SIMULATING THE EMERGENCE OF CRITICAL MASS IN ONLINE COMMUNITIES:

HOW FORUMS GROW BY STIMULATING USER CONTRIBUTIONS

WYBO WIERSMA

OXFORD INTERNET INSTITUTE

mail@wybowiersma.net

Supervisor: SANDRA GONZALEZ-BAILON

Thesis submitted in partial fulfilment of the requirements for the degree of MSc in Social Science of the Internet at the Oxford Internet Institute, University of Oxford.

The total length is 9987 words.

Contents

1 Introduction **1**
1.1 Previous work 2
1.2 Hypotheses 3

2 Theory and methodology **5**
2.1 Critical Mass Theory 5
2.2 Structural Individualism 6
2.3 Agent-Based Simulation 6

3 Sampling strategy and data **8**
3.1 Case study I: Boards.ie 8
3.2 Case study II: Hacker News 9
3.3 Forum growth patterns 10

4 Modelling and results **12**
4.1 The simulation model 12
4.2 Conversational reciprocity 14
4.3 Threads impact on engagement 18

5 Discussion **21**

6 Conclusion **24**

7 Acknowledgements **25**

List of Figures

1 *The four thread-structures.* . . . 4
2 *Colemans boat: Model of the interactions between micro and macro processes. Arrow 4 is not to be used, as it arises from arrows 1 to 3.* . . . 7
3 *Cycle in daily posts on Boards.* . . . 10
4 *UUPDs on Boards.* . . . 10
5 *No growth in UUPDs on HN.* . . . 11
6 *Unique daily visitors for HN.* . . . 11
7 *'Nets & Comms' (pc networks).* . . . 11
8 *'After Hours' (various).* . . . 11
9 *'Messages' (discontinued).* . . . 11
10 *'Hockey' (sport).* . . . 11
11 *Size of threads per hour.* . . . 13
12 *Threads created every hour.* . . . 13
13 *The model: actors, a forum, threads and posts.* . . . 13
14 *Threads generated by the model vary naturally.* . . . 16
15 *Chance of replying to prompting posts, per tie strength.* . . . 17
16 *Extent to which reciprocations exceed chance.* . . . 17
17 *Reciprocation exceeding chance, for each day since first reply.* . . . 17
18 *Reciprocation exceeding chance, in and outside of the thread.* . . . 17
19 *Reciprocation exceeding chance, after days for next user.* . . . 19
20 *Increase in chance of replying at distance (skipped posts) from reply.* . . . 19
21 *Growth for all thread-structures with 50 initial users.* . . . 20
22 *Growth with 100 initial users.* . . . 20
23 *Growth with 200 initial users.* . . . 20
24 *Less growth without reciprocity. Compare to 21.* . . . 22
25 *Larger communities are impacted less by reciprocity.* . . . 22
26 *Flat threads are relatively minority-friendly (top-4 topics shown).* . . . 22
27 *Ordered threads exclude minorities.* . . . 22

List of Tables

1 *Actions and their associated costs and rewards* . . . 15
2 *Actions and their chances (values as X in a thousand calculated at each cycle).* . . . 15
3 *Critical mass attainment for the four thread-structures* . . . 23

Simulating the emergence of critical mass in online communities: How forums grow by stimulating user contributions

Wybo Wiersma
Oxford Internet Institute
mail@wybowiersma.net

Abstract

Attaining critical mass is notoriously difficult for online platforms, and many online communities fail because of this. A common feature of most successful communities, are threads. Threads are found in newsgroups, forums, and on Quora, among other places. In this paper it is examined whether there is something in the structure of threads that makes them especially good at engaging users, improving forums chances of attaining critical mass. And where in previous work on critical mass, social factors have been the primary concern, here non-social factors, such as how threads bundle posts, and mimic conversations, are the focus.

The research was grounded in two data-sets, Boards.ie, a bulletin board site, and Hacker News, a news-site. Data from the latter was especially collected for this study. Agent-Based Simulation was then used as an analytical tool to examine the impact of variations in thread-structures on growth. What was found was that threads perform better than no threads at all, and thread-structures that allow users more control over what to read, successively improve engagement. In addition, based on the empirical data, it was found that most reciprocity on forums — rather than resembling social ties —, was limited to the duration of the conversation in the thread, and thus conversational, rather than social.

It can be concluded from this study, that structural factors impacting engagement are likely to be an important determiner for critical mass attainment. And threads, through improving the reading-experience, and by offering a suitable context for conversational reciprocity, are a powerful structure for enhancing engagement.

1 Introduction

Thread-based discussions are enormously popular on the internet. Not only are they found in usenet-groups, and on classical web-forums, but they also appear in mailing-lists, google-groups, on many news-sites, below blog-posts, and more recently, on Facebook walls and Quora[119]. Over decades several attempts have been made at developing alternative discussion-structures, but apart from Twitter and chat, these have not been successful[15,75,24,123,121]. This suggests there might be something about thread-structures that makes discussion-platforms successful.

Are threads ubiquitous because they were developed early on, and people have become used to this format? Is it because they are simple? Or is it because they are good at overcoming the critical mass problem by better engaging users? Probably a mixture of these, but here the latter is going to be the focus of research. Thus making the research-question:

What is it in the structure(s) of threaded forums, if anything, that makes them good at attracting users, and then retain these users, as it grows? Also, as some thread-based platforms are more successful than others, a comparison is going to be made between the growth-rates that variously structured threads provide.

An Agent-based Simulation Model (model henceforth) is built to clarify and formalize this process. The model is calibrated and tested against data-sets from Boards.ie and Hacker News, two real forums-communities. The latter of these data-sets has been especially collected for this paper. During the modelling, special attention is paid to the utility derived by actors from interacting with threads, and to reciprocity in reply-patterns. Agent-Based Simulation (ABS) as a method was employed because, apart from formalization, it allows for a controlled comparison between the performance of thread-structures, of a kind that otherwise only a large-scale experiment could provide.

The paper proceeds as follows: First its approach will be contrasted to previous work. Next, the hypotheses will be specified. Following that, a theoretical foundation will be laid, and the research-design introduced, and especially its method: ABS. Then the sampling strategy and data will be presented. Next, a simple model will be introduced. It will be extended with short-term reciprocity, based on findings from our data, in which only conversational, and no long-term social reciprocity was found. Then the impact of various thread-structures will be analysed. This is followed by a discussion of their positive impact on growth and critical mass attainment. The paper will be concluded with some ideas for further research.

1.1 Previous work

Previous work has tried to answer similar questions using a comparable methodology. Yuking and Kraut have recently used an ABS to test the influence of different moderation-regimes on user-contributions, and community ties[96,95,97]. A number of other works have employed ABS for studying how innovations reach critical mass across various types of networks, and how this influences the valuation of software companies[58,13,74]. Their analysis was focused on the impact of social network structures.

This paper contributes by focusing on one aspect of virtual communities so far not studied using ABS; thread-based discussions, and the impact of thread-structures on critical mass. Threads have been fruitfully studied using other methods before. In a work studying the shape of threads across topics on Slashdot, it was found, among other things, that political threads tend to be much wider and deeper than those about games[31]. Then there are a set of papers in which the size of Slashdot threads was predicted successfully, based on how fast they received their first few posts[53,55,54]. These findings suggest that threads have their own dynamics.

In this paper, threads were not taken as indicative of social structures, as was done by numerous papers where social networks were generated based on reply-structures[22,59,63,106,119,117,126,106,13,14,12,122]. Instead, it was examined here whether the structural properties of threads, such as them mimicking conversations, can be considered the primary drivers behind both the appearance of such networks, and the appearance of critical mass (more on this in section 1.2).

Then there are two papers that look at critical mass in online communities without employing ABS. A very early paper examines critical mass in (deal-in) BBS systems, which found symmetry in contribution-levels, and content-diversity to be positive predictors for critical mass[93]. The second looked at the success of Wikipedia, its accelerating production function (making later contributions ever more valuable), and heterogeneity of contributors as success-factors. They also note that forum-threads have a decelerating production-function, with early posts receiving much attention, and later ones in the thread being hardly read. Which makes forums good at sparking critical mass, but harder to maintain growth for[89,72].

Finally, a lot of work has been done on motivations for contributing to on-line communities[3,4,11,52,67,77,78,82,91,93,94,97,11,5,51], as well as on crucial factors for the success of online communities, such as their governance, usability, visual appearance, and other factors, including friending, and welcoming newcomers[127,67,90,88,89,60]. While acknowledging their importance, these will not be examined or discussed here, the former category because individual motivations are black-boxed in the model, and the latter because we focus on threads. Also, only the internal dynamics of forums will be looked at. External effects, such as competition between communities, people posting across communities, or even advertising-campaigns, will not be taken into account.

1.2 Hypotheses

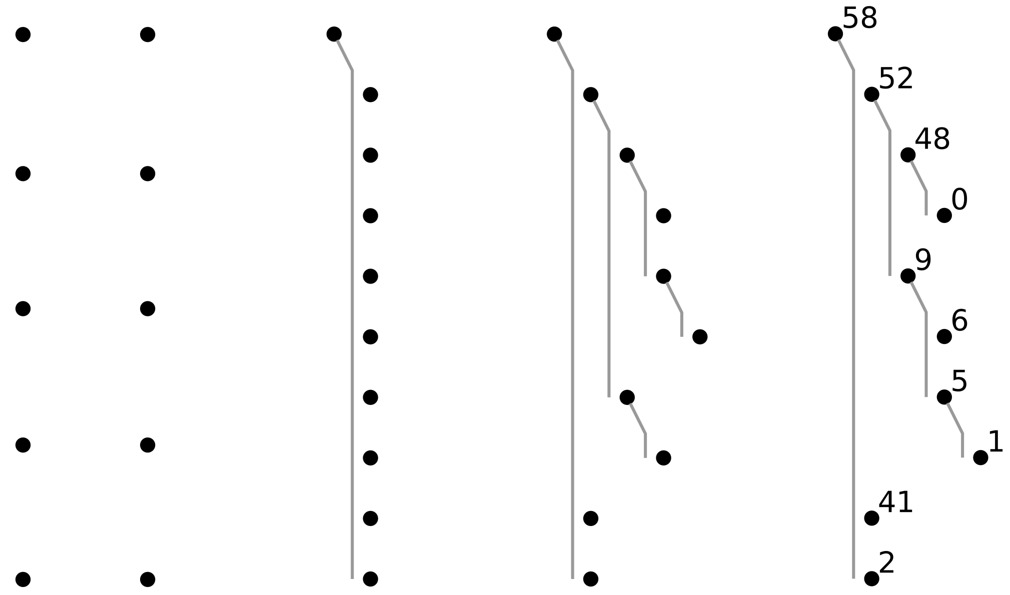
As noted, the central question asked in this paper is whether threads make critical mass easier to attain, and whether there is something in thread-structures that could be (partially) responsible for this.

The expectation is that structural factors, such as the way in which posts are bundled in threads, may help improve engagement. Threads bundle posts in two ways, first of all, by collating posts on a single page. And as moving between pages takes time, and requires a click (+ decision of whether to follow the link), having fewer pages, reduces overhead[66,112]. Secondly, threads generally collect posts on the same or similar topics, making it easier for users to find an interesting read. A third way in which forums might improve engagement, is by focussing user-attention on new threads, raising perceived activity, and ensuring timely replies to posts. All of these should increase the utility users derive from forums, and thus lead to a quicker attainment of critical mass. Hence our primary (H1) hypothesis is that thread structures improve engagement and enable forums to attain critical mass faster.

In order to test this hypothesis, an ABS will be built, and used to test the growth of, and propensity of attaining critical mass for forums using four increasingly refined thread structures. The structures being: T1) Not thread-based: Users are shown random posts. T2) Has threads, but they are simple and flat, as known from traditional bulletin board sites: So an opening-post, and all replies, ordered chronologically. T3) Has sub-threads; threads have multiple levels of indentation, allowing replies to specific posts. T4) Rated threads; the same as in 3, but posts on each level are now ordered by user-ratings (without breaking sub-threads). Each of the four thread-structures are displayed below (figure 1).

A second factor that is expected to be important for engagement, is reciprocity. Reciprocity was found to be important for user retention, by Kraut and Joyce, who respectively saw it as the most important determiner for whether new users would return to a web-community, and it leading to a 12 percentage points increase in them posting again[52,62,9]. Such improvements to user retention can make a big, cumulative difference to the attainment of critical mass. Therefore this factor needs to be modelled correctly in the ABS. In addition, examining reciprocity, also offers an opportunity to test whether the import that the primary hypothesis assigns to structural factors, is really warranted.

