-

1 Introduction

Online communities are used to aggregate and process information dispersed across many individuals. Pioneered in the 1980s, they have become more widespread with mass adoption of the Internet, and are now used across many different contexts in business [?, ?], politics and public decision making [?, ?, ?], expertise sharing [?, ?, ?], and education [?]. At the same time as they spread across domains, they did so geographically: for example, they have attracted large numbers of users and large venture capital investments in China [?] Most online communities lack a central command structure; despite this, many display remarkably coherent behaviour, and have proven effective at large tasks like writing the largest encyclopedia in human history (Wikipedia), providing an always-on free helpline for software engineering problems (StackOverflow), or building a detailed map of planet Earth (OpenStreetMap) [?].

Organizations running online communities typically employ community managers, tasked with encouraging participation and resolving conflict: this practice is almost as old as online communities themselves and predates the Internet [?], although it has become much more widespread as Internet access became a mass phenomenon. Though most participants to online communities are unpaid and answer to no one, a small number of them (only one or two in the smaller communities, many more in the larger ones) will recognize some central command, and carry out its directives. We shall henceforth call such directives *policies*.

Putting in place policies for online communities is costly. Professional community managers need to be recruited, trained and paid; software tools to monitor communities and make their work possible need to be developed and maintained. This raises the question of what benefits organisations running online communities expect from policies; and why they choose certain policies, and not others.

A full investigation of this matter is outside the scope of this paper; however, in what follows we outline and briefly discuss the set of assumptions that underpin our investigation.

1. In line with the network science approach to online communities, we model online communities as social networks of interactions across participants.

2. We assume that organisations can be modelled as economic agents maximising some objective function. The target variable being maximised can be profit (for online communities run by commercial companies); or welfare (for online communities run by governments or other nonprofit entities); or some combination of the two.

3. We assume that the topology of the interaction network characteristic of online communities affects their ability to contribute to the maximisation of the target variable.

4. We assume that such organisations choose their policies as follows:

   - Solve their maximisation problem over network topology. This yields a vector of desired network characteristics, where "desired" means that those characteristics define a maximum of the objective function. These solutions will be statements with the form "In order to best meet our ultimate [profit or welfare] goals, the interaction network in our online community should be in state Θ_D, where Θ is a vector of topology-related parameters".
   - Derive a course of action that community managers could take to change the network away from its present state Θ_0 to the desired state Θ_D.
   - Encode such course of action in a set of simple instructions for community managers to execute. Computer scientists might think of such instructions as algorithms; economists call them mechanisms; professional online community managers call them policies. In this paper we use this third term.

All this implies that the organisation running the online community has tools at its disposal to (attempt to) reach such preferred states of the interaction network. But rarely, if ever, is it explicitly argued that, in fact, it does. Network topology results from the behaviour of all the users of the online community. Most of them have joined the community by their own volition, can leave at any time, and are in no way answerable to the organisation. This makes topology an emergent property of the pattern of interaction. For recommendations about "better" states to be meaningful, we need to verify that organisations have the tools to achieve them.

This paper attempts to bridge that gap. To do so, it does not attempt to model the whole chain of decision starting from the maximisation problem. Rather, it focuses on those participants in an online community that *are* answerable to the organisation in charge of it: its online community managers. We conjecture the following:

1. Organisations can formulate policies and instruct online community managers to execute them.
2. Online community managers execute by communicating with users.
3. Communication with online community managers nudges users towards taking the course of action desired by the organisation.

The first two items in this list are assumed to fit into the mechanism design framework. This entails assuming that the organisation in charge of the online community knows both the desired network state Θ_D and the set of behaviours that, were they adopted by the online community's users, would result in it achieving Θ_D. We focus on the final item, which is equivalent to assuming that user behaviour is responsive to communication with online community managers. This assumption is implicit both in the literature and in managerial practice but has not, in practice, been proven.

