Online community management as social network design: testing for the signature of management activities in online communities

ALBERTO COTTICA*   GUY MELANÇON

University of Alicante, Edgeryders   University of Bordeaux
alberto@edgeryders.eu

BENJAMIN RENOUST

National Institute of Informatics, Tokyo

Abstract

Online communities are used across several fields of human activities, as environments for large-scale collaboration. Most successful ones employ professionals, sometimes called "community managers'" or "moderators", for a variety of tasks including onboarding new participants, mediating conflict, and policing unwanted behaviour. Network scientists routinely model interaction across participants in online communities as social networks. We interpret the activity of community managers as network design: they take action oriented at shaping the network of interactions in a way conducive to their community's goals. It follows that, if such action is successful, we should be able to detect its signature in the network itself.

Growing networks where links are allocated by a preferential attachment mechanism are known to converge to networks displaying a power law degree distribution. Growth and preferential attachment are both reasonable first-approximation assumptions to describe interaction networks in online communities. Our main hypothesis is that managed online communities are characterised by in-degree distributions that deviate from the power law form; such deviation constitutes the signature of successful community management. If true, this hypothesis would give us with a simple test for the effectiveness of community management practices. Our secondary hypothesis is that said deviation happens in a predictable way, once community management practices are accounted for.

We investigate the issue using (1) empirical data on three small online communities and (2) a computer model that simulates a widely used community management activity called onboarding. *We find that the model produces in-degree distributions that systematically deviate from power law behaviour for low-values of the in-degree; we then explore the implications and possible applications of the finding.*

I. INTRODUCTION

Organizations running online communities typically employ community managers, tasked with encouraging participation and resolving conflict [**?**]. Only a small number of the participants (one or two members in the smaller communities) will recognize some central command, and carry

---

*Corresponding author

out its directives. We shall henceforth call such directives *policies*. Putting in place policies for online communities is costly, in terms of recruitment, training, and software tools. This raises the question of what benefits organizations running online communities expect from policies; and why they choose certain policies, and not others. In what follows we outline and briefly discuss the set of assumptions that underpin our investigation.

1. We model online communities as social networks of interactions across participants.

2. We assume that organisations can be modelled as economic agents maximizing 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. Indications that this assumption might be reasonable are not difficult to find in the literature ([**?**],[**?**]).

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 decision to deploy a particular policy on an online community is a network design exercise. An organisation decides to employ a community manager to shape the interaction network of its community in a way that helps ist own ultimate goals. And yet, interaction networks in online communities cannot really be designed; they are the result of many independent decisions, made by individuals who do not respond to the organization's command structure. An online community management policy is then best understood as an attempt to "influence" emergent social dynamics; to use a more synthetic expression, it can be best understood as the attempt to design for emergence. Its paradoxical nature is at the heart of its appeal.

We are interested in detecting the mathematical signature of specific policies in the network topology. We consider a simple policy called *onboarding* [**?**, **?**]. As a new participant becomes active (*e.g.* by posting her first post), community managers are instructed to leave her a comment that contains (a) positive feedback and (b) suggestions to engage with other participants that she might share interests with.

We model online conversations as social networks, and look for the effect of onboarding on the topology of those networks. We proceed as follows:

1. We initially examine data from three small online communities. Only two of them deploy a policy of *onboarding*. We observe that, indeed, the shape of the degree distribution of these two differs from that of the third.

2. We propose an experiment protocol to determine whether onboarding policies can explain the differences observed between the degree distributions of the first two online communities and that of the third one.

3. Based on a generalized preferential attachment model [?], we simulate the growth of online communities. Variants to the model cover the relevant cases: the absence of onboarding policies and their presence, with varying degrees of effectiveness.

4. We run the experiment protocol against the degree distributions generated by the computer model, and discuss its results.

Section 2 briefly examines the two strands of literature that we mostly draw upon. Section 3 presents some data from real-world online communities; it then proceeds to describe our main experiment, a computer simulation of interaction in online communities with and without onboarding. Section 4 presents the experiment's results. Section 5 discusses them.

II. Related works

The extraordinary successes of online communities in deploying large-scale, decentralized projects has led many scholars to conjecture that online communities exhibit emergent behavior, and called such behavior collective intelligence, after an influential book by Pierre Lévy [?]. This name was adopted by a research community that aims at providing tools for better collective sense- and decision making such as argument maps (representations of the logical structure of a debate, with all redundancy eliminated) [?] and attention-mediation metrics (indicators that signal what, in an online debate, is worthiest reading and responding to. The number of Likes on Facebook is one such metric) [?].