As briefly mentioned, previous work interprets reciprocity in forums, mailing-lists, and other online communities socially, rather than structurally. Not only are social processes, such as triadic closure (friends of friends becoming friends), deemed central to the operation of online reciprocity, but social networks are even extracted from reply-structures. In that case each reciprocation (of



(a) *T1: None* (b) *T2: Flat* (c) *T3: Indented* (d) *T4: Ordered*

Figure 1: *The four thread-structures.*

a certain strength, say 4 messages in each direction) is explicitly interpreted as a relational tie, or kind of 'tit for tat', similar to reciprocity in offline communities[34].

The view held in this paper, is, that not all of these interactions might be meaningful socially. That is, most people probably simply reply to interesting posts they come across, largely ignoring whether they know the author of the post they reply to. Patterns in the replying itself could then be explained from the content of the post (whether arousing, incomplete, unclear, etc...) and from people being conditioned to respond to messages, and other conversational habits[104,125,97]. Key characteristics of this *conversational reciprocity*, would be that it is mostly short-term, and confined to the context of the thread, a single conversation only. Thus even for reciprocity, non-social properties of threads and posts are likely to be important.

In other contexts, such as Open Source projects, surveys have shown that the motivations of those contributing include many non-social motivations, such as the joy of coding, or scratching ones own itch[40,124,9,111,83,116]. Parallels to these in forums could be such things as the pleasure of writing, or following the flow of arguments. Bimber called this — contributing to a community without direct intention, mediated by structures created in the environment — second order communality[4]. In any case, social reciprocity can not be the whole story, as it cannot explain why the post initiating the (first half of the) relationship was made. On forums pre-existing ties are, after all, generally absent[9].

The social and conversational explanations of reciprocity will be distinguished by examining what happens in the following example: User A comes across a reply to his post by user B. Now: E1) There is true social reciprocity, and once B replied to As posts, A will be more likely to reply to B specifically, anywhere on the forum, and more or less evenly spread out over time, as would be expected when a relationship had formed between the two users (long-term, social reciprocity). E2) After B replied to As post, A is (much) more likely to post soon after, tapering off with time. This increased likeliness to post, then is mostly confined to the thread, and directed towards B, and possibly others, but mostly in the thread (short-term, conversational reciprocity).

The secondary (H2) hypothesis is that reciprocity is primarily confined to the thread, and thus conversational (E2). This hypothesis will be tested against the forum-data that was collected (see section 3). Findings with regard to it will then inform the way in which reciprocity is added to the model.

2 Theory and methodology

2.1 Critical Mass Theory

The most well-known theory of critical mass in the social sciences, is that by Granovetter[37,38]. It starts off by considering people that have to make a binary choice. For instance, whether to join a riot or not, or whether to spread a rumour. He then proposes that each individual has his own threshold in terms of how many other people (in their surrounding) should choose in a particular way, before they will follow suit. Thus every person that joins a riot will count, and potentially draw others in when his joining makes them meet their threshold. The critical mass is then the point at which the equilibrium — purely based on these individual thresholds — tips, say between very few rioting, and the majority eventually rioting.

In other disciplines critical mass is defined differently[81,18,29,71]. And because in this study the utility derived from direct engagement is considered of primary relevance, rather than merely users expected utility based on how many others chose to participate, several of these definitions are of interest here. The first is the simplest conception of critical mass: as an overall threshold denoting the minimum number of users required for an application to display sufficient network-effects[18]. Network effects are an (originally) economic term, denoting increased utility derived from a product when others are adopting it as well. An example is the Blue Ray player, which is made more valuable as more movies on blue-ray-disk become available. Another of-cited example is the telephone.

Network effects are usually considered a positive thing, as they can help spur adoption. Though, alternatively, with Jacob Goldenberg, one can see network effects as a constraint[29]. In his view the value of the product is fixed, assuming wide adoption, while network effects prevent adoption until a threshold is met. Interestingly enough, this view is comparable to Granovetters', in that Granovetter defines thresholds purely in social terms, implicitly assuming that the action the person is drawn into will be valuable once enough others are participating. Contrary, in this paper it is assumed that while an individuals threshold being met might spur him into joining for social reasons, the (variable) utility he derives from actually participating (network effects) will also importantly affect whether he stays.

A third concept of critical mass, and one that meshes well with considering both initial expectations and utility derived from engagement, is that of a minimum core group of active users needed to sustain the community. This concept adds the possibility of a relapse in activity levels that brings an application under its (overall) threshold of active users again. It is analogous to the concept of critical mass in physics: the smallest mass that will sustain a reaction[19]. In this paper a combination of these three conceptions of critical mass will be employed: a general threshold for the minimum group of active users required, determined both by the extent to which people meet their social threshold (number of other users required), and derive sufficient utility from the activity to sustain them, once they joined.

2.2 Structural Individualism

Another conceptual issue that is important for this research-project, but on an epistemological level, is whether to favour micro, or macro-level explanations. The principles of Structural Individualism are followed on this issue. Which means explaining social processes only based on the (inter)actions of individuals. Quoting Weber: *'In sociological work ...collectives must be treated as solely the resultants and modes of organisation of the particular acts of individual persons, since these alone can be treated as agents in a course of subjectively understandable action'*. Structural Individualism is neither about grand theory, nor about reporting statistical regularities between macro-processes. It calls for examining social mechanisms instead[41,42,44].

Social mechanisms are the 'cogs and wheels' of interactions between individuals that actually bring about the macro-effects. Whereby macro-effects should be understood as aggregates of individual actions, which can only have an impact through individuals' actions (supervenience)[41,45]. A model called Colemans boat can clarify this further (see figure 2): Arrow one is the individual being influenced by macro processes, for example by perceiving global trends, two represents the (inter)actions of individuals on the micro-scale, and three their alterations to the environment. Arrow four then shows the macro-level associations that arise from 1 – 3, which have no causal effect of their own, and thus should not be used in sociological explanations[16].

An advantage of focussing on social mechanisms is that it brings to the fore the actual causal mechanisms, allowing for better predictions and comparisons[46]. Also, Structural Individualisms non-exclusion of contextual factors allows it to encompass limits on action (or nudges) of a political or technical nature, and in the case of this research, also second order communality[86,110]. Leaving room for both social expectations based on perceived activity levels (would be arrow one in Colemans boat) and utility derived by individuals from engagement with threads (arrow two), to contribute to an aggregate growth pattern (arrow three). Similarly, our hypotheses are directly about mechanisms, about threads guiding reading-, and reply-patterns, and both affecting growth.

While Colemans boat is normally used in the context of macro effects on a societal or global scale, the online communities under study are large enough to make the distinction between micro and macro processes relevant. In addition, it should be noted that structural individualism is not equal to a psychological approach, as the specific motivational or intentional states of individuals may even be abstracted away, and be modelled stochastically, as they will be here[44,23]. It are the interactions between people, and the social processes emerging from this, that are its, and this papers' primary interest.

2.3 Agent-Based Simulation

ABS as a method is employed because analysing the (possible) impact of thread-structures on growth patterns would have been impossible with empirical data only. This because with such data, due to external events, and the many differences between forums (including Boards and HN, see section 3.3) excluding alternative explanations would be impossible. A controlled (field) experiment could have covered this, but these are hard to do well over the many months required for a forum to gain momentum. ABS, on the other hand, can function as an abstraction-layer between the data and analysis, so everything can be kept constant, except for the variable of interest. In addition, ABS allows for experimenting without having to rely on human subjects, quickly, affordably, and on a large scale[2].

There are a few well-known applications

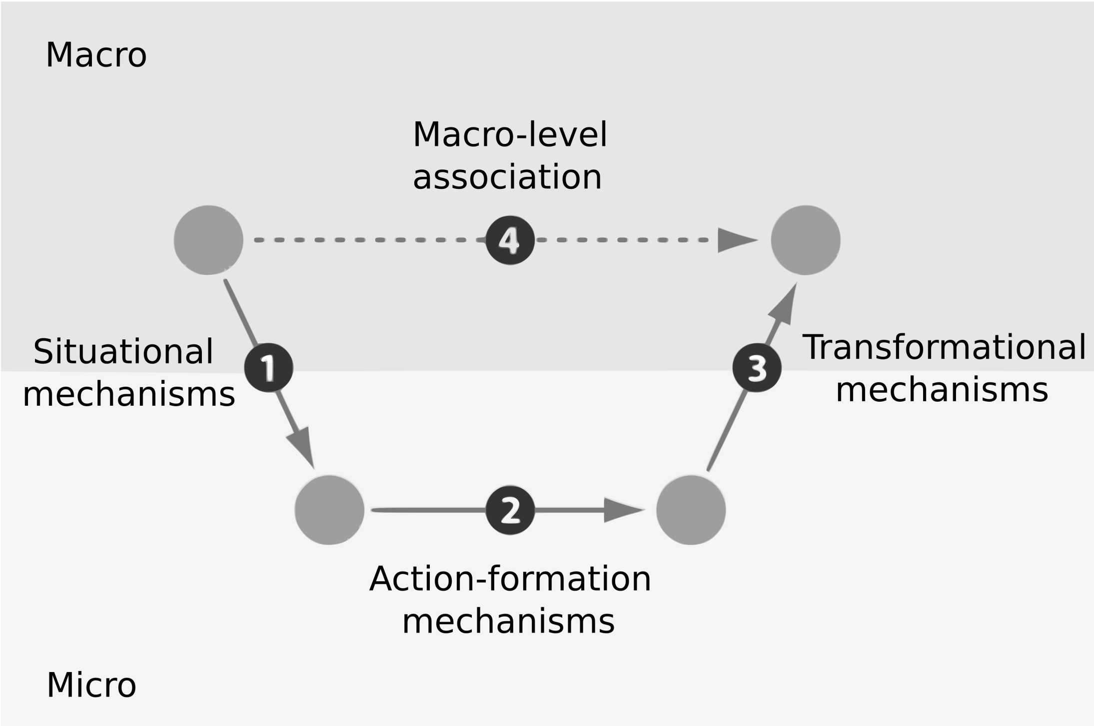


Figure 2: *Colemans boat: Model of the interactions between micro and macro processes. Arrow 4 is not to be used, as it arises from arrows 1 to 3.*

of ABS in sociology. The first is the modelling of segregation, done by Schelling. He found that small changes in individual preferences, such as a rise in preference for similar neighbours from 25 to 26%, led to a large rise in segregation at the macro-level. The second, is the modelling of the movement of pedestrians, which has already saved many lives by improving the design of safety fencing at festivals, in the London underground, and even around the Masjid al-Haram mosque in Mecca. The final, is Nigel Gilberts pioneering work on simulating primitive societies, and innovations-networks[27,100,101,47].

In ABS, at minimum, agents behave according to decision-rules, possibly enhanced by memory, computations, and observations from the environment. The latter of which, and its most important objects/structures, would then be modelled as well (threads and posts are in this study). The state of the agents, their position in the environment, as well as the actions of others, then determine the agents opportunities and actions. Sociological processes will then emerge from these processes alone[8,69]. This focus on individuals makes ABS a good fit with Structural Individualism, as social mechanisms can be directly modelled.

It is desirable for a simulation model to be grounded in theory. This is because a model that matches the data is not necessarily correct. Multiple models that generate the same growth-patterns/output can always be created, without any way to decide between them. While this problem of so called generative plurality, is not a new one (any empirical study trying to establish causality suffers from it), the fact that in ABS one works with multiple independent variables, some of which are hypothetical, instead of measured, makes it more of a problem. There is an alternative to complete theoretical grounding, however, acknowledging it as a limit of the methodology[20].

A simple, somewhat plausible model that produces the complex macro-behaviour still allows one to claim that, on an analytical level, the model offers a possible explanation: that is, it is enough to generate the complex behaviour; it has generative sufficiency[69,45]. And given the scope of the current work, and it being the first in which user-interactions at the level of threads are modelled, full grounding is not aimed for. Examining whether thread-structures can impact growth is this papers primary purpose, not being absolutely correct on the exact size or shape of the impact. Moreover, a separate advantage of targeting generative sufficiency only, is that this allows for a more parsimonious model, furthering its intelligibility.

