Supplementary Material for "All Talk: How Increasing Interpersonal Communication on Wikis May Not Enhance Productivity"

Sneha Narayan
snarayan@carleton.edu

Nathan TeBlunthuis
nathante@uw.edu

Wm Salt Hale
halew@uw.edu

Benjamin Mako Hill
makohill@uw.edu

Aaron Shaw
aaronshaw@northwestern.edu

This supplement presents additional descriptive statistics of our sample, as well as analyses we conducted in order to evaluate the validity of our findings in response to several potential threats.

1 Summary Statistics

Table 1 presents univariate summary statistics of our eight dependent variables. We find that almost all of our dependent variables have very skewed distributions with sample means that are considerably larger than the sample medians. As is typical of participation rates in many online communities, the vast majority of contributors and communicators in our sample participate very little, while a small handful participate a quite a lot.

2 Retaining Week 0

In our paper, we presented models estimated using data where we drop all observations from the week immediately after the transition (referred to as *week 0*). We did so because we identified a 'bump' in many of our dependent variables during week 0.

We present additional evidence for this bump in Table 2 by showing the change in medians in the weeks surrounding the transition. This is structurally similar to Table 1 in our paper, which reports the change in means.

	Outcome	Min.	Median	Mean	Max.	Std. Dev.
All Users	No. of messages	0	0	5.57	685	25.19
	No. of communicators	0	0	1.40	43	2.60
	No. of contributors	0	3	5.30	272	10.52
	No. of contributions	0	17	80.09	26488	470.55
New Users	No. of messages	0	0	0.86	61	3.10
	No. of communicators	0	0	0.42	20	1.06
	No. of contributors	0	1	2.15	177	6.86
	No. of contributions	0	2	8.57	849	31.52

Table 1: Summary statistics for the eight dependent variables across all weeks in our study. Each variable describes activity taking place within a wiki each week.

	Outcome	Week -1	Week 0	Week 1
All Users	No. of messages	1	2	1
	No. of communicators	1	2	1
	No. of contributions	23	37	31
	No. of contributors	3	4	4
New Users	No. of messages	0	0	0
	No. of communicators	0	0	0
	No. of contributions	3	4	3
	No. of contributors	1	1	1

Table 2: Median outcomes across all wikis for the weeks immediately around the transition.

	Outcome	Est.	SE
All Users	M1a: No. of messages	1.027 ***	0.10
	M1b: No. of communicators	0.64 ***	0.06
	M2a: No. of contributors	0.122 ***	0.03
	M2b: No. of contributions	0.468 ***	0.07
New Users	M3a: No. of. messages	0.724 ***	0.14
	M3b: No. of communicators	0.593 ***	0.10
	M4a: No. of contributors	0.055	0.05
	M4b: No. of contributions	0.097	0.09

Table 3: Estimates for the *msgwall* term in models that include data from the week after the transition to message walls.

In four out of our eight dependent variables, we see that the median value of the measure shows an increase in week 0 relative to the weeks immediately before and after week 0. Though these increases might appear small, an increase in the median by even one can indicate considerable variation in our sample because the distributions of all our outcome variables are quite skewed and include many zeroes.

The strongest evidence for the presence of this bump is seen in Table 3, where we report the results of an analysis that is identical the one included in our paper but estimated using a dataset with week 0 included. In M1a, M1b, M3a, and M3b, we see that estimate sizes are considerably higher—in some cases more than double of what they are without week 0. Although they are not statistically significant in the models reported in the body of our paper that exclude week 0, Table 3 suggests support for our hypotheses in M2a and M2b. Additionally, M4a and M4b fit using data on week 0 suggest positive estimates, rather than negative estimates shown in the paper.

Although these results suggest that message walls may have had a positive effect on contributions to article pages, we chose not to present them in our paper because in each case, these estimates were driven by a single extreme data point within each wiki. We are also concerned that a higher amount of interest and attention to wikis by administrators may have driven both a short term increase in activity and the decision to turn on message walls.