Collective intelligence scholars confirmed importance of online community management practices, indeed, they have tried to systematize it [?] and produce technological innovation to support it [?, ?]. These tools are meant to facilitate and encourage participation to online communities, to make it easier for individuals to extract knowledge from them. Studying human communities is a traditional focus of network science [?, ?], for which easily available datasets of online communities make an ideal ground for structural analysis: friendship in Facebook [?, ?], following/retweet/mentions for Twitter [?, ?, ?], or vote and comments in discussions [?, ?, ?, ?].

Starting in the 2000s, online communities became the object of another line of enquiry, stemming from network science. Network representation of relationships across groups of humans has yielded considerable insights in social sciences since the work of the sociometrists in the 1930s, and continues to do so; phenomena like effective spread of information, innovation adoption, and brokerage have all been addressed in a network perspective [?, ?]. As new datasets encoding human interaction became available, many online communities came to be represented as social networks. This was the case for social networking sites, like Facebook [?, ?]; microblogging platform like Twitter [?, ?, ?]; news-sharing services like Digg [?]; collaborative editing projects like Wikipedia [?]; discussion forums like the Java forum [?]; and bug reporting services for software developers like Bugzilla [?]. Generally, such networks represent participants as nodes. Edges represent a relationship or interaction. The nature of interaction varies across online communities: one edge can stand for friendship for Facebook; follower-followed relationship, retweet or mention in Twitter; vote or comment in Digg and the Java forum; talk in Wikipedia; comment in Bugzilla.

In contrast to collective intelligence scholars, network scientists typically do not address the issue of community management, and treat social networks drawn from online interaction as fully emergent. In this paper, we employ a network approach to investigate the issue of whether

the work of community managers leaves a footprint detectable by quantitative analysis. To our knowledge, no other work attempted this investigation. In particular, we exploit a result from the theory of evolving networks, from seminal work by Barabási and Albert [?] showing that the assumption of growth and preferential attachment, when taken together, result in a network whose degree distribution converges to a power law ([?, ?]). The model was later generalized in various ways and tested across a broad range of networks, including social networks [?].

We use this generalized model as a baseline. The degree distribution of the interaction network in an online community follows a power law by default. The action of online community managers, as they attempt to further the goals of the organisation that runs the online community, will result in its degree distribution deviating from the baseline power law in predictable ways. Such deviation can be interpreted as the signature that the policy is working well.

The most important difficulty with this method is the absence of a counterfactual: if a policy is enacted in the online community, the baseline degree distribution corresponding to the absence of the policy is not observable, and viceversa. This rules out a direct proof that the policy "works". Hence our choice to combine empirical data and computer simulations.

III. Materials and methods

In this section we introduce the empirical data, the experiment protocol and the simulation model we use in the experiment.

i.   Empirical data

We examine data from three real-world online communities. All three use the same software (Drupal 7), are roughly comparable in size and are used by practitioners and interested citizens to publicly discuss issues that have a collective dimension. They are modelled as interaction networks, in which nodes are registered users and edges represent comments. The presence of an edge from Alice to Bob indicates that Alice has commented content authored by Bob at least once. The resulting graphs are directed ("Alice comments Bob" is not equivalent to "Bob comments Alice") and weighted (Alice can write multiple comments to Bob's content; the edge's weight is equal to the number of comments written). Table **??** presents some descriptive statistics about them.

- InnovatoriPA is a community of (mostly) Italian civil servants discussing how to introduce and foster innovation in the public sector. It does not employ any special onboarding or moderation policy.
- Edgeryders is a community of (mostly) European citizens, discussing public policy issues from the perspective of grassroot activism and social innovation. It enacts the onboarding of new members policy.
- Matera 2019 is a community of (mostly) citizens of the Italian city of Matera and the surrounding region, discussing the city's policies. It, too, enacts an onboarding policy.

The communities are modeled as interaction networks (summarized in Table **??**) in which nodes are users and edges represent directed comments from A to B, weighted by the number of comments written. A glance at their respective visualizations (Figure **??**) suggests that the networks of the three communities have very different topologies. Innovatori PA displays more obviously visible hubs than the other two.