ABS is different from two other kinds of modelling. The first is equation-based modelling. Equation-based modelling, rather than modelling the local interactions of individuals, approaches sociological processes as unified systems, and specifies relationships between macro-outcomes. Secondly, it is also different from game-theory, in that it relaxes the rationality requirements on the part of agents and allows for more extensive modelling of the environment. The latter of which makes it very suitable for our purpose of analysing the impact of thread structures. Both these differences, however, come at the cost of not having easily calculable outcomes or optima: One has to run the model to see what happens[69,8,30].

Another advantage of ABS is that it allows one to integrate multiple, specific, smaller theories, combining them into a middle-level framework[26]. And all this, while still being able to open the black-box of micro processes. A final advantage is that the model/framework is not just written down in a narrative (as it otherwise often would have been), but is rather formalized in the code of the model. This both enables it to be more strictly defined, and allows for the model to be informed by, and even be tested against hard data[7]. These properties are what allows a model to mitigate between the data and the analysis.

3 Sampling strategy and data

3.1 Case study I: Boards.ie

As not much research has been done on critical mass in web-forums yet, and data-sets are relatively labour intensive to create/pre-process, our sampling strategy was not based on a random selection of web-forums/news-sites. Instead, two typical, yet different sites were picked as our case studies. The first data-set is 30 gigabytes of data from Boards.ie (Boards), the largest Irish bulletin board site[6]. This site offers hundreds of sub-forums on various topics, from films and real estate, to hunting. It was founded in 1998, and the data-set runs until 2008, making it relatively longitudinal.

The Boards data was collected and provided by John Breslin, from NUI Galway[103]. Even though pre-processed, and stored in XML, it had to be processed further to extract thread-structures, posts, posting-times and other meta-data. The XML pages were parsed using the Nokogiri HTML scraping library[76]. Then some hoops had to be jumped through to re-connect posts to threads, as they were stored separately. Next, tools were written for calculating various statistics. Also, the sheer amount of data (30 gigabytes) required tools optimized for speed and memory usage.

As for the statistics, the Boards data-set consists of 7,755,000 posts, made in 623,000 threads, by 72,000 users. In terms of user-activity, an average of 18.5 posts and 1.2 threads were created by 12,500 users in the datasets last month. The distribution of these is left-skewed (the medians are 4 and 1). Secondly, circadian cycles can be observed for post-creation, with each cycles peak occurring during waking/work hours of the sites dominant demographics (Irish users). Secondly, weekly cycles are visible as well, with dips during the weekend (figure 3). These cycles warrant further study, as by concentrating user-activity around certain times, they could potentially bump perceived activity levels over a critical mass threshold, but in the model built here, they are not accounted for.

As for growth, a measure for active user-participation was devised. In the literature the number of daily posts is often reported, with another alternative being counting users between their first and last post on the forum. The number of unique users posting per day (UUPD), was chosen here. Its main advantage is that it can also be calculated on the fly in the ABS (for which, while it runs, it is unknown when an users last post will be). In figure 4, the growth of the number of UUPDs is shown for the whole Boards community. As can be seen, for the last decade there was a slow exponential growth on Boards. Interesting to note are the flat line near the beginning of the graph, and several dips to near zero later in Boards' lifetime. These were caused by interruptions in server uptime.

3.2 *Case study II: Hacker News*

The second data-set is the Hacker News (HN) startup/technology news-site, which is comparable to Slashdot, in that people can post links to news stories, and comment on them[105,39]. It has recently surpassed Slashdot in visitor-numbers, and is quite large, with 100,000 unique visitors per day, most of whom are lurkers[36]. It is ran by Paul Graham, an internet millionaire and a prominent venture capitalist. This data was collected especially for this paper. It runs from February 2011 until the end of May.

More scripts had to be written to gather and process this data, which was done in stages. An initial script fetched the threads and user-pages from the web, and a second parsed the HTML-pages on the local machine using Nokogiri. All custom scripts together (including the ABS) totalled 7,500 lines, which, according to the Ohloh code-analysis site, is over 14 man-months worth of programming[80]. All code is open source, and available on Github[28]. Apart from enabling a thorough analysis, some scripts were also needed to overcome a special challenge with the HN data.

This challenge is that HN does not provide time-stamps with posts and threads, but rather phrases such as 'X hours ago', which complicates getting a handle on the time dimension. This is especially true for material older than a day, as hours are no longer reported after a day. Luckily enough, all items (posts and threads) on the site are assigned an unique ID, which increases predictably with time. This allows us to deduce the time at which items were posted from the order of their IDs. By knowing of two items that one had been posted six hours ago, while the other was posted four hours ago, it can be deduced that a third item, with an ID in between these, must have been posted before the latter, and after the former, so approximately five hours ago.

Having this knowledge makes it possible to reconstruct the time at which all comments were posted, without having to sample every thread every few hours in the hope of recording the time of any new comments. By another strike of luck, the site has a 'comments' page, which listed all new comments. This page was harvested every 10 minutes (from 2 separate machines, in case one would go down), producing a sample of IDs and their timing dense enough not to leave gaps larger than two minutes. Threads thus had to be harvested only once.

In all, the HN data-set contains 299,000 posts, made by 23,000 users in 10,700 threads. In the data-sets last month, on average 5.9 posts and 0.2 threads were created by 11,000 users. These figures were, as expected, left-skewed again (medians 2 and 0), and cycles were found again as well (figure 12). However, contrary to Boards, the HN community did not grow over the four months

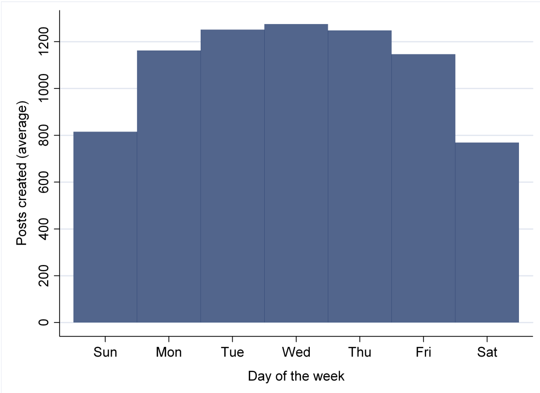


Figure 3: *Cycle in daily posts on Boards.*

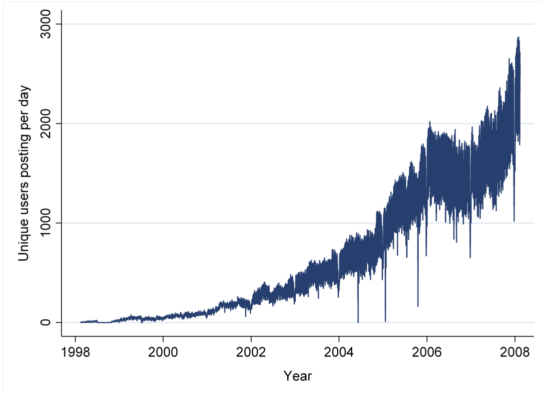


Figure 4: *UUPDs on Boards.*

for which there is data, as can be seen in figure 5. However, from the graph of unique daily visitors provided by Graham (figure 6), it can be seen that there was growth from 2007 onwards, when the site opened. Also interesting to note, is that there were far more readers on HN than posters (over 98.5%). A site having this many lurkers is exceptional, compared to figures reported for newsgroups (normally up to 90%)[77]. This suggests that HN, in line with it being a news-site, is predominantly a community of readers.

3.3 Forum growth patterns

The reason for drawing on both the Boards and HN data-sets, in addition to increased robustness, is that Boards uses flat threads, while HN features indented threads that are ordered by ratings. These correspond to our thread-structures T2 and T4 (section 1.2), and thus provide good coverage for calibrating the model. On the other hand, a difference between the two data-sets that makes them harder to interpret, is that while HN is a single forum, Boards consists of 505 sub-forums. Graphs for 4 Boards sub-forums are shown in figures 7 to 10. As can be seen, the forums that are created later on, show high initial user counts. Which can be explained by them benefiting from users already frequenting other sub-forums.

Also, Boards largest sub-forum (figure 8), has on average ‘only’ about 200 UUPDs, even in its last (and best) year (compared to HN’s 1,300+). This makes it seem as if (flat-threaded) forums have a size-limit, which Boards circumvented by having many forums in parallel, compounding their growth. This makes the Boards data not very representative for the growth of independent forums. In addition, as noted, no posting-data was available for HN’s initial years. These things together make that growth-figures were not used to calibrate the model directly. Instead, a general exponential growth-curve was aimed for.

A number that is directly used in the model, is the fraction of UUPDs that arrived newly on the forum every day. In percentages, for HN this figure was found to be 9.5% on average, while for Boards it is 4.5%. As this number appeared slightly higher for younger sub-forums, 9% is used in the model. It should also be noted that the percentage of new users varies considerably from day to day (a standard deviation of 4.9% for HN), due to external events such as changes in Google rankings and the forums being featured on other sites. These effects are not included in the model.

Another observed feature that informed the design of the model, is the ratio of threads relative to the number of posts that are created. On Boards the average number of posts per thread was 12.4 (median 6). While on HN each thread received many more posts, at 28.0

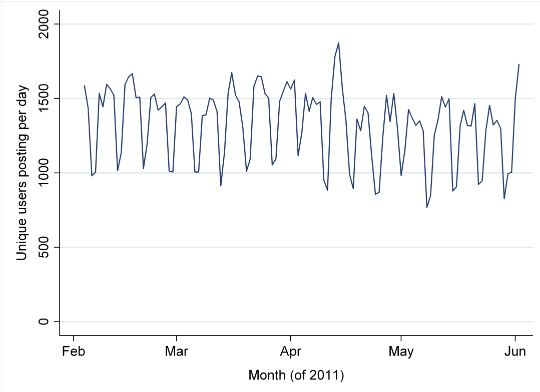


Figure 5: *No growth in UUPDs on HN.*

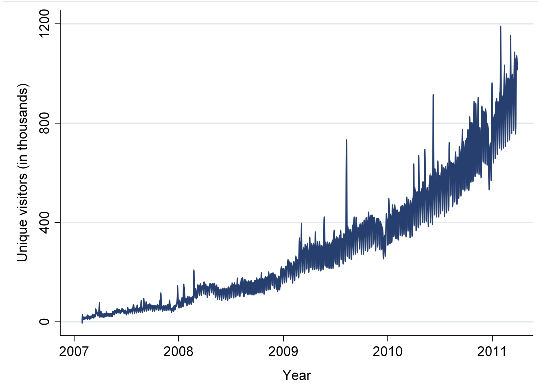


Figure 6: *Unique daily visitors for HN.*

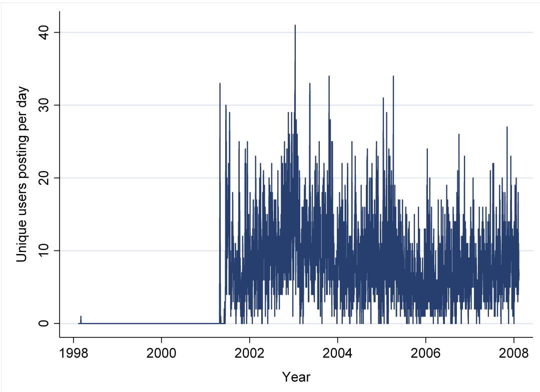


Figure 7: *'Nets & Comms' (pc networks).*

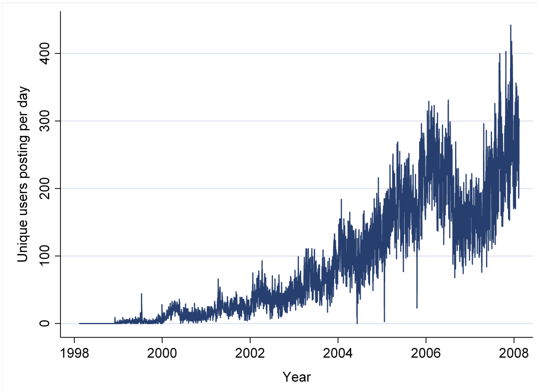


Figure 8: *'After Hours' (various).*

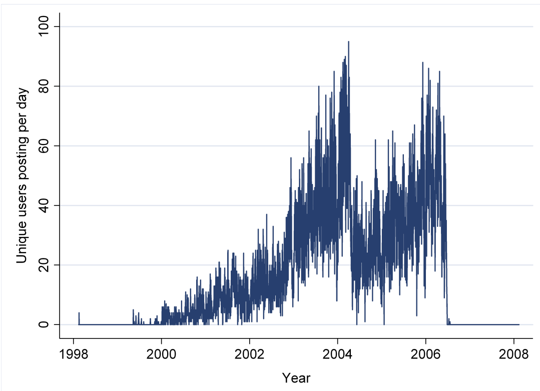


Figure 9: *'Messages' (discontinued).*

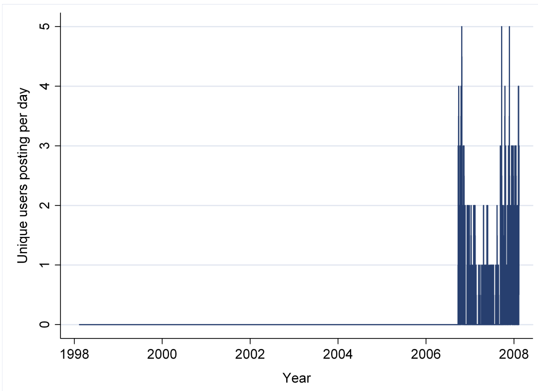


Figure 10: *'Hockey' (sport).*

(median 15). These ratios were found to be remarkably stable over the life of sub-forums, and even across circadian cycles on HN. As can be seen in figures 11 and 12, thread-churn simply increases in busy periods, offsetting extra posts. The ratio is thus kept constant in the model, and for consistency reasons, is set to the same value across all thread-structures: Posting posts was made 20 times as likely as creating threads.