We consider a policy called *onboarding*. As a new participant becomes active (for example by posting her first post, or commenting somebody else's post for the first time), professional community managers are instructed to leave her a comment that contains (a) friendly, positive feedback and (b) suggestions to engage with other, existing participants that she might have interests in common with[1], perhaps one of the simplest and most common in online community management [**?**, **?**]. A new participants, after her first contribution, would get a comment like this by one of the moderators:

"Welcome, Alice! That was a very interesting point. It definitely resonates with my own experience in the field. In our community, the people who are most involved in the matter are Bob [link] and Charlie [link]. You might be interested in this post [link] by Bob, where he relates his own experience: if you leave him a comment, I am sure an interesting conversation will ensue."

If the new participant engages (by replying or asking a question), she gets another message (an acknowledgement or an answer); after that, the new participant is generally considered onboarded, and is not made the object of any special attention.

In the rest of the paper, we consider the issue of whether user behaviour does, indeed, respond to onboarding. Organizations running online communities invest considerable resources in onboarding, and user responsiveness is an important parameter in making the decision of how much to spend in this activity as opposed to other, competing ones.

Section 2 briefly examines the literature that we mostly draw upon. Section 3 presents some data from a real-world online community called Edgeryders; then proceeds to describe a statistical model of that community's user behavior in the presence of onboarding. Section 4 presents the estimation's results. Section 5 discusses them. -

2 Literature survey

We are interested in the responsiveness of the behaviour of participants in online communities when policies, aiming to change the topology of the interaction network to a desired state, are enacted by the organisations in charge of the online community itself. This topic finds itself at the intersection of two different strands of academic literature.

The first strand is mainly interested in the determinants of the behaviour of individuals participating in online communities. Many scholars, moving from a psychology or business studies background, take the view that individual behaviour in online community both responds to social norms and contributes

[1] Katherina Fake, founder of Flickr, a popular website to share photographs, is reported to have deployed the company's employees as the website's first users and the initial core of its community. According to her "We learned you have to greet the first ten thousand users personally". [**?**]

to shape them. Several attempts have been made to conceptualize this view into testable models ([?], [?], [?]). Such attempts use survey data to estimate structural equation models. The models' parameters quantify the influence of both individual incentives (such as obtaining valued information) and social incentives (such as group norms) in participating in online communities.

The second strand moves from a social network science background, and looks rather at the interplay between network topology and user behaviour. The best-known example is probably Ronald Burt's theory of structural holes ([?]), which generalises to any community, both online and offline ones. There exist several studies dedicated to specific online communities, such as Slashdot ([?]), Java developers forums ([?]) and Linux ports developer forums ([?]). Some authors have attempted to merge these two strands of literature, treating network topology characteristics as variables to incorporate into their structural equation models ([?],

Authors from both strands agree that some topological characteristics are more conducive than others to the organisation's goals. For example, Burt ([?]) suggests that densely connected clusters of individuals are useful to better focus on goals and targets, whereas less dense networks with some individuals bridging across clusters are more conducive to innovation; Ganley and Lampe ([?]) propose that densely connected clusters of "power users" lead to social tensions and a loss of the egalitarian spirit that makes many online communities attractive; Kim and collaborators ([?]) show that, in some circumstances, a tendency towards communication reciprocity ("intimacy") leads to membership loss and, potentially, network breakdown. Dholakia and collaborators are perhaps clearest in indicating that their findings are relevant to designing and enacting policies:

> Understanding the antecedents of social influence is important since it is likely to provide significant managerial guidance regarding how to make virtual communities useful and influential for their participants. ([?], p. 242)

Most authors do not explicitly envision pathways leading from the formulation of a policy to attaining the desired change in network topology. Some do point to the relevance of online community managers, explicitly ([?]) or implicitly ([?]), by referring to cases in which such professionals are prominent ([?]); but we have been unable to find an explicit discussion of how few online community manager can wilfully alter the behaviour of the many participants to the same community. Departing from this practice, in a previous paper ([?]), we have proposed such a pathway for the onboarding policy: online community managers communicate one-to-one with new members of the online community and suggest they initiate communication with existing members, indicating those existing members that seem to match the newcomer's interests best. This paper provides an empirical test that such behaviour does indeed prompt users to interact with each other, underpinning the usefulness and profitability of engaging in online community management activities.