We fitted power laws in-degree distributions of these three online communities, as of early December 2014. Next, we tested the hypothesis that degree distributions follow a power law, as

	Innovatori PA	Edgeryders	Matera2019
Policy	*"no special policy"*	*"onboard new users"*	*"onboard new users"*
In existence since	December 2008	October 2011	March 2013
Accounts created	10,815	2,419	512
Active participants (nodes)	619	596	198
Number of edges (weighted)	1,241	4,073	883
Average distance	3.77	2.34	2.51
Maximum degree	155	238	46
Average degree	2.033	6.798	4.454
Goodness-of-fit for $k \geq 1$			
exponent	1.611	1.477	1.606
p-value	0.21	0.00 (reject)	0.00 (reject)
Goodness-of-fit for $k \geq k_{min}$			
k_{min}	2	5	6
exponent	1.834	2.250	2.817
p-value	0.76	0.45	0.94

Table 1: *Comparing interaction networks of the three online communities and testing for goodness-of-fit of power functions to degree distributions. "Exponent" refers to the power law's scaling parameter. "p-value" to the result of the test that the degree distribution of the community was generated by a power law with that exponent.*

predicted by [?]. To do so, we first fitted power functions to the entire support of each in-degree distribution[1]. We next fitted power functions to the right tail of each in-degree distribution, *i.e.* for any degree $k(n) \geq k_{min}$, where k_{min} is the in-degree that minimizes the Kolmogorov-Smirnov distance (hereafter denoted as D) between the fitted function and the data with in-degree $k \geq k_{min}$.

Finally, we ran goodness-of-fit (hereafter *GoF*) tests for each in-degree distribution and for fitted power functions. The method we followed throughout the paper is borrowed from Clauset *et al* [?]. The null hypothesis tested is that the observed distribution is generated by a power function with exponent α. We compare the D statistic of the observed distribution with those of a large number of synthetic datasets drawn by the fitted power function. Such comparison is summarized in a *p*-value, that indicates the probability of the D statistic to exceed the observed value conditional to the null hypothesis being true. *p*-values close to 1 indicate that the power function is a good fit for the data: the null hypothesis is not rejected. *p*-values close to zero indicate that the power function is a bad fit for the data, and reject the null hypothesis. The rejection value is set, conservatively, at 0.1. Results are summarized in Table **??**.

As we consider the interval $k \geq 1$, we find that the in-degree distribution of the Innovatori PA network – the unmoderated one – is consistent with the expected behavior of an evolving network with preferential attachment. We cannot reject the null hypothesis that it was generated by a power law. For other two communities, both with onboarding policies, the null hypothesis is strongly rejected. On the other hand, when we consider only the tail of the degree distributions, i.e. $k \geq k_{min}$, all three communities display a behavior that is consistent of a setting with preferential attachment.

These results are consistent with the objectives of the onboarding policy, consisting in helping newcomers find their way around a community that they don't know yet. A successfully onboarded

---

[1]We emphasize in-degree, as opposed to out-degree, because directedness is implicit in the idea of preferential attachment, and because the in-degree distribution is the one to follow a power law in online conversation networks ([?]).
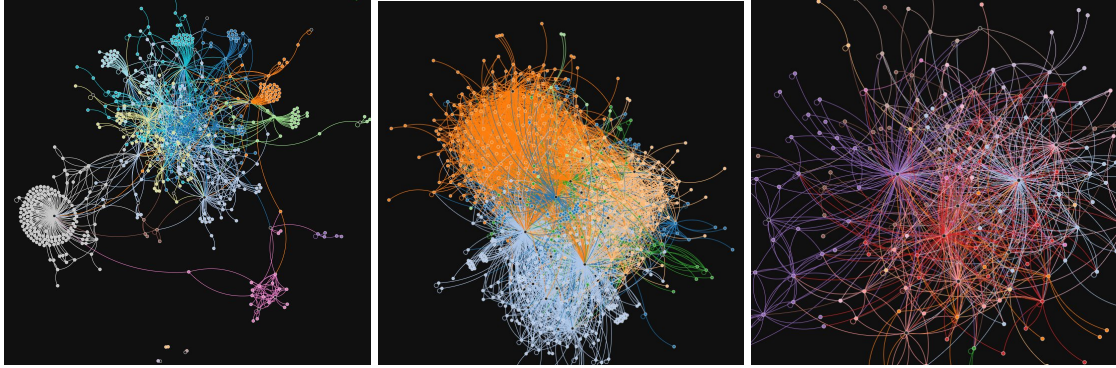
Figure 1: *Interaction networks of three small online communities. Innovatori PA (left) does not have an onboarding policy in place, whereas the two others do (Edgeryders: center, Matera: right).*

new user will generally have some extra interaction with existing active members. All things being equal, we can expect extra edges to appear in the network, and interfere with the in-degree distribution that would appear in the absence of onboarding – explaining the non-power law distribution of Edgeryders and Matera2019. Extra edges target mostly low connectivity nodes: onboarding targets newcomers, and focuses on helping them through the first few successful interactions. Highly active (therefore highly connected) members do not need to be onboarded. This may explain why all three communities display power law behavior in the upper tail of their in-degree distributions, regardless of onboarding.