Although we believe that the results included in our paper are the more conservative estimates that are more likely to be robust to this threat, the results in Table 3 provide some evidence in support of the idea that message

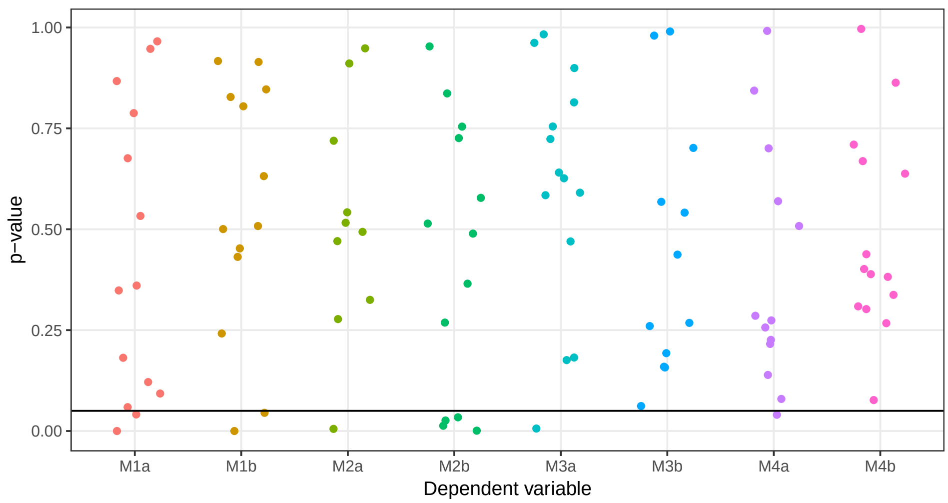

Figure 1: Placebo test results. P-values for the effect of message walls from models fit to our placebo datasets grouped by dependent variable.

walls may have a very short-term effect in the hypothesized directions.

3 Placebo Tests

We wished to determine that the relationships in our reported models were driven by the transition to message walls and not by spurious correlations caused by underlying noise in our outcome measures. We reasoned that if we were seeing significant effects on our outcome variables that were brought about the message wall transition, we would not see those effects in other 16-week periods on wikis where there was no such transition.

To this end, we conducted a number of 'placebo' tests[1] by running our models on 14 new datasets—drawn from real data—but constructed around fictional message wall transition dates. We made these datasets by transposing the real transition date by non-overlapping periods of approximately 120 days before and after the true cutoff. This resulted in non-overlapping 16-week periods that bounded 14 different fictional transition dates. These datasets placed a fictional cutoff at 120, 241, 363, etc. days after the real cutoff and another where it was placed 120, 241, 363, etc. days before. We estimated all of our models in each of these datasets.

Our results of these models are summarized in Figures 1 and 2 where each point reflects the p-value associated with a hypothesis test for the effect

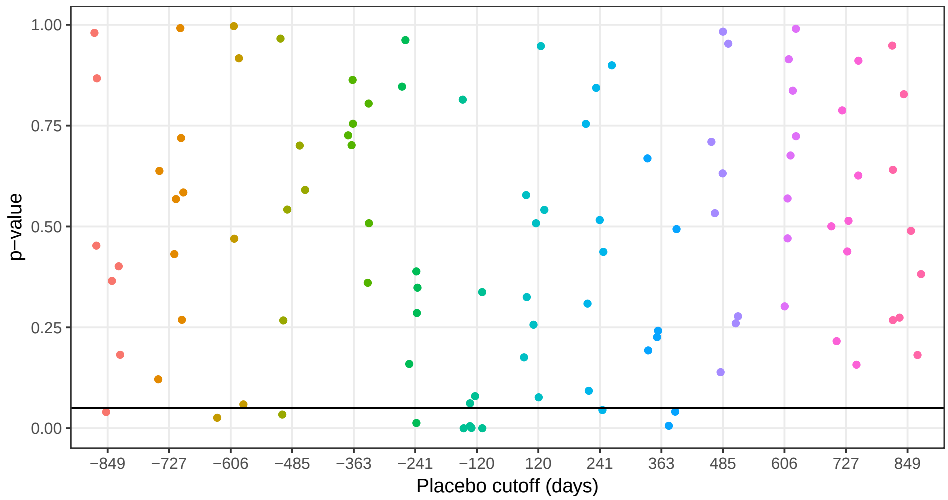

Figure 2: P-values of models run in placebo datasets with various artificial cutoffs.

of message walls in a given model across a given placebo dataset. Figure 1 shows the results of hypothesis test clustered by model or dependent variable. Figure 2 shows the same data clustered by fictional cutoff or dataset. Overall, we found that 90% of the models we ran yielded no significant effects around the fictional transition dates. This suggests that although there is some evidence of underlying noise in the data that could be driving spurious results, the effects we estimate around the true transition dates are likely due to the variation introduced by the message wall feature.