The present paper is also related to the growing mass of literature on the treatment of time in network analysis. Early work in networks, both in mathematics (for example [?]) and in sociology (for example [?]), focused on the topology of static networks. The beginning of the 21st century marked a surge of interest in evolving networks (for example [?] and related work; see [?] for a survey). Scholars sharing this interest investigate the growth paths that produce certain notable topologies often encountered in real-world networks. Later still, a literature on dynamic networks developed, interested in dynamical processes that happen in networks, like information diffusion (for example [?] or epidemics (for example [?]). Many of these studies direct their attention towards the statistical properties of the sequences of event that make up spreading dynamics. Human communication dynamics, they find, turns out to be ”bursty”, with the timings between communication events deviating significantly from uniform or Poissonian statistics (see [?] for a survey).

Our focus does not lie in the dynamics of communication events in online communities. Rather, we wish to establish whether community managers can really influence the behaviour of participants in the online community. The pattern of relationships across the community is likely to have some influence on users' willingness to initiate communication. Since such patterns varies across time, we adopt the stance of representing temporal data as a sequence of static graphs, tracing the evolution of the interaction network. This allows us to treat the topology of the network as a control variable, whose effect on the users likelihood to initiate communication must be kept separate from that of the action of community managers. It is a common enough technique, deemed appropriate when, as is our case, topology is more central to the analysis than time sequence ([?]). -

3 Materials and methods

3.1 Data

We consider an online community called Edgeryders, whose members discuss and try to implement various projects connected to social innovation in a decentralized manner. The community is run by a British company, also called Edgeryders, which we cast as the policy maker of the community. The affordances of the online platform that hosts the community and its discussion deeply influence the social dynamics thereof ([?]). Those of Edgeryders are as follows: the discussion is hosted on a web platform based on an open source framework called Drupal 7. Drupal is one of the largest open source projects globally; in this sense, Edgeryders can be though of as fairly typical from a technology point of view. It consists of users posting reflections or status reports on their social innovation projects; these are then commented upon by other members of the community. The Edgeryders community website supports what is known in online community parlance as threaded comments: in other words, comments are themselves commentable. Of course, posts are commentable too, so that a comment can be directed either towards a post or towards another comment. Both posting and commenting are restricted to registered users only. This architecture has not changed since inception of Edgeryders in October 2011.

We extracted a snapshot of the database on October 7th 2015. At the time, the community had 2,904 registered users, who had authored 4,062 posts and 18,285 comments. 23 of the users had, at some point in time, reported to the Edgeryders company; we cast them as the online community managers for this particular community.

All events in our dataset (creating accounts, creating posts, creating comments) are timestamped. This allows for the possibility of modelling interaction in the online community dynamically.

3.2 A network model of interaction in online communities

We need to consider carefully how we incorporate dynamics in our model ([**?**]). We make the following assumptions:

1. The duration of interactions is negligible. This means that events can be represented by a contact sequence, i.e. triples (i, j, t) where $i, j \in V$, the set of interactants (vertices in the network) and t denotes time. This is a standard assumption in studies of person-to-person communication online ([**?**]).

2. First-time interaction produces a permanent change in the social relationship between the source and the target of the interaction (the author and the recipient of the comment). All Edgeryders users can be assumed to share some common interests around social innovation, that motivated them to join Edgeryders. When they interact directly, however, they begin to unpack and specify such common interests, and they (often explicitly) signal interest in each other's activities. This produces a shift in the nature relationship between the interactants, which becomes one of actual, as opposed to potential, interaction.

3. Subsequent interaction also carry social significance. They have the effect of further strengthening the relationship between the two interactants. The strength of the relationship increases monotonically with the number of interactions recorded. The effect of interactions is permanent.

We then proceeded to build a social network of the interactions within the community as follows:

1. Users that have posted at least one post or one comment are included as nodes in the network.

2. We interpret each comment as one interaction event. Whenever a comment is posted, an edge is induced. The new edge's source node is the node associated to the author of the comment. The new edge's target node is the node associated to the author of the post or comment being commented.

3. If an edge with the same source and target already exist, we do not add a new one, but rather increase the edge's weight by one.