ii.   Experiment protocol

The difference observed between the two communities with onboarding policies and the one without might be caused not by the policy itself, but by some other unobserved variable. To explore the issue further, we generate and compare computer simulations of interaction networks in online communities that are identical except for the presence and effectiveness of onboarding policies. Communities are assumed to grow over time, with new participants joining them in sequence; at each point in time, new edges appear; their probability of targeting an existing node grows linearly with that node's in-degree. Additionally, communities might have or not have onboarding policies. For those communities that do have them, they are modelled by means of two scalar parameters ν_1 and ν_2, that vary between 0 and 1. The first one captures onboarding effectiveness; the second one captures community responsiveness. As they get closer to 1, the community manager's onboarding action gets closer to having the desired effects. In the next subsection, we specify the model and define more specifically the meaning of both parameters.

We proceed as follows.

First, we simulate the evolution of the interaction network of a large number of online communities. Divide them into a control group (no onboarding policy) and a treatment group (presence of onboarding policy). Specifically, we simulate the evolution of the interaction network of:

- 100 communities with no onboarding policy. These will constitute the control group of our simulated communities.
- 100 communities for each couple of values of ν_1 and ν_2, with $\nu_1, \nu_2 \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$. These will constitute our treatment groups.
- For each of these networks, we compute the in-degree distribution.
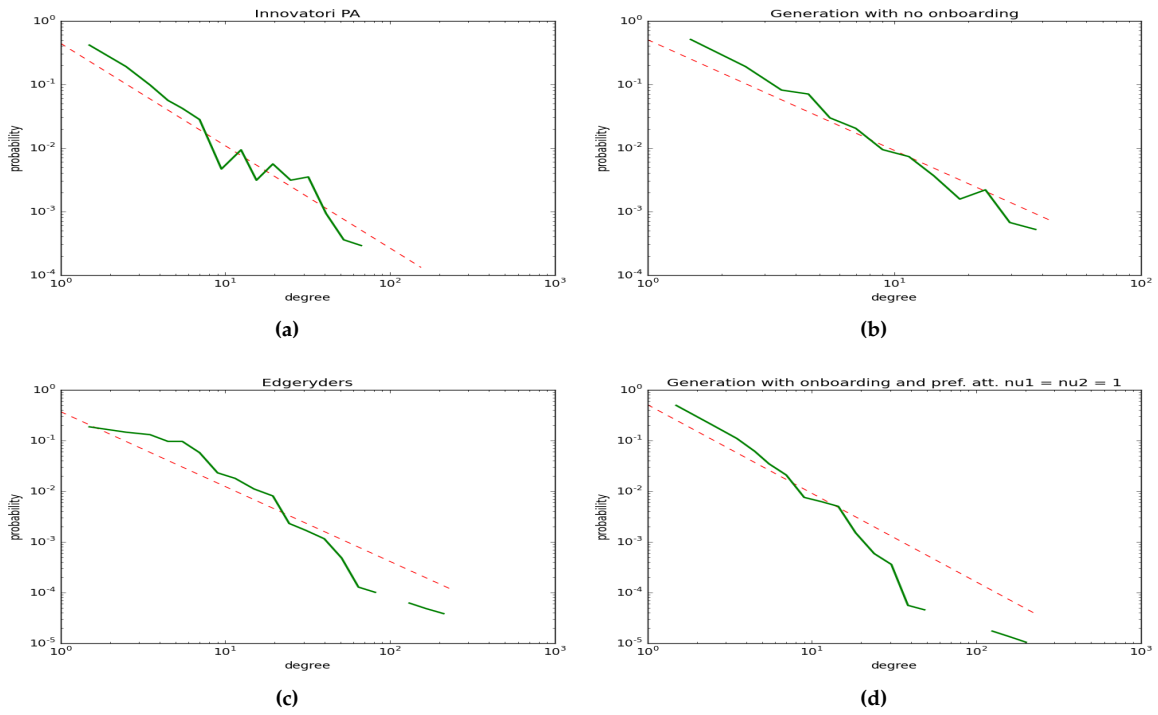
Figure 2: *(log - log) Probability density function from the degree distributions of: (a) the Innovatori PA network without onboarding policy in place versus (b) a simulated network with preferential attachment and no onboarding. (c) The Edgeryders network with onboarding and preferential attachment versus (d) a simulated network with preferential attachment and fully effective onboarding ($v_1 = v_2 = 1$).*

Next, we define the following hypotheses.