4 Shorter and Longer Analytic Windows

Another possible threat to our findings concerns the number of observations we choose to make before and after the transition to message walls. It is theoretically possible that our results are impacted by the length of the study-period we chose to examine, rather than just the transition to message walls.

To test our results' sensitivity to this threat, we ran our models on datasets where we changed the analytic windows of our study from 8 weeks before and after the transition to 6 and 12 weeks. Table 4 and Table 5 show the results of these alternate specifications.

Table 4 shows that reducing the analytic window to 6 weeks before and

	Outcome	Est.	SE
All Users	M1a: No. of messages	0.414 ***	0.12
	M1b: No. of communicators	0.178 *	0.07
	M2a: No. of contributors	-0.03	0.04
	M2b: No. of contributions	-0.064	0.08
New Users	M3a: No. of. messages	0.256	0.17
	M4a: No. of contributors	-0.195 **	0.06
	M4b: No. of contributions	-0.418 ***	0.11

Table 4: Estimates for the *msgwall* term in models that include data from six weeks before and after the transition to message walls.

	Outcome	Est.	SE
All Users	M1a: No. of messages	0.59 ***	0.09
	M1b: No. of communicators	0.309 ***	0.06
	M2a: No. of contributors	0.049	0.03
	M2b: No. of contributions	-0.002	0.06
New Users	M3a: No. of. messages	0.544 ***	0.13
	M3b: No. of communicators	0.418 ***	0.10
	M4a: No. of contributors	-0.01	0.05
	M4b: No. of contributions	-0.07	0.08

Table 5: Estimates for the *msgwall* term in models that include data from twelve weeks before and after the transition to message walls.

after the transition generally produces a pattern of results that are similar to those from our original models. We see the same pattern of findings in M1a, M1b, M2a and M2b as we did in the original models, with significant, positive estimates for θ in M1a and M1b and insignificant estimates for θ in M2a and M2b. For M3a, M4a and M4b, we see estimate sizes for the effect of message walls all have the same signs as in the original models, yet there is some difference in which models had statistically significant estimates for θ. In this case, the estimate for θ in M3a is insignificant, but significant for M4a. The opposite was true in the original models. We were unable to fit model M3b with this specification because there was very little variation in the dependent variable during this shortened analytic window.

Table 5 shows that increasing the analytic window to 12 weeks before and after the transition shows the same pattern of significance as the estimates

	Outcome	Est.	SE
All Users	M1a: No. of messages	0.577 ***	0.10
	M1b: No. of communicators	0.318 ***	0.06
	M2a: No. of contributors	0.049	0.03
	M2b: No. of contributions	0.045	0.06
New Users	M3a: No. of. messages	0.467 **	0.15
	M3b: No. of communicators	0.437 ***	0.11
	M4a: No. of contributors	-0.053	0.05
	M4b: No. of contributions	-0.187 *	0.09

Table 6: Estimates for the *msgwall* term in models where data is binned in four day long ‘weeks’.

for θ reported in Table 2 of our paper for all models but one—M4b is no longer shows a significant effect for the transition to message walls in this specification. We concluded that the length of the analytic window could impact the significance of some of our findings, but did not change the overall pattern of our results.

5 Smaller and Larger Time Windows for Binning Data

The analysis we presented was based on measures constructed by binning a week’s worth of activity in each data point. It is possible that our findings could be driven by this choice. While larger bins could over-smooth data and lose variation, smaller bins could introduce more variation in ways that could drive the results. Both these outcomes could affect the way we estimate the local trend around the transition.

To estimate how sensitive our results were to the decisions we made about binsize, we ran our models on data binned with alternate definitions of ‘weeks.’ In Tables 6 and 7, we report the results of running our models on 4-day long ‘weeks’ and 10-day long ‘weeks’ respectively. We find that the size and significance of our estimates are unchanged by binning our data into 4-day weeks. However, increasing the week length to 10 days does indeed cause a loss of variation in our dataset and results in estimates of θ in M1b, M3a and M3b that are no longer significant. Additionally, we find a significant negative estimate for M4a.