4. If the source and target of the edge are the same, the edge is discarded.

5. The sequence of timestamped interaction events allows us to model interaction in Edgeryders as an evolving network. We divide the observation period in one week-long intervals, and consider users to be acting on the basis of the interaction network that they find themselves in at that period. In other words, rather than modelling explicitly the dynamics on the network, we focus on the time evolution of the network's static structure, and assume such static structure to influence user behaviour. This reflects our preoccupation with topological aspects over temporal ones ([**?**], [**?**]).

This process results in a sequence of 212 networks, each one a step on the growth path of the interaction network in Edgeryders. At the end of week 1 the network had only 6 nodes and 1 edge; at the end of week 212 it had 789 nodes and 4,861 edges.

3.3 Model

We model user activity in an online community as a function of three groups of variables:

$$A_{i,t} = f(NN_{i,t}, EN_{i,t}, GN_t) \tag{1}$$

Where:

- $A_{i,t}$ denotes the activity of user i at time t, defined as the number of posts plus the number of comments authored by user i during the period.

- $NN_{i,t}$ denotes a vector of non-network variables associated with user i at time t. These include writing a post or a comment; receiving a comment from a community manager; receiving a comment from another user who is not a community manager; the total number of posts and comments written by other users users $A_{j,t}, j \neq i$; and i's experience as a user of Edgeryders, as measured by the number of weeks elapsed since she joined the community. The week of joining can predate the week in which the user writes her first post or comment.

- $EN_{i,t}$ denotes a vector of variables pertaining to the user's ego network at time t. These include her own in-degree (the number of users linking to her); her out-degree (the number of users she links to); her betweenness centrality (the fraction of shortest paths across any two users $j, k \neq i$ that i finds herself on); her PageRank (the probability that a random walker across the network would end up at this particular node[2]; and her clustering coefficient (the fraction of i's neighbours that are also neighbours to each other).

[2] PageRank was originally developed as a measure of network centrality for the World Wide Web ([**?**]).

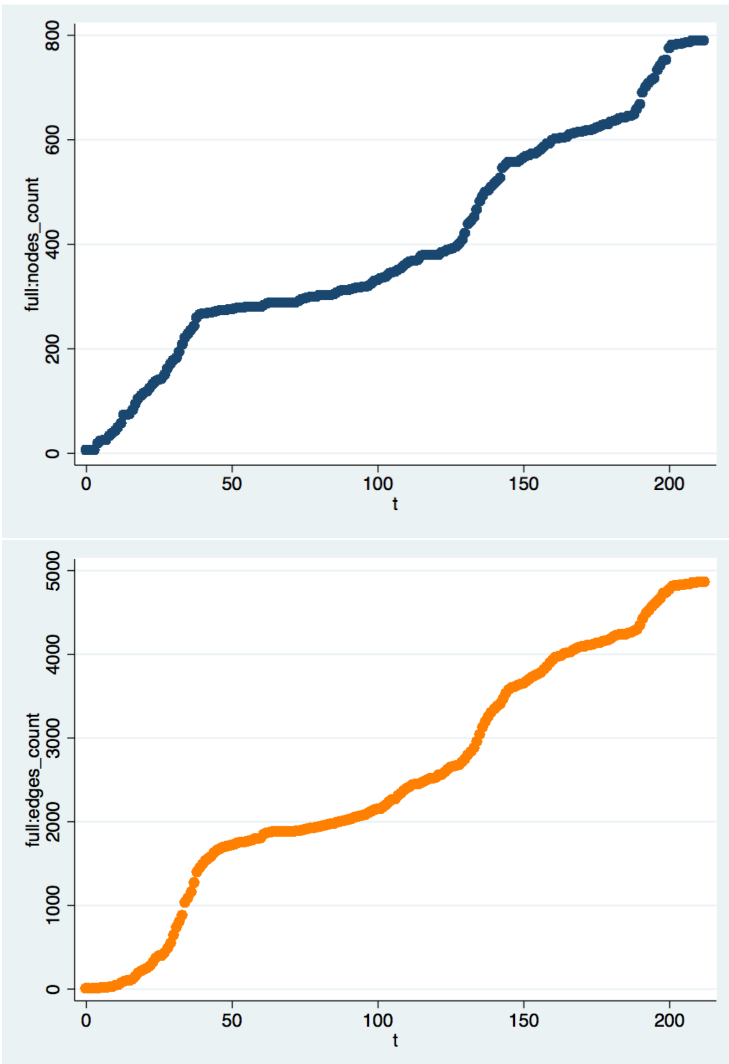

Figure 1: Growth of the number of nodes (top) and edges (bottom) in the Edgeryders interaction network.