- Let C be the network of interaction in an online community. Denote the in-degree of nodes in the network by k. Let F be the best-fit power-law model for the distribution

$$q(k) = k + mA \qquad (1)$$

  where $k($ the in-degree distribution of C, m the number of nodes that join the network at each timestep and A a node attractiveness parameter.
- *Hypothesis 1*. The distribution of $q(k)$ is generated by F for any $k > 1$.
- *Hypothesis 2*. The distribution $q(k)$ is generated by F for any $k \geq k_{min}$, where k_{min} is the in-degree that minimizes the Kolmogorov-Smirnov distance between the fitted function and the data over $k \geq k_{min}$.

Both hypotheses are based on the asymptotic form taken by the stationary in-degree distribution of networks growing by preferential attachments in [**?**]. The result holds even if preferential attachment is not the sole mode of network evolution, and for any edge sources.

Finally, we test Hypothesis 1 and 2 on each of the 3700 in-degree distributions generated. We do this using the goodness-of-fit tests proposed by Clauset et.al. [**?**] and illustrated in detail in the Appendix. We expect to obtain the following:

- In the control group, both Hypothesis 1 and Hypothesis 2 are true.
- In the treatment group with fully effective onboarding Hypothesis 1 is false and Hypothesis 2 is true.
- In the intermediate situations of partially ineffective onboarding, Hypothesis 1 can be true or false, according to the value of v_1 and v_2. Hypothesis 2 is true.

IV. Results

Following the protocol outlined above, we evolved 100 networks for each of the 37 variants of the model. For all networks, we set network size to 2000 nodes; $A = 1$; and $m = 1$. These choices are discussed in the appendix.

i. Goodness-of-fit of the power-law model

For each network evolved we computed two best-fit power-law models, one for $k > 1$ and the other for $k \geq k_{min}$ where k_{min} is the in-degree the minimizes the Kolmogorov-Smirnov distance between the fitted function and the data over $k \geq k_{min}$. On each of these models, we ran a goodness-of-fit test as described in section **??**. This resulted in two distributions of p-values for our control group, plus two more for each of our six treatment groups. Table **??** and **??** report descriptive statistics for these distributions.

From Table **??**, we conclude that onboarding seems to have some effect on the goodness-of-fit of the generated data to their respective best-fit power-law models when $q > 1$. The effect goes in the direction of reducing the p-values and increasing the number of rejects to almost 100%.

It is worth looking at the average p-values generated by each combination of v_1 and v_2. These are shown in Table **??**.

Running t-tests of the null hypothesis that the average p-value in the control group is equal to the average p-values in the different treatment groups results in a strong rejection of the null for any combination of v_1 and v_2 ($6.5 < T < 7.5$ in all cases). It seems unquestionable that introducing

Table 2: *Number of rejects (out of 100 runs) for goodness-of-fit tests of power-law models to in-degree distributions of interaction networks in online communities, with no onboarding (control group) and with onboarding. Power-law models are estimated over all nodes with degree $k > 1$*

	$v_2 = 0.0$	$v_2 = 0.2$	$v_2 = 0.4$	$v_2 = 0.6$	$v_2 = 0.8$	$v_2 = 1$
Control group: 61						
$v_1 = 0.0$	99	97	100	99	99	100
$v_1 = 0.2$	100	100	99	98	97	98
$v_1 = 0.4$	98	98	96	99	100	98
$v_1 = 0.6$	96	96	99	99	99	98
$v_1 = 0.8$	98	97	98	99	98	98
$v_1 = 1$	98	98	100	100	99	98

Table 3: *Average p-values for goodness-of-fit tests of power-law models to in-degree distributions of interaction networks in online communities, with no onboarding (control group) and with onboarding. Power-law models are estimated over all nodes with degree $q > 1$*

Average p-value	$v_2 = 0.0$	$v_2 = 0.2$	$v_2 = 0.4$	$v_2 = 0.6$	$v_2 = 0.8$	$v_2 = 1$
Control group: 0.183						
$v_1 = 0.0$	0.005	0.010	0.005	0.007	0.006	0.007
$v_1 = 0.2$	0.005	0.006	0.009	0.011	0.014	0.012
$v_1 = 0.4$	0.009	0.009	0.015	0.005	0.006	0.008
$v_1 = 0.6$	0.013	0.012	0.008	0.009	0.010	0.008
$v_1 = 0.8$	0.009	0.015	0.012	0.009	0.012	0.009
$v_1 = 1$	0.009	0.011	0.009	0.008	0.010	0.013

onboarding to an online community has a measurable negative impact on the probability of a power-law model to be a good fit for its interaction network's in-degree distribution.

We now turn to the question of the role played by v_1 and v_2 within the treatment group. Figure **??** show the cumulate density functions of the p-values in the control and treatment groups as v_1 and v_2 vary.