4 Modelling and results

4.1 The simulation model

The model is built to test the primary hypothesis (H1), that thread-structures benefit forum growth. The model consists of four entities: actors, a forum, threads, and posts (see figure 13). Actors represent the actual users that visit, and participate in the forum. In each simulation an initial number of actors enters the simulation. This number is varied between experiments, between zero and 200. In addition, every day a certain number of new actors show up. This is the fraction of current users arriving. In experiments testing not just growth, but also social thresholds (see section 2.1), a fixed number of actors will be created instead, but more on the experiments in the next section. 'Days' in the model consist of 240 cycles, with every cycle representing 6 minutes, and allowing every actor one action.

Then there is the forum object, of which there is only one in each simulation. It keeps track of threads and actors, and keeps various counts. The most important parameter for the forum is the maximum number of threads that are visible to actors. This should be pictured as, for example, the list of recent news-items on HN. Two important assumptions are made here. The first is that actors, once on the forum, will not browse beyond the front-page. In the case of Google results pages, this was found to be true for at least 80 percent of searches[49,50]. It is assumed to be always true here. The second assumption is that actors will actually arrive through the front-page, and not, for example, through a direct link to a thread, or a search. This assumption, while reasonable for news-related forums such as HN, could be less so for Boards. At least it was found for HN and Boards, that less than 11.0%, and 25.4% respectively, of replies came in after a thread had left the homepage.

Actors move along threads while reading the forum. Threads define the structure in which posts are displayed. As noted, four thread-structures are tested: without threads, flat threads, indented threads, and rating-ordered threads: making threads' structure-type the most important variable of the model. In addition, threads, as well as posts, have topics. There are 6 topics in the model. Actors have a preference for certain topics. These preferences are distributed according to Zipf's law: $1/n$, with n being the topic's rank. This as, besides the frequency of words in the English language, Zipf's law was also found to approximate the frequency of tags in online communities such as Delicious[10]. Topic 1 is thus preferred by 41% of the actors, topic 2 by 20.4%, roughly halving for each from there. Topic preferences determine both how likely an actor is to post on the given topic, and the utility he derives from reading a post on that topic. For simplicities sake, post-quality is not modelled separately, and is thus part of the topic-preference.

The utility derived from different actions is, mostly following Hedstrom, modelled in terms of a desire. This desire is split in two along a temporal dimension. First of all, there is the desire to participate during the current session, and secondly there is a mounting desire to engage with the forum at a later time. This split was made to provide for a natural moment to go offline: once the actors current desire is depleted. When offline, the actor has a chance of coming online equal to his next desire over 1000. Upon going online, his next desire becomes his current. Costs are always deducted from the current desire, while util-

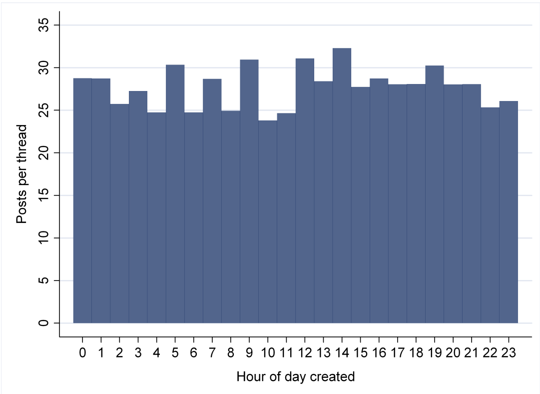


Figure 11: *Size of threads per hour.*

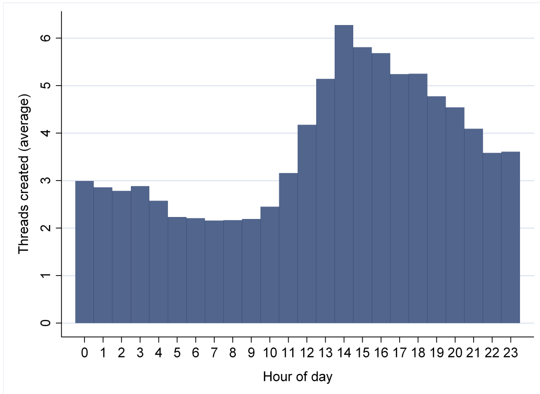


Figure 12: *Threads created every hour.*

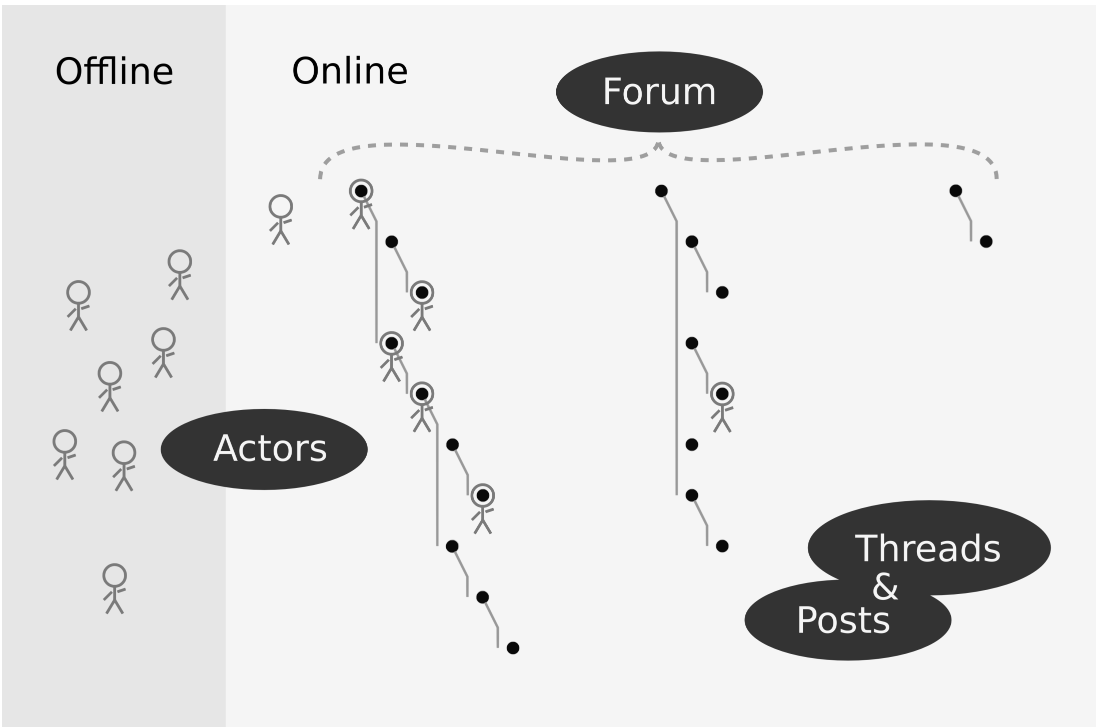


Figure 13: *The model: actors, a forum, threads and posts.*

ity rewards are mostly credited to the next-desire. Costs can both be pictured in terms of satisfying desires (such as being updated on ones topic) and of annoyances or time spent. The actions that have associated costs and rewards, are listed in table 1, along with their respective costs/rewards in utility points.

To give a better picture of the simulation, the following is a simple walk-through: An actor is offline in the simulation. His next desire is 12, and he comes online as he hits his 12 in 1000 chance of doing so. His next desire now becomes his current desire (his new next desire is 0). He then reads a post, loses 1 current to this, and another 0.2 for loading the page, but as the post is on-topic, gains 1.8 for his next-desire. Over the next 9 cycles he reads 9 more posts in the thread. As all of them are on-topic apart from two, his total next-desire balance becomes $17.2i$ $(8 * 1.8 + 2 * 0.5)$. Finally, he wites one post in the 11th cycle, and then logs out, as his current desire has become -1.7 $(10.8 - 1 * 10 = 0.8, 0.8 - 2.5 = -1.7)$. He won't come online again until he hits his new 17.2 in 1000 chance.

Decisions on which action to perform, such as whether to read or to post, are, as the example shows, fully stochastic. Table 2 lists the respective chances for all the actions that the actor can perform. The chances are mostly derived from our data. Notably, things such as expected utility, are not taken into consideration by actors, nor do actors learn from their behaviour, evolve, or change their behaviour over time for other reasons than changes in their desire-balances. Such factors were left out of the model because, as noted, parsimony and communicability, rather than complete realism, were priorities in the modelling strategy. Secondly, as the main topic of this study is the impact of structural effects on aggregate behaviour, users psychological considerations could be black-boxed without sacrificing much of the models explanatory power.

Pseudo-code for the model can be viewed in appendix A. Screenshots, and links to an online demo and the full source-code, and all output, are provided there as well. Before any results were generated, the model was verified internally and externally. Internal validation consists of checking the models implementation. While nothing can ensure that software is bug-free, the impact of all parameters was tested by giving each a very extreme value in turn, and checking that the model behaved as expected[7,48]. Two bugs were discovered (and fixed) this way. Sensitivity testing was done as well, which means exploring each variables impact, as well as the ranges at which the model still displays normal behaviour. What was found, is, that the model is most sensitive to the `on_topic` and `off_topic` settings, which is not surprising, as these are intuitively linked to users utility earnings.

Its external validity then, consists of hypotheses being formulated beforehand, and the model being appropriate for testing these[7]. Finally, the model also seems to be plausible, both in terms of the variables it considers, and its output (see figure 14), though we leave judgement of this to the reader.

4.2 Conversational reciprocity

Our secondary hypothesis (H2) is that reciprocity on forums is not tie-like, but confined to the thread. In addition, whether people receive a reply, is an important determiner for whether they come back to a forum, making it an important factor for the model as well (which tests H1). In line with Kraut's results (section 1.2), in our data receiving a reply made new users 8.4 respectively 16.1 percent points (Boards, HN) more likely to post again. And as in our data-sets about 24% of users never come back after their first post (in some newsgroups it was even 60%), capturing even a fraction of these users can make a big difference to growth[94]. In addition, it was found that receiving replies causes users to post more frequently[65,64]. On HN, the median interval in hours between posts was re-

Table 1: *Actions and their associated costs and rewards*

Name	Reward/cost	Description
	Current desire being spent	
read	-1	Satisfied reading desire
create	-3	Time/cost of writing
page_load	-0.2	Page loading time cost
	Building the next desire	
on_topic	1.8	Reading ontopic post (net 0.8)
off_topic	0.5	Reading uninteresting post (net -0.5)

Table 2: *Actions and their chances (values as X in a thousand calculated at each cycle).*

Action	Chance	Notes
create reply	25	
create new thread	1.25	1/20th as likely as replies
move to next thread	U * 4	Uninteresting posts seen times 4
move to next post	*remainder*	

duced from 36 to 21 by a reply. On Boards the reduction was less (47 to 45), but there, more users received replies anyway (59.2 versus 46.9% on HN).

It will now be examined how the user A from the example (see section 1.2) reacts to receiving a reply; with either social reciprocity (E1) or conversational reciprocity (E2). Making the distinction is not easy, as there are many things going on at the same time in fora. For example, merely using increased posting frequency as a proxy for the motivational impact of replies, has limited value. This because any burst in activity (possibly for other reasons, such as more time off, online for the first time, etc.), would also lead to more posts (and likely some replies). Also, people often don't wait until they receive a reply, but keep posting, so the effect of replies to each individual post cannot be untangled. Therefore, both aggregation over many users, and triangulation, is necessary to get a better picture of reciprocity.