- $GN_{i,t}$ denotes a vector of variables pertaining to the global network at time t. These include the total number of nodes; the total number of edges; the average degree; and the network's Louvain modularity[3].

3.4 Estimation

We estimate the behaviour of a binary variable $A_{i,t}$ that take value 1 if user i engages actively in interaction (i.e. writes a post or a comment at time t), and 0 otherwise. We make use of a logit model:

$$PR(A_{i,t} = 1|X_{i,t}) = \frac{e^{X_{i,t}\beta}}{1 + e^{X_{i,t}\beta}} \tag{2}$$

Where β_0 is a constant term; $X_{i,t}$ is a vector of observed explanatory variables; β is a vector of parameters.

An advantage of the logit model is that its coefficient lend themselves to being interpreted in terms of marginal effects on the log-odds ratio ([**?**]). We rewrite equation 2 as

$$ln\frac{Pr(A_{i,t} = 1)}{1 - Pr(A_{i,t} = 1)} = X_{i,t}\beta \tag{3}$$

Next, we proceed to estimating equation 3 with fixed effects on the 766 Edgeryders users who have no moderator or site administrator role. To do this, we drop the 23 users that have such roles. We add to the right hand-side c_i, an unobserved time-invariant effect specific to user i and $u_{i,t}$ is a residual error term, with mean zero and uncorrelated with right-hand side variables. We take care to employ lagged variables when failure to do so might cause endogeneity issues.

$$ln\frac{Pr(A_{i,t} = 1)}{1 - Pr(A_{i,t} = 1)} = X_{i,t}\beta + c_i + u_{i,t}$$

3.5 Software stack

Primary data from the Edgeryders database were obtained via a call to the platform's APIs. We used a modified version of a software called Edgesense[4]. Modifications concerned adding support for computing some extra network metrics, like the clustering coefficient. to extract data from then Edgeryders database. We then enriched the data so obtained with non-network metrics computed directly from the primary data, and exported the results in tabular form. Such results were then imported into Stata for estimation. -

[3]The modularity of a network is a measure of how much it differs from an Erdos-Renyi random network with the same degree distribution. Values close to 0 indicate it is indistinguishable from a random network; values close to -1 or 1 indicate structure ([**?**]). Measuring modularity is computationally hard; it is customary to use algorithms to compute an approximate value. Of these, the Louvain algorithm is the most widely used in the literature for its attractive computational efficiency [**?**].

[4]https://github.com/Wikitalia/edgesense

4 Results

4.1 Hypotheses

We wish to test the following hypotheses.

Hypothesis 1 *The more comments a user receives from community managers, the higher her probability to become active.*

The organisations running online communities wish for their users to do certain things, most of which imply being active in the community itself: writing posts and comments. They cannot order them to do so, since users are not on the payroll and remain unanswerable to those organisations. Online community managers are then tasked to prompt users into action without using either monetary incentives or command power. That leaves interaction as the main tool online community managers have at their disposal. Rejecting Hypothesis 1 would imply that users do not respond to cues from online community managers.

Hypothesis **??** requires that the coefficient on the number of comments received by the user from community managers be positive. Denote said coefficient by β_{cmrec}, and the related variable by x_{cmrec}. It is immediate to check that:

$$\frac{\partial Pr(A=1|x)}{\partial x_{cmrec}} > 0 \Rightarrow \frac{\partial LOR}{\partial x_{cmrec}} > 0 \tag{5}$$

where LOR denotes the logarithm of the log-odds ratio as per equation **??**. The same equation implies that:

$$\frac{\partial LOR}{\partial x_{cmrec}} > 0 \Rightarrow \beta_{cmrec} > 0 \tag{6}$$

Hypothesis 2 *Receiving comments from community managers has a larger effect on the probability to become active than receiving comments from users that are not community managers.*

Online community managers are professionals. They are likely to have better communication skills than the average user, and they certainly have stronger incentives to craft their interaction modes so as to drive users to being more active in the online community. We therefore expect, on average, the effect of incoming communication from online community manager to have a larger positive effect on the probability of becoming active than that of incoming communication from other users.