Onboarding effectiveness v_1 and community responsiveness v_2 do not seem seem to affect the goodness-of-fit to power law of in-degree distributions much. This is likely to be simply an effect of the large impact of onboarding: the percentage of non-power law distributions is already close to 100% and cannot increase any further.

Regression analysis confirms the intuition from Figure **??**. We generated 6 dummy variables, each taking value 1 when $v_1 = c$ and 0 otherwise, with $c \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1\}$; next we generated 6 more dummy variables for the same vaules of v_2. We then estimated a linear regression model with the p-value of our goodness-of-fit test as the dependent variable and the 12 dummy variables as its predictors. The results are:

1. Coefficients on predictors corresponding to different values of v_1 are positive, but generally non-significant. The coefficient on the variable corresponding to $v_1 = 0.6$ is significant (p-value: 0.009).

2. Coefficients on predictors corresponding to different values of v_2 are also generally positive and non-significant. The coefficient on the variable corresponding to $v_2 = 0.0$ is negative, but also non-significant; that on the variable corresponding to $v_2 = 0.6$ is positive and significant (p-value: 0.018).

3. Coefficients on several interaction terms between v_1 and v_2 are negative and significant.

4. A F-test of joint significance of the group of predictors corresponding to different values of v_1 does not reject the null hypothesis of non-significance (p-value: 0.1541).

5. A F-test of joint significance of the group of predictors corresponding to different values of v_2 strongly rejecta the null hypothesis of non-significance (p-value: 0.0009).

6. A F-test of joint significance of the interaction terms strongly rejects the null hypothesis of non-significance (p-value: 0.0000).

Full regression results are available in the appendix.

When we consider only the upper tail of the the distribution generated by equation **??**, the effect of introducing onboarding on the goodness-of-fit is much less clear. In Table **??** we show what happens when we choose the scaling range so as to minimize the Kolmogorov-Smirnov distance between the degree distributions themselves and their best-fit power-law models. In the control group, the goodness-of-fit-to-power-law test fails in 13 of the 100 runs. In the treatment groups, rejections very from 18 to 36, depending on the values of v_1 and v_2.

Table 4: *Number of rejects (out of 100 runs) for goodness-of-fit tests of power-law models to in-degree distributions of interaction networks in online communities, with no onboarding (control group) and with onboarding. Power-law models are estimated over all observations with $k \geq k_{min}$*

	Control group: 13					
	$v_2 = 0.0$	$v_2 = 0.2$	$v_2 = 0.4$	$v_2 = 0.6$	$v_2 = 0.8$	$v_2 = 1$
$v_1 = 0.0$	34	35	35	25	22	28
$v_1 = 0.2$	35	24	30	34	29	29
$v_1 = 0.4$	28	22	25	34	27	26
$v_1 = 0.6$	29	27	18	23	28	19
$v_1 = 0.8$	26	27	28	36	32	18
$v_1 = 1$	28	28	18	27	21	27

Average p-values of goodness-of-fit tests when $k \geq k_{min}$ are shown in Table **??**. They are all well within the do-not-reject range. The control group has an average p-value which is *lower* than that of the treatment group, which is somehow counterintuitive.

Tables **??** and **??** tell two different stories. Table **??** is unconclusive: in both the control and the treatment groups, we do not reject Hypothesis 2 in the treatment group most of the time, as expected, but must still reject in a relatively large number of cases (13 in the control group, 18-36 in the treatment groups). Table **??** indicates that the average p-value in all groups is comfortably within the do-not-reject range, and in this sense behaves as expected.

V. Appendix

i. A1. Testing for goodness-of-fit of a power law distribution

The goodness-of-fit tests we employed were built following a procedure indicated by Clauset, Shalizi and Newman [**?**, pp. 15-18]. What follows summarizes it in the context of the paper. The test's null hypothesis is that the empirical data are distributed according to a power law model; the alternative hypothesis is that they are not.

First, we fit the data for the degree distribution of a network generated by our model to a discrete power-law model, using maximum likelihood estimation. When we are testing for

Table 5: *Average p-values for goodness-of-fit tests of power-law models to in-degree distributions of interaction networks in online communities, with no onboarding (control group) and with onboarding. Power-law models are estimated over all observations with $k \geq k_{min}$*