The first way to examine whether reciprocity is relational, or conversational, is by tie-strength: Are people more likely to reply to someone if they received multiple replies from them? Relational reciprocity would suggest this, analogous to how people are more likely to spend time with close friends, than with acquaintances. However, because some users post more often than others, simply counting the number of replies that strong ties receive, compared to weak ties, will not work, as people that leave a lot of posts on the forum are expected to receive more replies, purely by chance.

Therefore, the chance of replying to each user was calculated, based on how many posts they had put out as possible prompts on the forum. Next, for each user, and for each tie strength (to all of the other users), it was calculated whether the number of replies they received, was more or less, than chance. And in order to make this calculation more precise, all of it was done upto the time of each individual reply, because posts created after the reply in question could not have been prompts for it. This calculation resulted in figures 15 and 16, which, as a side-note, only contain results for HN, as the data from Boards shows almost the same pattern.

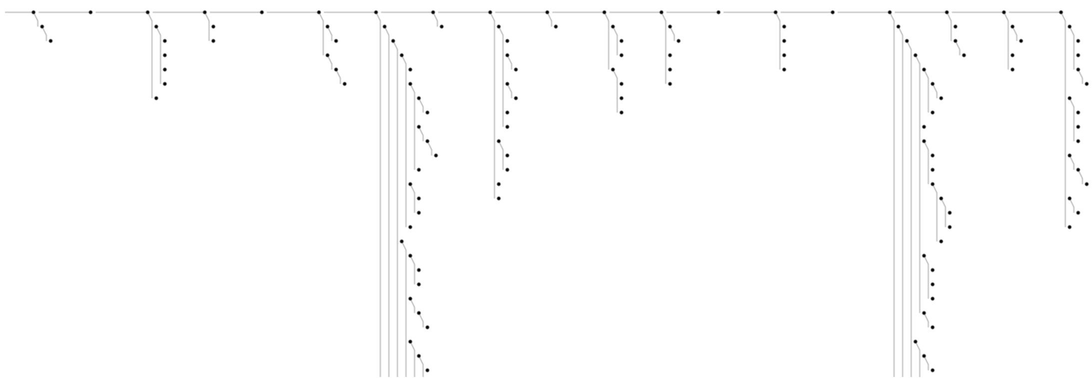


Figure 14: *Threads generated by the model vary naturally.*

In the graph on the left, it can be seen that users which have (strong) ties with others, post a lot more often, and through this alone, increase the chance of receiving replies to their prompts, by upto 6.7 times. The other graph shows to what extent reciprocation goes beyond this chance. As can be seen, the average user replies to those he never received a reply from slightly less often than expected by chance (multiplier of 0.83, 97.6% of replies). While he is is 8.23 times more likely to reply to those that did reply to him once (1.9% of replies). However, though replies remain higher than chance, the multiplier goes down for users that have replied to him more often (that he would have stronger ties with), rather than up. And, though the multiplier seems to go up again for ties stronger than 8 (upto 6.8), data is sparse for this, as only 0.04% of replies fall in those categories. Putting first doubt on the social reciprocity hypothesis.

The second dimension along which to distinguish whether reciprocity is relational or not, is time. Relational reciprocity suggests stable, or intensifying reciprocity. While conversational reciprocity would show a spike in replying, tapering off in a matter of days. Therefore, the above calculation was repeated, but now with the time in days since the first reply along the X-axis, rather than the tie-strength. Where chance was calculated against the prompts created on that day. The picture that emerges (figure 17) is exactly what the conversational hypothesis would suggest, in that replies to repliers exceed chance on the first day (64.8 times!), but taper down quickly, to 32.7, 2.4, and then further down on the second, third, and later days. If users that replied once or twice are excluded, results hardly differ, so reciprocity seems non-relational for stronger 'ties' as well.

The third, and most important dimension that can shed light on the issue, is whether reciprocating replies are being made mostly inside threads, or not. Figure 18 shows the same calculation, but now split out between replies to the replier inside threads that the user received the reply (or any other replies) in, and outside of such threads. As can be seen, replies inside threads go beyond chance much more at 90.2, while outside they exceed it 'only' 45.9 times, about half as much. Interestingly, however, inter-thread replies taper down within a day, while those inside are still at 85.4 on the second day, and go down after that. This seems to suggest a tendency towards conversational reciprocity, though some measure of reciprocation happens outside of threads as well.

Finally, instead of calculating the chances for replies to the replier, they were also calculated for replies to a third user, C, the first to post after the replier, conditional on him

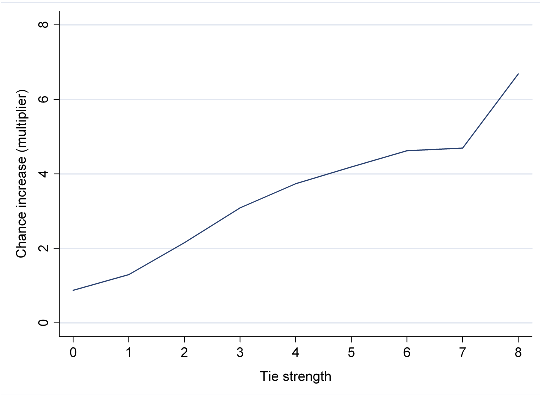


Figure 15: *Chance of replying to prompting posts, per tie strength.*

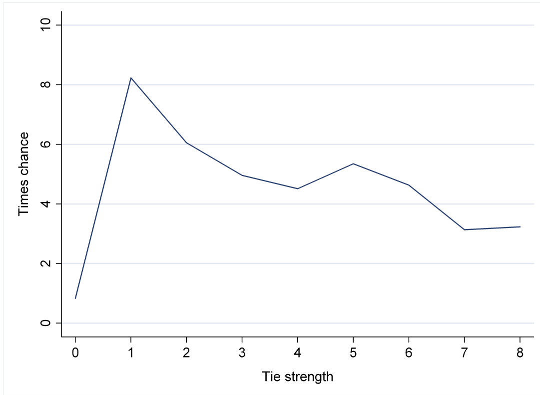


Figure 16: *Extent to which reciprocations exceed chance.*

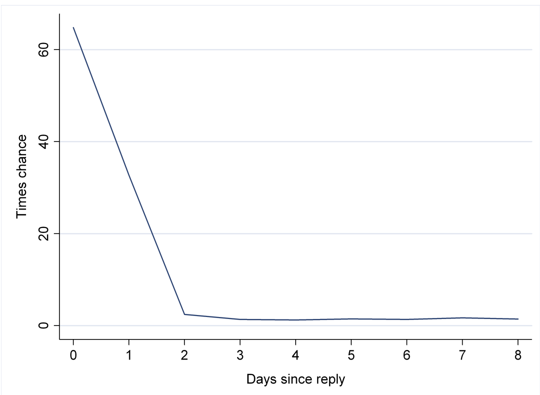


Figure 17: *Reciprocation exceeding chance, for each day since first reply.*

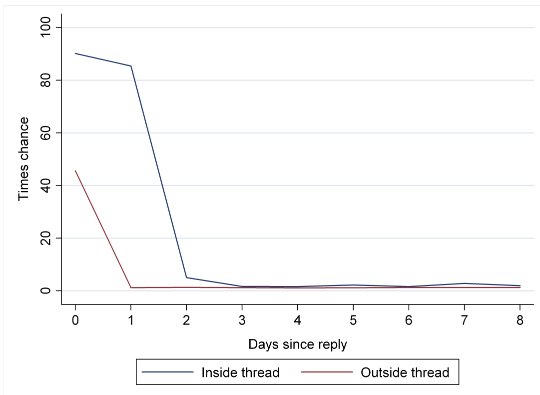


Figure 18: *Reciprocation exceeding chance, in and outside of the thread.*

not having replied to A as well. As can be seen in figure 19, apart from chance being exceeded to a lesser extent, the picture is pretty similar, with most posts happening inside the thread, and them tapering off faster outside the thread. And while being lower, the multipliers are still pretty high, at 31.6 for in-thread replies, and 12.4 for out-of-thread ones. This indicates that there is a significant general spike in replying, to all posts in the thread, no matter by whom they were written. As can now be concluded, though replies to repliers are about three times more likely (multiplier of 90.2 versus 31.6) than replies to others, this is almost exclusively true at the short term level, and mainly within threads. Which suggests an explanation for reciprocation that is primarily conversational (E2), thus confirming H2.

A concrete, mechanism-based explanation for conversational reciprocity could be that threads, as the context in which the user discovers that his post has received a reply, also present him with prompts to reply to, notably first in sight, and most relevantly, that of his replier. Such immediate reciprocation is well known from so called flame-wars, where people respond to one another in quick succession. But this data suggest that short-term reciprocity may be the common form in which it appears.

This mechanism was added to the model. Where short-term was interpreted as the actor remembering the last ten repliers to his posts that he came across. And because replies to repliers were found to be three times as likely as replies to other users, they were made to be. Next, to model the general burst in replying, it was established how much more likely an user is to reply after receiving a reply, per model-cycle (rather than per day). To do this, for each distance from the prompt, the relative rate of posting when a reply was received, was compared to that for when none was received, producing figure 20. As posts are 3.7 times more likely directly after reading a reply, users receive an additional 60 points reply-desire when they see they have been replied to. This desire drops off by 6 on each cycle to make it match the graph. If the user posts, it is reset to 0 immediately. In addition, users receive a 20 points bonus to their current desire, when they see a reply. While arbitrary, this represents the motivational increase provided by a reply.

4.3 Threads impact on engagement

Now that reciprocity is included in the model, the impact of thread-structures on critical mass attainment can be tested (H1). This is done in two stages. General growth patterns will be examined first for each thread-structure, and in the second stage social thresholds will be added. For testing growth, a fraction of current actors enters the model every day (9% of UUPDs), each of which comes online at some point, and engages with the forum, gaining and/or losing utility (see section 4.1). This should give an impression of how good each thread-type is at engaging users through the two mechanisms of bundling, and facilitating conversational reciprocity. All results reported in this section are averages over 100 runs, with 95% confidence intervals.

The first of the thread-structures, the no threads case, serves as a baseline. As its name implies, it does, not have any threads. Rather, every post is its own thread. Individual blog-posts (without commenting), or hyperlinked homepages, are an example of this structure. In the model, actors move between posts randomly, without knowing in advance whether the page they move to is going to be interesting to them. In addition, actors lose a certain amount of utility per page they visit, because of loading-times, and lack of a conversation or narrative that draws the actor to the next page. Also, as hyperlinks are normally unidirectional (no trackbacks, which is a recent feature of blogs to make them bidirectional), actors will not be able to learn about replies to their posts. So no reciprocity bonus is dealt

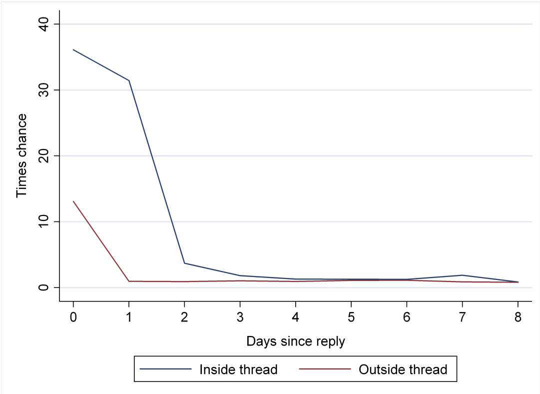


Figure 19: *Reciprocation exceeding chance, after days for next user.*

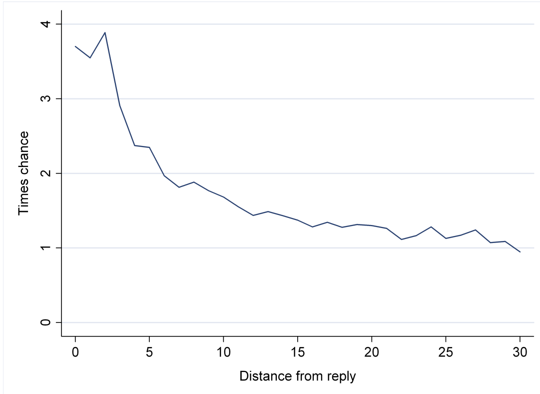


Figure 20: *Increase in chance of replying at distance (skipped posts) from reply.*

out. Together, these factors make that, without threads, there is no growth, and the actors that enter the platform run out of utility quickly. In figures 21 to 23 the growth-patterns of all thread-structures are shown, with 50, 100, and 200 initial users added to the model. While interpreting them, it should be noted that not all initial users post on the first day, so the number of UUPDs is lower.