	Outcome	Est.	SE
All Users	M1a: No. of messages	0.345 **	0.12
	M1b: No. of communicators	0.104	0.07
	M2a: No. of contributors	-0.026	0.03
	M2b: No. of contributions	-0.157	0.08
New Users	M3a: No. of. messages	0.118	0.16
	M3b: No. of communicators	0.117	0.12
	M4a: No. of contributors	-0.136 *	0.06
	M4b: No. of contributions	-0.341 ***	0.10

Table 7: Estimates for the *msgwall* term in models where data is binned in ten day long 'weeks'.

	Outcome	Est.	SE
All Users	M2b: No. of contributions	0.011	0.07
No Admins	No. of contributions	-0.205 **	0.08
New Users	M4b: No. of contributions	-0.205 *	0.10

Table 8: Estimates for the *msgwall* term in models where contributions from administrators are dropped

6 Dropping Administrators

We were interested in understanding if administrators on Wikia were driving our findings. Given the sparse editor base of many Wikia wikis, it is often the case that the vast majority of activity in each wiki is driven by administrators. As we described above, we were also concerned that administrator activity that coincided with turning message walls on may have driven some of our measures.

Table 8 compares findings from M2b and and M4b (related to the number of contributions made by different user groups) with a model that estimates the effect of message walls on the contributions made by all non-administrative editors. We find that the estimates for non-administrative editors is virtually the same as the ones for newcomers. This demonstrates that the estimate we see in M2b in our full models is driven almost entirely by administrator activity.

	Outcome	Est.	SE
New Users	M3a: No. of. messages	0.424 **	0.16
	M3b: No. of communicators	0.308 **	0.12
	M4a: No. of contributors	-0.083	0.06
	M4b: No. of contributions	-0.196 *	0.10

Table 9: Estimates for the *msgwall* term in models where newcomers have edited for less than two months and made fewer than 20 edits.

	Outcome	Est.	SE
New Users	M3a: No. of. messages	0.415 **	0.16
	M3b: No. of communicators	0.29 *	0.12
	M4a: No. of contributors	-0.09	0.06
	M4b: No. of contributions	-0.193 *	0.10

Table 10: Estimates for the *msgwall* term in models where newcomers have edited for less than four months and made fewer than 20 edits.

7 Varying the Definition of Newcomers

As explained in the text of our paper, we define newcomers as any user having made fewer than 20 edits and having edited for less than 3 months. Although we think these are reasonable ways of measuring newcomer status that are in line with what other social computing researchers have done, these thresholds are necessarily arbitrary.

We sought to understand the extent to which our results were driven by the way we defined newcomers by rerunning the models in our analysis that pertained to newcomers (i.e. M3a, M3b, M4a and M4b) on datasets where we varied the definition of newcomer by altering either the time since account creation or the number of edits made so far.

Results from these models are shown in Tables 9, 10, 11, and 12. The

	Outcome	Est.	SE
New Users	M3a: No. of. messages	0.481 **	0.16
	M3b: No. of communicators	0.27 *	0.13
	M4a: No. of contributors	-0.109	0.06
	M4b: No. of contributions	-0.166	0.09

Table 11: Estimates for the *msgwall* term in models where newcomers have made fewer than ten edits in less than three months.

	Outcome	Est.	SE
New Users	M3a: No. of. messages	0.238	0.15
	M3b: No. of communicators	0.212	0.11
	M4a: No. of contributors	-0.081	0.06
	M4b: No. of contributions	-0.226 *	0.10

Table 12: Estimates for the *msgwall* term in models where newcomers have made fewer than thirty edits in less than three months.

ways in which we varied the definition of newcomer is detailed in the captions for each of the Tables. We find that in almost every case, variations to the definition of newcomer do not significantly alter our findings. We do see a difference in Table 12, where although M3a and M4a have positive estimates for the effect of message walls on their respective outcomes, they are no longer statistically significant.

References

[1] Lee, D. S., and Lemieux, T. Regression discontinuity designs in economics. *Journal of Economic Literature 48*, 2 (2010), 281–355.