Control group: 0.451						
Average p-value	$v_2 = 0.0$	$v_2 = 0.2$	$v_2 = 0.4$	$v_2 = 0.6$	$v_2 = 0.8$	$v_2 = 1$
$v_1 = 0.0$	0.341	0.341	0.345	0.368	0.411	0.355
$v_1 = 0.2$	0.328	0.399	0.364	0.339	0.324	0.381
$v_1 = 0.4$	0.382	0.408	0.367	0.341	0.414	0.372
$v_1 = 0.6$	0.348	0.372	0.4087	0.381	0.413	0.409
$v_1 = 0.8$	0.370	0.383	0.382	0.324	0.359	0.436
$v_1 = 1$	0.383	0.401	0.458	0.413	0.4	0.393

goodness-of-fit of the entire degree distribution, we set the fitted power-law model lower bound to 1; when we are testing for goodness-of-fit of the distribution's upper tail only, we choose a lower bound such that the Kolmogorov-Smirnov distance D between the power law model and the empirical data is minimized. Formally, define

$$D = \max_{k \geq k_{min}} |S(k) - P(k)|$$

Here, $S(k)$ is the cumulative density function of the data for the observations with value at least k_{min}, and $P(k)$ is the cumulative density function for the power-law model that best fits the data in the region $k \geq k_{min}$. The value of k_{min} that minimizes the function D is the estimate for the model's lower bound.

Next, we generate a large number of power-law distributed synthetic datasets with the same scaling parameter, standard deviation and lower bound as those of the distribution that best fits the empirical data. We fit each of these synthetic datasets to its own power-law model and calculate the D statistics of each one relative to its own model. Finally, we count what fraction of the values of D thus computed is larger than the value of D computed for the empirical data. This fraction is interpretable as a p-value: the probability that data generated by our estimated best-fit power-law model will be more distant from the model than our empirical data ("distant" in the Kolmogorov-Smirnov sense). A p-value close to zero indicates that it is quite unlikely that the estimated power-law model would generate empirical data so distant from the fitted power function; a p-value close to one, on the contrary, indicates that the estimated power model is quite likely to generate empirical data that are further away from the fitted power function than the ones we collected.

Generating artificial datasets requires a treatment for the region below k_{min} that differs from that of the one above it. We proceed as follows. Assume that our observed dataset has n observations total and n_{tail} observations such that $k \geq k_{min}$. To generate a synthetic datasets with n observations, we repeat the following procedure n times:

- With probability n/n_{tail} we generate a random number k_i with $k_i \geq k_{min}$, drawn from a power law with the same scaling parameter as our best-fit model.
- Otherwise, with probability $1 - n/n_{tail}$, we select one element uniformly at random from among the elements of the observed dataset in the region $k < k_{min}$.

At the end of the process, we will have a synthetic dataset that follows the estimated power-law model for $k \geq k_{min}$, but has the same non-power law distribution below k_{min}.

This test requires we decide how many synthetic datasets to generate for each test; and what is the threshold value below which we reject the null hypothesis. Again based on [?] we make the following decisions:

- We set the number of artificial datasets generated to 2500. This corresponds to an accuracy of about 0.01, based on an analysis of the expected worst-case performance of the test.
- We conservatively set the rejection threshold at 0.01.

A2. Choosing parameter values

The simulation's computational intensity prevented us from conducting a thorough exploration of its behaviour across the whole parameter space. It follows we had to pick values from some parameters. In this section we discuss briefly our choice of parameter values. The choice of $m = 1$ implies that the number of edges in the networks in our control group will be equal to the number of nodes; we initialize the network with two nodes connected by two edges (one in each direction), then add one node and one edge at each time step. A glance at Fig. **??** shows that this is unrealistic. The real-world online communities described in section 3.1 all display a number of edges with is a multiple of the number of nodes.

We justify this choice as follows: we have no pretence at realism. Rather, we are interested in pitting against each other two phenomena, that of preferential attachment, that tends to generate rich-gets-richer dynamics; and that of onboarding, that tends to introduce a measure of equality. The way we modeled onboarding is by having one single incoming edge targeting the only newcomer to the community at each timestep; we therefore chose to have one single non-onboarding generated edge at each timestep. It seems reasonable that our choice would make these two forces roughly equivalent to each other, and make the impact of onboarding on the in-degree distribution easier to detect.

The choice of $A = 1$ follows from another, and more fundamental, modelling choice. We mimic Dorogovtsev's and Mendes's approach, where the network being modeled is directed and the probability of a new edge to target a node with in-degree k is proportional to k [?]; this contrasts with Barabási's and Albert's approach, that models the network as undirected and assumes that the probability of a new edge to target a node is proportional to its total degree. In a Dorogovtsev-Mendes type model, new nodes have, by construction, in-degree zero, whereas in a Barabási-Albert type model new nodes have total degree one. It follows that, in a Dorogovtsev-Mendes type model, the parameter A tunes the "traction"ÃĎÃź of preferential attachment: the higher its value, the weaker the grip of pure preferential attachment. For $A = 0$ Dorogovtsev-Mendes type models degenerate into "multiple star networks", where the probability of newcomers to receive an edge is zero, and all edges target the nodes initially in the network for all time.