The flat structure is the first real thread-structure that was modelled. It is also the simplest one. Actors navigate from post to post along the thread in linear fashion. In addition, all posts in each thread are collated on a single page, so no page-loading penalty applies between posts. Also, threads, as mentioned, have topics, defined by their opening-post. Actors will occasionally create off-topic posts (one third of posts will be on-topic, in a third of the cases the topic is determined by the post they respond to, leaving the rest for the usual topics distribution). In addition, actors have a 1/3rd chance of entering uninteresting threads. So threads only provide a rough guidance on topics. As briefly noted, actors also have a chance of leaving the thread prematurely, which increases with every uninteresting post they see. As can be seen in the graphs, thanks to these modest structural improvements, flat threads do engage users when the number of initial users is sufficient.

Indented threads, then, have (recursive) subthreads, formed by the replies to each post (see figures 1 and 14). The main advantage of sub-threads is that they not only allow readers to skip uninteresting threads, but also replies to uninteresting posts. Their downside is that, where in flat threads almost all posts receive a reply (as in those replies are appended at the bottom), in indented threads fewer do (also true in our data, see section 4.2), making those actors miss out on their reply-bonus. This makes indented threads especially suitable for communities that have proportionally more readers. The graphs show that indented threads speed up growth as expected. In addition, indented threads might be better for larger communities as well, because in those, thread-churn is greater (threads are bumped off the frontpage sooner), so some users might not find out about having received a reply, and thus miss their reply bonus anyway.

As for the impact of reply-bonuses, if they are disabled, only communities with the largest number of initial users (200) survive, suggesting the important role reply-bonuses play while communities are small. In addition, when conversational reciprocity is disabled (maintaining reply-bonuses and the same likeliness of people posting replies, merely to anyone, rather than repliers), small communities are also more disadvantaged

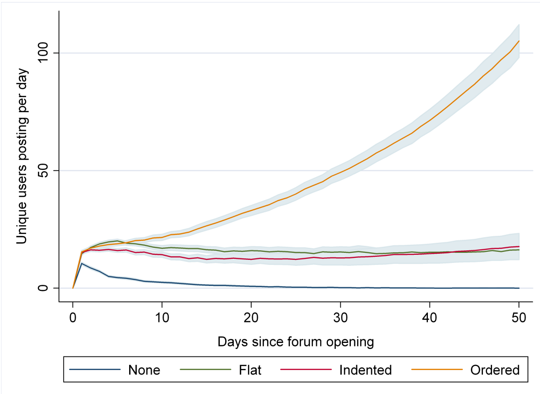


Figure 21: *Growth for all thread-structures with 50 initial users.*

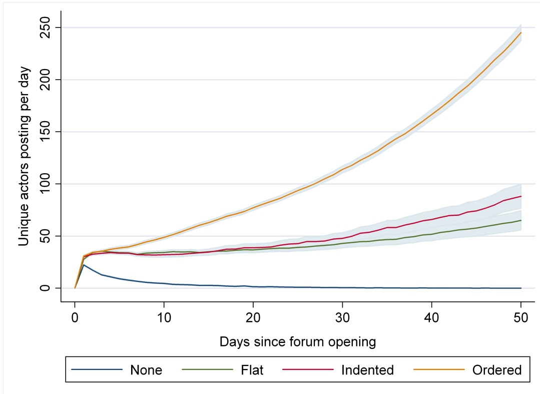


Figure 22: *Growth with 100 initial users.*

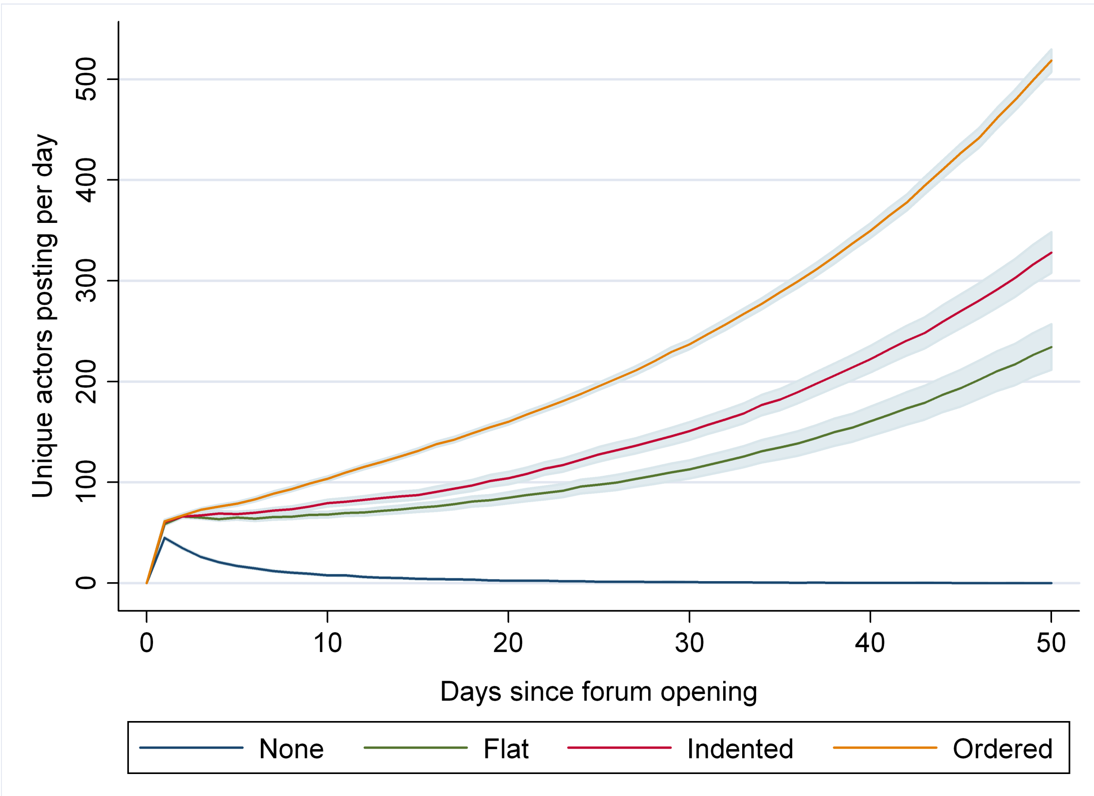


Figure 23: *Growth with 200 initial users.*

than larger ones (see figures 24 and 25). This is interesting, not only because it confirms that conversational reciprocity matters in addition to reply-bonuses, but also because it implies that who exactly receives a (reciprocated) reply matters. In the model, the only explanation for this can be that it allows those that have just invested in writing a post (-3 current desire), to recuperate their investment (+20 to next desire), before becoming inactive because of the loss. Threads thus not only make posting worthwhile by ensuring timely replies, but also — through mechanisms of conversational reciprocity — ensure that such replies are sent to the users that really need them to remain engaged.

The final thread-structure that is analysed, is the rating-ordered one. This is the structure used by HN. Posts in it are not ordered by ratings alone, but by a function of time and rating ($(r-1)/(t+2)^{1.8}$, with rating r, and t being the time in hours)[35]. Thus in addition to ratings, the rate at which posts receive votes, matters — giving newer posts a chance. Threads are ordered on the homepage using the same algorithm. This makes rating-ordered threads even better for readers, while being about the same for authors as unordered indented threads. The only group it disadvantages are users with minority interests, as it makes non-mainstream threads leave the homepage faster (see figures 26 and 27). Thus, if, as was found for BBSes, content-diversity is still important for attaining critical mass, then ordered threads would be at a disadvantage[93]. Nevertheless, by allowing most users an even better reading-experience, ordered threads perform better than all other thread-types, confirming H1 for the first stage; growth: thread-structures do improve engagement.

Arriving at the second stage, it is examined how many days it takes for critical mass to be attained when, in addition to engagement, individual social thresholds are taken into account. The thresholds are normally distributed, with a standard deviation of 15, around an average of 25. Every day a number of new actors are created (from 10 to 50), and sent to the forum. Only those actors (including those already on the forum) for whom the number of UUPDs is higher than their threshold, join the forum. The day-counter is stopped when a general threshold is met, here set to the average individual threshold (25 UUPDs), or when a simulated year has passed. The results are shown in table 3. As can be seen, threads do better than no-threads, and successively more advanced thread-structures do better, especially rating-ordered threads. Ordered threads consistently attain critical mass in less than 80 days, when 30 potential actors arrive every day, as opposed to almost six months for flat-threads (and only in 37% of cases). Thus also confirming H1 for when individual thresholds are considered; thread-structures do increase the chance of attaining critical mass, by better engaging users.

5 Discussion

Two hypotheses were tested in this paper. The secondary hypothesis (H2) was tested first, and served both to test the relative importance of structural- as opposed to social factors, and to inform the way reciprocity was added to the model. It consisted of two alternative scenarios (E1, E2, see section 1.2). No long-term, persistent tie-like reciprocity was found, so E1 was rejected. Instead, the replies that exceeded chance most were confined to the context of the thread, and short-term only, thus conversational (E2). Where conversational means that it follows the rules of conversation, in which, minding the context, one responds to utterings directed at (ones own, sometimes others') previous remarks.

And as with conversations at many a conference or party, making conversation does not necessarily constitute, or even lead to, a tie. Online communities might thus mostly

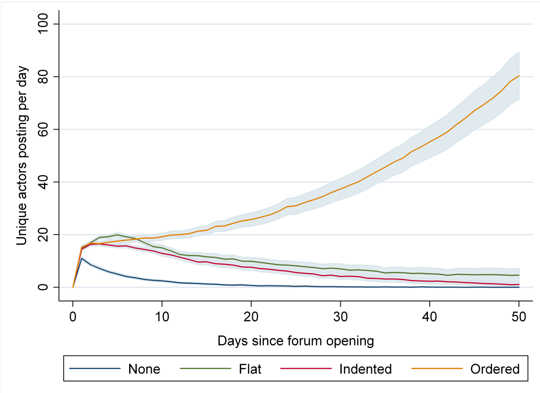


Figure 24: *Less growth without reciprocity. Compare to 21.*

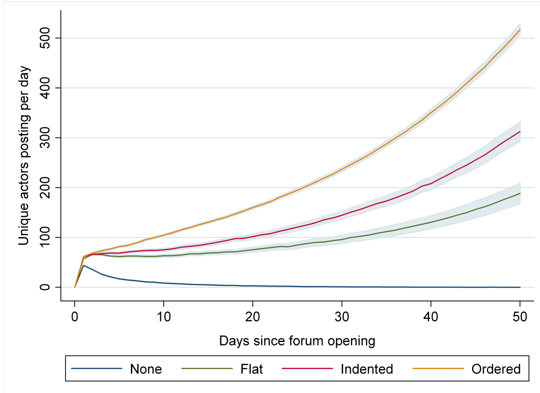


Figure 25: *Larger communities are impacted less by reciprocity.*

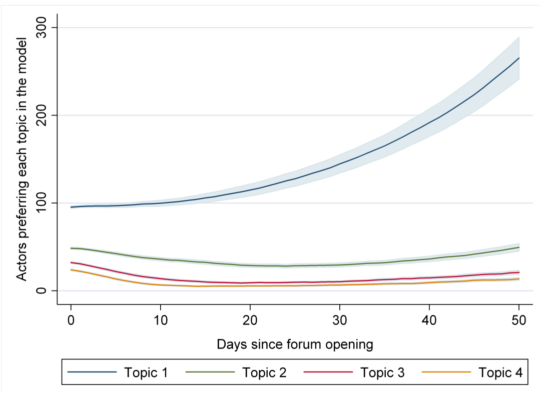


Figure 26: *Flat threads are relatively minority-friendly (top-4 topics shown).*

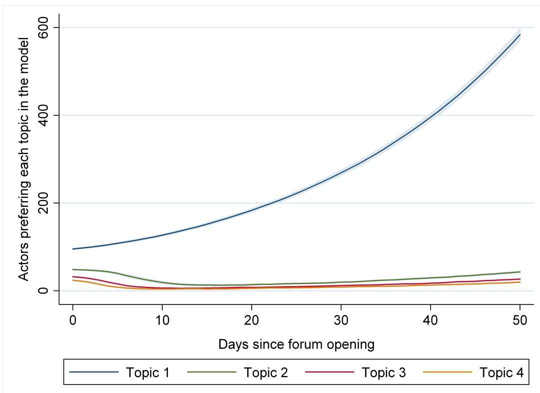


Figure 27: *Ordered threads exclude minorities.*

Table 3: *Critical mass attainment for the four thread-structures*

Arrivals	Days until attained (95% confidence interval) Lower	Average	Upper	Achieved
		T1: No threads		
10	X	X	X	0%
20	X	X	X	0%
30	X	X	X	0%
40	X	X	X	0%
50	X	X	X	0%
		T2: Flat threads		
10	X	X	X	0%
20	X	361	X	1%
30	144	174	204	37%
40	106	123	139	85%
50	76	89	102	100%
		T3: Indented threads		
10	X	X	X	0%
20	X	177	X	5%
30	188	154	222	33%
40	120	140	160	84%
50	76	89	102	96%
		T4: Ordered threads		
10	X	203	X	19%
20	135	149	164	93%
30	67	73	79	100%
40	45	49	53	100%
50	36	39	42	100%

thrive on second order communality. Which, to speak with Benedict Anderson, would make online forum communities, predominantly imagined communities[1]. Anderson uses the term to describe modern nations as 'communities' imagined by the people perceiving themselves as part of them. This as opposed to real communities, of people that know one-another personally, and interact on a regular basis. The short-lived nature of the reciprocity found on forums, and it in reality being mostly confined to threads, makes this a fitting analogy.

