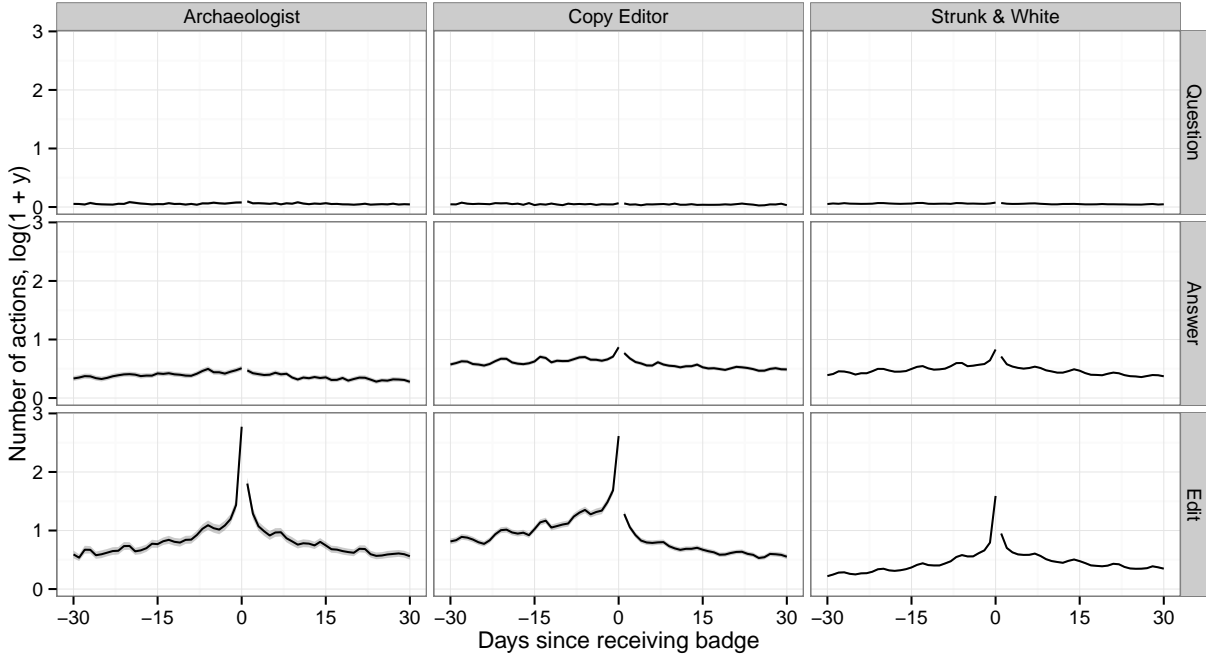

Figure 1: Mean number of actions performed over time

Conclusion

As a fan of Holland and Rubin's motto "no causation without manipulation", it is important to note that this paper does not identify the causal effect of badges (Holland 1986). To reliably estimate the causal impact of
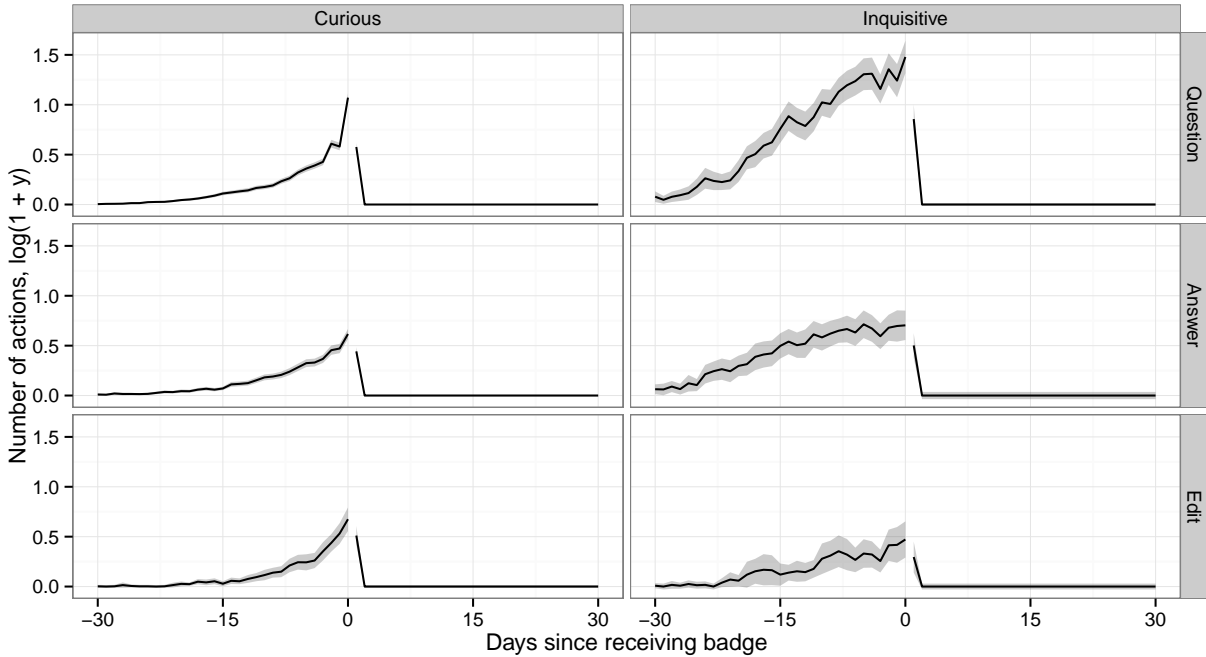
Figure 2: Event study around badges awarded for asking questions

badges of user activity we need a source of exogenous variation (Miller 2013). This paper contributes to the literature by describing more clearly how users behave around receiving badges.

References

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: Envisioning Future Media Environments*, 9–15.

Grant, Scott, and Buddy Betts. 2013. "Encouraging User Behaviour with Achievements: An Empirical Study." In *2013 10th IEEE Working Conference on Mining Software Repositories (MSR)*, 65–68.

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

Jacobson, Louis, Robert LaLonde, and Daniel Sullivan. 1993. "Earnings Losses of Displaced Workers." *The American Economic Review*, 685–709.

Miller, Ashley. 2013. "Principal Turnover and Student Achievement." *Economics of Education Review* 36 (October): 60–72.

Williams, Richard. 2012. "Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects." *Stata Journal* 12 (2).