Setting $A = 1$ we make the probability of a newcomer to receive its first edge equal to one half that of an incumbent participant who already has one incoming edge to receive its second one, one third of that of an incumbent participant who already has two incoming edges to receive its third one and so on. One can check that this behaviour mimics that of the simplest, and best known, Barabási-Albert type model.

References

[Figueredo and Wolf, 2009] Figueredo, A. J. and Wolf, P. S. A. (2009). Assortative pairing and life history strategy - a cross-cultural study. *Human Nature*, 20:317–330.
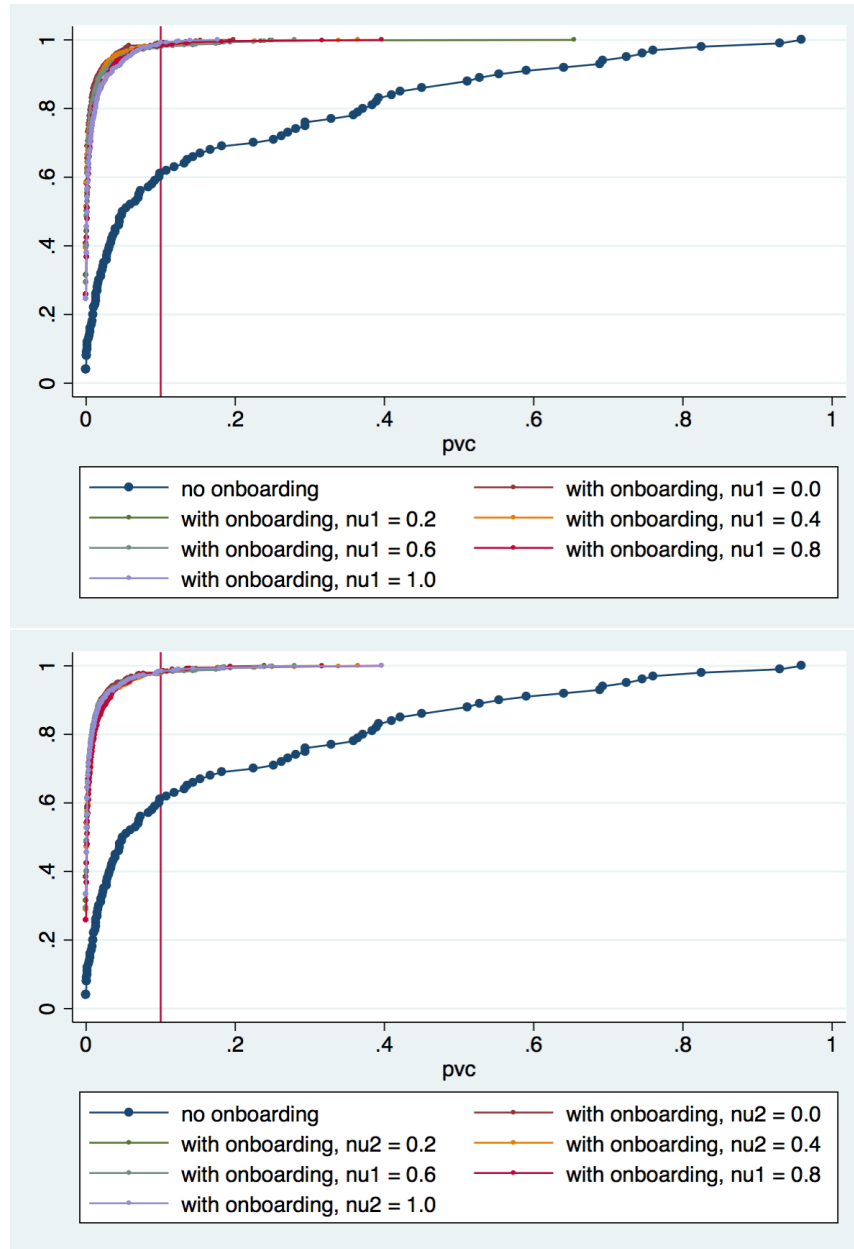
Figure 3: *Cumulate Density Functions of p-values returned by goodness-of-fit tests to the (best-fit) power-law models for in-degree distributions of the interaction networks in the control and treatment groups. 60% of the networks evolved without onboarding (dark blue) have degree distributions that test negatively for H1. When onboarding is introduced, that percentage rises to almost 100%. Above, the treatment group interaction networks have been grouped according to the value taken by v_1; below, they have been grouped according to the value taken by v_2*