It, nevertheless, is likely that a small number of real and valuable forum-based relationships do exist, but that these are simply not detected among the noise of the many brief conversations. Though, for ties up to strength 8, a lower reciprocity was found than for those of strength 1, suggesting that most 'ties' come to be through pure chance, simply because some people post a lot, or enjoy reading about the same topics. Either way, these findings strongly question the assumption that reply-relationships are primarily social relationships, in the sense of persistent ties. And most relevant for modelling, are replies generated by structural factors; as most people interact with the topic under discussion, and follow the flow of the conversation, replying to posts, rather than authors. Naturally, this does not preclude people from deriving their 'reply bonus' from imagined social reciprocity. After all even lurkers identify with 'their' online communities[98,78].

In any case, as far as the facts are concerned, threads, and structural factors, such as the differences between thread-structures, were not a bad place to look for mechanisms explaining forums relative success at attaining critical mass. And while it should be noted again that the second part of the study cannot go beyond claims of generative sufficiency, threads were indeed found to be good at engaging users. In all cases they performed better than the baseline (no threads) model, both because of bundling (on pages, and of posts with similar topics), and timely replies brought about by conversational reciprocity. And as for differences between thread-structures, more advanced structures performed better, because they offered improved ways of skipping uninteresting posts. Finally, better engagement also translated into a faster attainment of critical mass when social threshold were introduced. Especially flat threads performed better than expected on this, possibly because they retained minorities longer, meeting more social thresholds early on.

6 Conclusion

To conclude, in the first section this paper set itself apart from previous work by focussing on threads and thread-structures, rather than social factors, as possible mechanisms behind the growth of online communities. Then the hypotheses were introduced. These were followed by a clarification of the way in which critical mass was defined here; namely as dependent on both social thresholds, and utility derived from forum-participation itself. Next, the research-method, ABS, was chosen and explained against a background of Structural Individualism, stressing the importance of thread-structures in the way individuals derive utility from forums. The ABS was used as an abstraction-layer between the data, and the analysis.

The secondary hypothesis was tested first; regarding to whether reciprocity in forums is social or conversational. Conversational reciprocity was found to be the most likely mechanism. The model was extended for this, and then used to test the primary hypothesis, whether thread-structures make a difference for the attainment of critical mass. It was found that they do, with rating-ordered threads performing best, but all thread-structures being beneficial for community growth. Thus — with the provision that only generative sufficiency was proven — it can be said that threads have a positive impact on the attainment of critical mass.

As for further research, there are many opportunities. The model could be informed by surveying, interviewing or recording people that regularly interact with forums. It could employ more advanced probability distributions, and be tweaked to make it better fit empirical data. Examining how reciprocity varies with community-size could also be valuable in this respect, and if extended, the beyond-chance measure used here, might prove useful for detecting meaningful social networks in a range of studies. Factors could be added to the model as well, such as malicious actors (trolls), emotional polarity of posts, and circadian cycles. The model could also be generalized, and extended to other domains; such as the modern blogosphere, or even the journal-based publication system.

A more practical use-case could be the testing of various possible improvements to forums. One such possible improvement may be showing new threads to new users first, allowing them to rake in the likely reply-bonus that comes with posting early. Another interesting endeavour could be to model communities in open competition with one-another. As competition could possibly explain why, in succession, flat threads, indented threads, and now rating-ordered threads, have grown in popularity. That is, initially most, or all forums provided flat threads, then indented threads won out against them, and now increasingly, by better being able to serve narrow audiences, rating-ordered threads attract many people interested in the niches for which they exist.

The continuous appearance of new on-line platforms, such as Twitter and Quora, also reminds us of the fact that the world of online communities is still very much in flux. What makes them tick, and determines their success, is still ill understood. But tapping into conversational reciprocity, and getting structural factors right, could be a large part of it. Even though providing users with a sense of community, will probably remain important as well, even if much of it is imagined.

7 Acknowledgements

I wish to thank Sandra Gonzalez-Bailon for the guidance and supervision that made this paper possible. In addition, I am grateful to the faculty and staff at the Oxford Internet Institute, for providing much insight and support during my coursework and research. I would also like to thank the organizers and presenters at the Social Networks seminar series at Nuffield College, and the CabBDyN series at the Said Business School, for sharing and discussing their work on ABS and other social modelling techniques in such a vibrant way. Finally, I am grateful to my employer, Richard Price, All Souls Prize Fellow, and CEO of Academia.edu, for giving me two days off a week this July, enabling me, new on the job in Silicon Valley, to finish this paper in time.

Bibliography

1. Anderson, B. R., *Imagined communities: Reflections on the origin and spread of nationalism.* (Verso Books, 2006).
2. Baldassarri, D., Hedstrom, P., & Bearman, P., in *The Oxford handbook of analytical sociology* (2009).
3. Beenen, G., 'Using social psychology to motivate contributions to online communities.', *Proceedings of the 2004 ACM conference on Computer supported cooperative work* 212–221 (2004).
4. Bimber, B., Flanagin, A. J, & Stohl, C., 'Reconceptualizing collective action in the contemporary media environment.', *Communication Theory* 15, 365–388 (2005).
5. Bishop, J., 'Increasing participation in online communities: A framework for human-computer interaction.', *Computers in Human Behavior* 23, 1881–1893 (2007).
6. *Boards.ie.* <http://boards.ie/>.
7. Boero, R., & Squazzoni, F., 'Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science.', *Journal of Artificial Societies and Social Simulation* 8, 6 (2005).
8. Bonabeau, E., 'Agent-based modeling: Methods and techniques for simulating human systems.', *Proceedings of*

the National Academy of Sciences of the United States of America 99, 7280 (2002).

9. Butler, B., Sproull, L., Kiesler, S., & Kraut, R., ‘Community effort in online groups: Who does the work and why.’, *Leadership at a Distance* 1 (2002).
10. Cattuto, C., ‘Semiotic dynamics in online social communities.’, *The European Physical Journal C-Particles and Fields* 46, 33–37 (2006).
11. Chen, G., & Chiu, M. M, ‘Online discussion processes: Effects of earlier messages’ evaluations, knowledge content, social cues and personal information on later messages.’, *Computers & Education* 50, 678–692 (2008).
12. Chen, H., Shen, H., Xiong, J., Tan, S., & Cheng, X., ‘Social network structure behind the mailing lists.’, *Proceeddings of the 15th Text REtrieval Conference (TREC 2006)* (2006).
13. Choi, H., Kim, S. H, & Lee, J., ‘Role of network structure and network effects in diffusion of innovations.’, *Industrial Marketing Management* 39, 170–177 (2010).
14. Choi, J. H, & Danowski, J. A, ‘Making a global community on the net - global village or global metropolis?: A network analysis of usenet newsgroups.’, *Journal of computer-mediated communication* 7, 0–0 (2002).
15. *Cohere.* <http://cohere.open.ac.uk/>.
16. Coleman, J. S, *Foundations of social theory.* (Belknap Press, 1994).
17. Creswell, J. W, & Clark, V. L., *Designing and conducting mixed methods research.* (Sage Publications, Inc, 2007).
18. Economides, N., & Himmelberg, C., ‘Critical mass and network size with application to the us fax market.’, (1995).
19. Elam, C., Stratton, T., Hafferty, F., & Haidet, P., ‘Identity, social networks, and relationships: theoretical underpinnings of critical mass and diversity.’, *Academic Medicine* 84 (2009).
20. Epstein, J. M, *Generative social science: Studies in agent-based computational modeling.* (Princeton Univ Pr, 2006).

21. Eynon, R., Fry, J., & Schroeder, R., ‘The ethics of internet research.’, *The SAGE handbook of online research methods* 23 (2008).

22. Fisher, D., Smith, M., & Welser, H. T, ‘You are who you talk to: Detecting roles in usenet newsgroups.’, *System Sciences, 2006. HICSS’06. Proceedings of the 39th Annual Hawaii International Conference on* 3, 59b (2006).

23. Freese, J., Hedstrom, P., & Bearman, P., in *The Oxford handbook of analytical sociology* 245–268 (2009).

24. Garrett, L. N., Smith, K. E., & Meyrowitz, N., ‘Intermedia: issues, strategies, and tactics in the design of a hypermedia document system.’, *Proceedings of the 1986 ACM conference on Computer-supported cooperative work* 163–174 (1986).

25. Garton, L., Haythornthwaite, C., & Wellman, B., ‘Studying online social networks.’, *Journal of Computer-Mediated Communication* 3, 0–0 (1997).

26. Gilbert, G. N, *Agent-based models.* (Sage Publications, Inc, 2008).

27. Gilbert, N., Pyka, A., & Ahrweiler, P., ‘Innovation networks-a simulation approach.’, *Journal of Artificial Societies and Social Simulation* 4, 1–13 (2001).

28. *GitHub.* <https://github.com/wybo/Forum-Tools>.

29. Goldenberg, J., Libai, B., & Muller, E., ‘The chilling effects of network externalities.’, *International Journal of Research in Marketing* (2009).

30. Gonzalez-Bailon, S., ‘Networks and mechanisms of interdependence: Theoretical developments beyond the rational action model.’, *Revista Internacional de Sociologia* 67 (2008).

31. Gonzalez-Bailon, S., Kaltenbrunner, A., & Banchs, R. E, ‘The structure of political discussion networks: A model for the analysis of online deliberation.’, *Journal of Information Technology* 25, 230–243 (2010).

32. Gonzalez-Bailon, S., 'The positive effects of negative emotions in online forums.', *Forthcoming* (2011).

33. *Google Wave.* <http://googleblog.blogspot.com/2010/08/update-on-google-wave.html>.

34. Gouldner, A. W, 'The norm of reciprocity: A preliminary statement.', *American sociological review* 161–178 (1960).

35. Graham, P., *Hacker News | How Hacker News ranking algorithm works.* <http://news.ycombinator.com/item?id=1781013>.

36. Graham, P., *Hacker News news.* <http://ycombinator.com/newsnews.html>.

37. Granovetter, M., 'Threshold models of collective behavior.', *American journal of sociology* 1420–1443 (1978).

38. Granovetter, M., & Soong, R., 'Threshold models of diversity: Chinese restaurants, residential segregation, and the spiral of silence.', *Sociological Methodology* 18, 69–104 (1988).

39. *Hacker News.* <http://news.ycombinator.com/>.

40. Hars, A., & Ou, S., 'Working for free?-motivations of participating in open source projects.', *34th Annual Hawaii International Conference on System Sciences* 7, 7014 (2001).

41. Hedstrom, P., *Dissecting the social: on the principles of analytical sociology.* (Cambridge University Press Cambridge, UK, 2005).

42. Hedstrom, P., 'Explaining the growth patterns of social movements.', *Understanding Choice, Explaining Behaviour* (2006).