Hypothesis 2 requires that the difference between the coefficient on the number of comments received by the user from community managers and the coefficient on the number of comments received by other (non-community managers) users be positive. Formally, denote the latter coefficient by β_{urec}, and the related variable by x_{urec}. Since the transformation from probability to odds is monotonic, we have:

$$\frac{\partial Pr(A=1|x)}{\partial x_{cmrec}} > \frac{\partial Pr(A=1|x)}{\partial x_{urec}} \Rightarrow \frac{\partial LOR}{\partial x_{cmrec}} > \frac{\partial LOR}{\partial x_{urec}} \tag{7}$$

Applying again equation **??**:

$$\frac{\partial LOR}{\partial x_{cmrec}} > \frac{\partial LOR}{\partial x_{urec}} \Rightarrow \beta_{cmrec} > \beta_{urec} \Rightarrow \beta_{cmrec} - \beta_{urec} > 0 \tag{8}$$

4.2 Regression

Table 1 summarizes the estimation's results.

Variable	**Coefficient**	**(Std. Err.)**
number of comments received from community managers	1.250^{**}	(0.053)
number of comments received from other users	0.215^{**}	(0.037)
number of comments and posts written by ego (lagged)	$0.126^{\dagger}$	(0.066)
weeks since creating the account	$-0.011^{\dagger}$	(0.006)
total number of posts and comments written by users excluding ego	0.037^{**}	(0.004)
total number of posts and comments written by community managers	0.001	(0.006)
user out-degree (lagged)	0.028^{*}	(0.013)
user in-degree (lagged)	-0.003	(0.013)
user betweenness centrality	-107.870^{**}	(20.778)
user pagerank (lagged)	-5.592	(14.610)
user clustering coefficient	-0.654^{**}	(0.195)
network average distance	-4.156^{*}	(2.007)
network average betweenness centrality at $t-1$	-9.260	(209.212)
network Louvain modularity	8.379^{**}	(2.520)
number of nodes	-0.001	(0.004)
number of edges	0.000	(0.001)

Table 1: Estimation results

The coefficients on the first two variables are positive and highly significant ($p < 0.001$). This supports the conventional wisdom that users of online communities tend to engage with each other: when made the object of comments, they are more likely to become active than when they are not. By implication, our Hypothesis **??** cannot be rejected.

To test Hypothesis 2, we start by noting that the coefficient on the number of comments received from community managers is larger than that on the number of comments received from other (non-community managers) users. We next run a Wald test on the null hypothesis that the coefficients on our first two variables are identical. The null is strongly rejected ($p < 0.0001$). This, in turn, means we cannot reject Hypothesis 2.

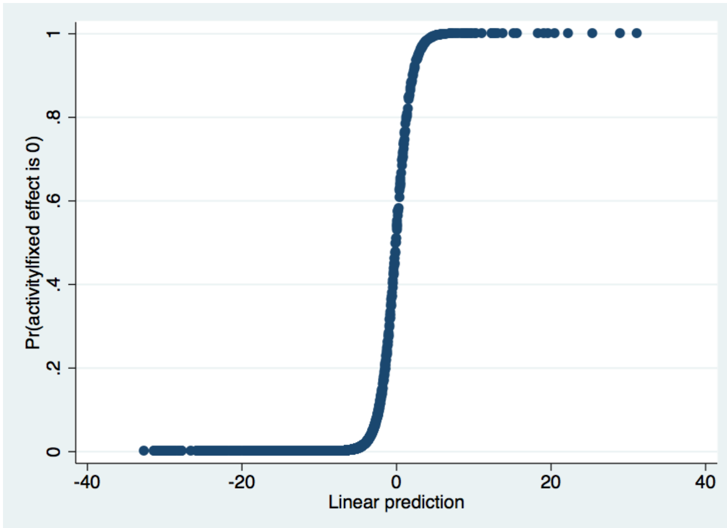

Figure 2: Predicted probability of Edgeryders users to become active as the linear prediction $x\beta$ based on number of comments received by community managers and other users increases.