43. Hedstrom, P., & Bearman, P., *The Oxford handbook of analytical sociology.* (Oxford University Press, USA, 2009).

44. Hedstrom, P., & Bearman, P., 'What is Analytical Sociology all about.', *The Oxford Handbook of Analytical Sociology. Oxford University Press. G. Manzo* 306 (2009).

45. Hedstrom, P., & Udehn, L., 'Analytical sociology and theories of the middle

range.', *The Oxford handbook of analytical sociology* 25–47 (2009).

46. Hedstrom, P., & Ylikoski, P., 'Causal mechanisms in the social sciences.', *Annual Review of Sociology* 36, 49–67 (2010).
47. Helbing, D., & Molnar, P., 'Social force model for pedestrian dynamics.', *Physical review E* 51, 4282–4286 (1995).
48. Hermann, C. F, 'Validation problems in games and simulations with special reference to models of international politics.', *Behavioral Science* 12, 216–231 (1967).
49. Hindman, M., *The myth of digital democracy.* (Princeton University Press, 2010).
50. Jansen, B. J., & Spink, A., 'How are we searching the World Wide Web? A comparison of nine search engine transaction logs.', *Information processing & management* 42, 248–263 (2006).
51. Johnson, S. L, & Faraj, S., 'Preferential attachment and mutuality in electronic knowledge networks.', *26th International Conference on Information Systems* 287–299 (2005).
52. Joyce, E., & Kraut, R. E, 'Predicting continued participation in newsgroups.', *Journal of Computer-Mediated Communication* 11, 723–747 (2006).
53. Kaltenbrunner, A., Gomez, V., & Lopez, V., 'Description and prediction of slashdot activity.', (2007).
54. Kaltenbrunner, A., Gonzalez-Bailon, S., & Banchs, R., 'Communities on the web: Mechanisms underlying the emergence of online discussion networks.', (2009).
55. Kaltenbrunner, A., 'Homogeneous temporal activity patterns in a large online communication space.', *Arxiv preprint arXiv:0708.1579* (2007).
56. Karau, S. J, & Williams, K. D, 'Social loafing: A meta-analytic review and theoretical integration.', *Journal of Personality and Social Psychology* 65, 681 (1993).
57. Karnstedt, M., 'Churn in social networks.', *Handbook of Social Network*

Technologies and Applications 185–220 (2010).

58. Kemper, A., *Valuation of network effects in software markets: A complex networks approach.* (Physica-Verlag Heidelberg, 2009).

59. Kiesler, S. E, *Culture of the Internet.* (Lawrence Erlbaum Associates Publishers, 1997).

60. Kim, A. J, 'Community building on the web: Secret strategies for successful online communities.', (2000).

61. Klamma, R., Spaniol, M., & Denev, D., 'Paladin: A pattern based approach to knowledge discovery in digital social networks.', *Proceedings of I-KNOW* 6, 457–464 (2006).

62. Kraut, R. E, & Resnick, P., 'Evidence-based social design: Mining the social sciences to build online communities.', (2011).

63. Kunegis, J., Lommatzsch, A., & Bauckhage, C., 'The slashdot zoo: Mining a social network with negative edges.', *Proceedings of the 18th international conference on World wide web* 741–750 (2009).

64. Lampe, C., & Johnston, E., 'Follow the (slash) dot: Effects of feedback on new members in an online community.', *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work* 11–20 (2005).

65. Lampe, C., & Resnick, P., 'Slash (dot) and burn: Distributed moderation in a large online conversation space.', *Proceedings of the SIGCHI conference on Human factors in computing systems* 543–550 (2004).

66. Lang, K., 'Newsweeder: Learning to filter netnews.', *in Proceedings of the 12th International Machine Learning Conference (ML95* (1995).

67. Lazar, J., & Preece, J., 'Social considerations in online communities: Usability, sociability, and success factors.', *Cognition in a digital world* 127–151 (2002).

68. Leimeister, J. M, Sidiras, P., & Krcmar, H., 'Success factors of virtual communities from the perspective of mem-

bers and operators: An empirical study.', *Proceedings of the 37th Hawaii International Conference on System Sciences* 7 (2004).

69. Macy, M. W, Flache, A., Hedstrom, P., & Bearman, P., in *The Oxford handbook of analytical sociology* 245–268 (2009).
70. Margetts, H., & John, P., 'Experiments for web science: examining the effect of the Internet on collective action.', (2009).
71. Margetts, H., John, P., Escher, T., & Reissfelder, S., 'How many people does it take to change a petition? Experiments to investigate the impact of on-line social information on collective action.', *5th ECPR General Conference: European Consortium for Political Research, Potsdam, Germany on* 10–12 (2009).
72. Marwell, G., & Oliver, P., *The critical mass in collective action: A micro-social theory.* (Cambridge Univ Pr, 1993).
73. McKenna, K. Y. A., 'Influences on the nature and functioning of online groups.', *Psychological aspects of cyberspace: Theory, research, applications* 228–242 (2008).
74. Montanari, A., & Saberi, A., 'The spread of innovations in social networks.', *Proceedings of the National Academy of Sciences* 107, 20196 (2010).
75. Nelson, T. H., *Literary machines : the report on and of, project Xanadu, concerning word processing, electronic publishing, hypertext, thinkertoys ...* (1992).
76. *Nokogiri.* <http://nokogiri.org/>.
77. Nonnecke, B., & Preece, J., 'Lurker demographics: Counting the silent.', *Proceedings of the SIGCHI conference on Human factors in computing systems* 73–80 (2000).
78. Nonnecke, B., & Preece, J., 'Why lurkers lurk.', *Americas Conference on Information Systems, Boston* (2001).
79. Novak, T. P, Hoffman, D. L, & Yung, Y. F, 'Measuring the flow construct in online environments: A structural modeling approach.', *Marketing science* 19, 22–42 (2000).

80. *Ohloh.* <http://www.ohloh.net/>.

81. Oliver, P. E., & Marwell, G., 'Whatever happened to critical mass theory? A retrospective and assessment.', *Sociological theory* 19, 292–311 (2001).

82. Onnela, J. P, & Reed-Tsochas, F., 'The spontaneous emergence of social influence in online systems.', *Arxiv preprint arXiv:0912.0045* (2009).

83. Ortega, F., & Gonzalez Barahona, J. M, 'Quantitative analysis of the wikipedia community of users.', *Proceedings of the 2007 international symposium on Wikis* 75–86 (2007).

84. Osborne, L. N, 'Topic development in usenet newsgroups.', *Journal of the American society for information science and technology* 49, 1010–1016 (1998).

85. Outhwaite, W., & Turner, S. P, *The SAGE handbook of social science methodology.* (SAGE Publications Ltd, 2007).

86. Petersen, T., Hedstrom, P., & Bearman, P., in *The Oxford handbook of analytical sociology* (Oxford University Press, 2009).

87. Platt, J., 'Social traps.', *American Psychologist* 28, 641–651 (1973).

88. Prasarnphanich, P., & Wagner, C., 'Creating critical mass in collaboration systems: Insights from Wikipedia.', *2nd IEEE International Conference on Digital Ecosystems and Technologies, 2008. DEST 2008* 126–130 (2008).

89. Prasarnphanich, P., & Wagner, C., 'Explaining the sustainability of digital ecosystems based on the wiki model through critical mass theory.', (2009).

90. Preece, J., 'Sociability and usability in online communities: determining and measuring success.', *Behaviour & Information Technology* 20, 347–356 (2001).

91. Preece, J., Nonnecke, B., & Andrews, D., 'The top five reasons for lurking: Improving community experiences for everyone.', *Computers in Human Behavior* 20, 201–223 (2004).

92. *Quora.* <http://www.quora.com/>.

93. Rafaeli, S., & LaRose, R. J, 'Electronic bulletin boards and "public goods" explanations of collaborative mass media.', *Communication Research* 20, 277 (1993).

94. Ren, Y., Kraut, R., Kiesler, S., 'Applying common identity and bond theory to design of online communities.', *Organization studies* 28, 377–408 (2007).

95. Ren, Y., & Kraut, R. E, 'A simulation for designing online community: Member motivation, contribution, and discussion moderation.', *Information Systems Research* (2011).

96. Ren, Y., & Kraut, R. E, 'Agent-based modeling to inform online community theory and design: Impact of discussion moderation on member commitment and contribution.', (2011).

97. Ren, Y., 'Increasing commitment to online communities: Designing from theory.', *University of Minnesota, Minneapolis, MN* (2010).

98. Roberts, T. L, 'Are newsgroups virtual communities?', *Proceedings of the SIGCHI conference on Human factors in computing systems* 360–367 (1998).

99. Ruggiero, T. E, 'Uses and gratifications theory in the 21st century.', *Mass Communication & Society* 3, 3–37 (2000).

100. Schelling, T. C, 'Dynamic models of segregation.', *The Journal of Mathematical Sociology* 1, 143–186 (1971).

101. Schelling, T. C, *Micromotives and macrobehavior.* (WW Norton & Company, 2006).

102. Sewell Jr, W. H, 'A theory of structure: Duality, agency, and transformation.', *American journal of sociology* 1–29 (1992).

103. *SIOC.* <http://data.sioc-project.org/>.

104. Skinner, B. F., 'Verbal behavior.', *Language* 35 (1957).

105. *Slashdot.* <http://slashdot.org/>.

106. Smith, M. A, 'Netscan: Measuring and mapping the social structure of usenet.', (1997).

107. Smith-David, J., 'Social networks on both sides of the transition point.', *Elec-*

tronic Commerce Research and Applications (2009).

108. *Stata.* <http://www.stata.com/>.

109. Tedjamulia, S. J., Dean, D. L, Olsen, D. R, & Albrecht, C. C, 'Motivating content contributions to online communities: Toward a more comprehensive theory.', *38th Annual Hawaii International Conference on System Sciences* (2005).

110. Thaler, R. H, & Sunstein, C. R, *Nudge: Improving decisions about health, wealth, and happiness.* (Yale Univ Pr, 2008).

111. Valverde, S., & SolÃl', R. V, 'Self-organization versus hierarchy in open-source social networks.', *Physical review* 76, 046118 (2007).

112. Van Zandt, T., 'Information overload in a network of targeted communication.', *RAND Journal of Economics* 35, 542–560 (2004).

113. Varian, H. R, *Intermediate microeconomics: a modern approach.* (WW Norton New York, 1993).

114. Varian, H. R, Farrell, J., & Shapiro, C., *The economics of information technology: An introduction.* (Cambridge Univ Pr, 2004).

115. Wagner, C., Liu, L., Schneider, C., Prasarnphanich, P., & Chen, H., 'Creating a successful professional virtual community: A sustainable digital ecosystem for idea sharing.', *Digital Ecosystems and Technologies, 2009. DEST'09. 3rd IEEE International Conference on* 163–167 (2009).

116. Wagner, C., & Prasarnphanich, P., 'Innovating collaborative content creation: the role of altruism and wiki technology.', *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on* 18–18 (2007).

117. Wellman, B., 'An electronic group is virtually a social network.', *Culture of the Internet* 179–205 (1997).

118. Wellman, B., 'Computer networks as social networks: Collaborative work, telework, and virtual community.', *Annual review of sociology* 213–238 (1996).

119. Welser, H. T, Gleave, E., Fisher, D., & Smith, M., ‘Visualizing the signatures of social roles in online discussion groups.’, *Journal of Social Structure* 8 (2007).

120. Westland, J. C, ‘Critical mass and willingness to pay for social networks.’, *Electronic Commerce Research and Applications* 9, 6–19 (2010).

121. Wiersma, W., ‘LogiLogi: The quest for critical mass.’, *Unpublished thesis* (2010).

122. Wiersma, W., ‘The invisible hands of time: How timezones shape online communities.’, *LogiLogi Foundation blog* (2011).

123. Wiersma, W., & Sarlo, B., ‘LogiLogi: a webplatform for philosophers.’, *Digital Humanities 2008 Book of Abstracts* 221–222 (2008).

124. Wilkinson, D. M, & Huberman, B. A, ‘Cooperation and quality in wikipedia.’, *Proceedings of the 2007 international symposium on Wikis* 157–164 (2007).

125. Zhang, J., Ackerman, M. S, & Adamic, L., ‘Communitynetsimulator: Using simulations to study online community networks.’, *Communities and Technologies 2007* 295–321 (2007).

126. Zhongbao, K., & Changshui, Z., ‘Reply networks on a bulletin board system.’, *Physical Review E* 67, 036117 (2003).

127. Zittrain, J., ‘The rise and fall of sysopdom.’, *Harvard Journal of Law & Technology* 10, 495 (1996).