4.3 Marginal effects

Regression analysis shows that the activity of community managers does indeed have a positive impact on the probability that the users they engage will become active. This, however, does not tell us how large that impact is. Since online community management is costly, this is likely to be a question of some relevance to an organisation trying to make a decision to invest in it. Coefficient estimates are a poor indicator of marginal effects, because the logit model we employ here is not linear. The relationship between the linear prediction (as defined by the sum of each coefficient multiplied by the regressor it refers to, and assuming the fixed effects are zero) and the probability that the dependent variable is equal to 1 follows a logistical curve (figure 2). An increase in the number of comments received by community managers has a small effect for small or large values of $x\beta$. When $x\beta$ is between -5 and 5, however, the probability that the user becomes active increases quickly as community managers engage more with the user.

Table 2 shows a point estimate the marginal effect of comments from the two sources (online community managers vs. other users) on the probability that a user will become active. Neither is significant. This turns out to be an artefact of the computation: the size of the marginal effect is computed fixing the value of regressors at their mean. The means of the variables in question are low. In the average period, the average Edgeryders user received 0.08 comments from online community managers and 0.09 comments from other users. This is consistent with what we know about patterns of human communication, which is sparse and bursty ([?]).

Variable	dy/dx	$Std.Err.$	$P > \|z\|$
number of comments received from community managers	.0000959	.0004127	0.816
number of comments received from other users	.0000165	.000071	0.816

Table 2: Marginal effects of the number of comments received by a user (both from community managers and from other users) on the probability of that user to become active. The estimates are computed under the assumption that regressors be fixed at their means.

A more intuitive approach is to estimate the elasticities of the probability of becoming active with respect to the number of comments received from each source. These are shown in table 3. They are both highly significant. In absolute terms, they are both small, but quite different. Receiving an extra comment by a community manager increases the probability that a user will become active by 10%; but receiving one from another user will increase it by less than 2%.

Variable	ey/ex	$Std.Err.$	$P > \|z\|$
number of comments received from community managers	.1024245**	.0043113	0.000
number of comments received from other users	.01882**	.0032799	0.000

Table 3: Elasticities of probability that a user becomes active with respect to the number of comments received from community managers and from other users. The estimates are computed under the assumption that regressors be fixed at their means.

This estimation procedure turns out to have a flaw. The mean of the predicted probability of a positive outcome (defined as the user becoming active in the period) does not coincide with the proportion of users actually active across the dataset. This is accounted for by the fact that $x\beta = \sum_i x_i\beta_i$ is different from zero (table 4).

Variable	**Obs**	**Mean**	**Std. Dev.**
prob	84262	.0023325	.0410453
active	84262	.0181577	.1335221
$x\beta$	84262	-9.[illegible]17972	1.502639

Table 4: Descriptive statistics for the probability of users to become active as predicted by the model (*prob*); users actually being active (*active*); and the model's linear predictions ($x\beta$).

To correct this, we need to estimate a new logit model:

$$PR(A_{i,t} = 1|X_{i,t}) = \frac{e^{X_{i,t}\beta+\alpha}}{1 + e^{X_{i,t}\beta+\alpha}} \tag{9}$$

where α is a constant chosen so that

$$\frac{e^{mean(X_{i,t}\beta)+\alpha}}{1+e^{mean(X_{i,t}\beta)+\alpha}} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} active_{i,t}}{NT} \tag{10}$$

where $active_{1,t}$ takes value 1 if user i is active at period t, and 0 otherwise. The corrected model's linear predictions are unbiased, allowing us to compute marginal effects correctly. The right-hand side of equation 10 is identical to the mean of the variable *active* in table 4. Replacing the appropriate values from table 4 yields:

$$\frac{e^{-9.52+\alpha}}{1+e^{-9.52+\alpha}} = 0.18 \tag{11}$$

$$\alpha \simeq 8 \tag{12}$$