



Article

Poverty Traps in Online Knowledge-Based Peer-Production Communities

Andrew Vargo ^{1,*} , Benjamin Tag ² , Chris Blakely ³ and Koichi Kise ¹

¹ Department of Core Informatics, Graduate School of Informatics, Osaka Metropolitan University, Sakai 599-8531, Japan

² Department of Human Centred Computing, Faculty of Information Technology, Monash University, Clayton 3800, Australia

³ School of Applied Information Technology, The Kyoto College of Graduate Studies for Informatics, Kyoto 606-8225, Japan

* Correspondence: awv@omu.ac.jp

Abstract: Online knowledge-based peer-production communities, like question and answer sites (Q&A), often rely on gamification, e.g., through reputation points, to incentivize users to contribute frequently and effectively. These gamification techniques are important for achieving the critical mass that sustains a community and enticing new users to join. However, aging communities tend to build “poverty traps” that act as barriers for new users. In this paper, we present our investigation of 32 domain communities from Stack Exchange and our analysis of how different subjects impact the development of early user advantage. Our results raise important questions about the accessibility of knowledge-based peer-production communities. We consider the analysis results in the context of changing information needs and the relevance of Q&A in the future. Our findings inform policy design for building more equitable knowledge-based peer-production communities and increasing the accessibility to existing ones.

Keywords: collaboration; peer production; reputation systems; virtual communities



Citation: Vargo, A.; Tag, B.; Blakely, C.; Kise, K. Poverty Traps in Online Knowledge-Based Peer-Production Communities. *Informatics* **2023**, *10*, 61. <https://doi.org/10.3390/informatics10030061>

Academic Editor: Long Jin

Received: 10 May 2023

Revised: 26 June 2023

Accepted: 10 July 2023

Published: 13 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Online knowledge-based peer-production communities often rely on reputation systems to encourage high-quality contributions [1], promote collaboration on difficult problems [2,3], and facilitate effective community management [4,5]. The idea is simple; the reputation system manages the community and removes the burden of top-down administration. Instead of using authoritarian moderators, the community can vote on, edit, and remove content and users. This is beneficial for any system that seeks to have a large scale of users and content, as the crowd’s actions mitigate the costs and efforts of managing the community. However, one challenge to this type of system is whether it is accessible to new users.

Many communities, among which Stack Exchange is probably the most popular, exhibit reputation pyramids with the power law in effect [6]. This means that very few users own the majority of reputation, which is problematic for the continuation of a reputation system’s ability to motivate users. The power law is considered a natural occurrence, where it is expected that a small group of users will contribute the majority of content in any collaborative, knowledge-based peer-production community [7]. It makes sense that the users who contribute the most would have the most reputation, especially if these are the best users.

A concern for these kinds of collaborative communities is to make sure that the entry barrier is as low as possible for new entrants and that situations where there is a digital “poverty trap” are avoided, e.g., where newer users can never compete against entrenched users. Since earning reputation is an important motivating factor for continuous

contributions [1,3], it would make sense that these communities optimize the ability to help new entrants compete in the knowledge-collaboration system. If skill alone is the barrier for new entrants, there is little that these communities can do beyond their current formation. However, if there is a structural problem with the decreasing reputation point value of Q&A exchanges, there might be potential to introduce new design mechanisms.

We analyzed a cross-section of 32 communities from Stack Exchange. We found that all communities showed a positive relationship between how early in the communities' life cycle the question was asked and the amount of reputation earned by a question. We found that a longer membership had a positive predictive relationship within 31 of the 32 communities for answerers. Within these results, our analysis elicits significant differences between the types of communities based on the focus of the domain. In particular, we found that the permanence of information sought within a community was closely tied to the bias towards longer tenure.

The main contributions of this work are as follows:

- Our analysis presents an important practical approach towards understanding reputation systems for collaborative knowledge-based peer-production communities by treating reputation points as currency and Q&A exchanges as scarce resources.
- The results raise important questions about the structure of reputation systems and move toward developing mechanisms that increase equity and lower the barrier for new users.
- We present a set of design recommendations to inform the creation and maintenance of information-sharing and aggregation communities, such as Q&A sites, based on our findings.

To the best of our knowledge, this research is the first of its kind to directly investigate and identify the existence of poverty traps within online knowledge-based peer-production communities.

2. Background

Q&A communities, which are some of the most prominent types of knowledge-based peer-production communities online, have seen the emergence of several new types in recent years. One type of Q&A community that is known for its archival value and community structure is the collaborative Q&A model on Stack Exchange, where users can ask, answer, and edit questions [2,4]. The benefit of this type of forum lies in its ability to tap into the vast array of experts in the field who can provide help. This applies not only to the questioner but also to other members of the community and outside users [2,8,9].

2.1. Stack Exchange Overview

Stack Exchange is a network of Q&A sites that run and maintain the same reputation system mechanics. There are over 100 sites dedicated to specific domains, where users are expected to ask and answer questions around narrowly defined topics (<https://stackexchange.com/about>; (accessed on 28 March 2023)). All the sites share the same reputation system, which has remained remarkably consistent since its inception with Stack Overflow in 2008. Users are motivated by reputation points, whether they recognize the system's impact or not [3], and experts are motivated to answer questions quickly [1]. Anyone can propose new sites in *Area 51* (<https://area51.stackexchange.com/?tab=beta> (accessed on 1 May 2023)), a quarantine area where new communities can be discussed and tested. In *Area 51*, the community can collaborate and, if successful, facilitate the launch of the new community through a beta phase to a fully-fledged Q&A site. This community involvement process has been cited as a reason for Stack Exchange's success, especially with its biggest site, Stack Overflow [5,10], which boasts over 20 million users.

Within the Stack Exchange ecosystem, there are a number of actors within the user community. The first actor is the question-asker, who typically seeks information from others. The nature of this information request (i.e., concrete or theoretical) can vary within communities [11]. Next, we have the question-answerer. This actor attempts to provide

information to the asker and the entire community, which helps to resolve the question. Answers can be driven by the desire to earn reputation points within the community and by an altruistic motive to help the asker and the community [2,3,11]. Finally, there are actors that act as community moderators. These moderation tools are given as privileges to users as they gain reputation; these privileges include the ability to vote, edit, and comment on questions and answers. Most editing appears to be an altruistic act for the benefit of the community [12].

Stack Exchange's success has elevated it to the status of an interactive community. Users can build complex profiles to show their achievements; top users are the sources of many professional and academic studies, and the system has an effective job-hunting site connected to the profiles. Consequently, these features mean that the performance of a Stack Exchange site could have real-life impacts on the users.

For these reasons, it is important to understand whether Stack Exchange domains have "poverty traps", i.e., situations where individuals or groups lack the necessary capital to escape poverty. The concept of a *poverty trap* in this work revolves around the idea that there is a structural bias towards rewarding longer tenure within a community. In an online Q&A community, we can see reputation as a form of currency that provides status and, in the case of Stack Exchange, power. Q&A contributions are investment vehicles where effort is either rewarded or punished. "Owning" part or all of the most valuable investment vehicles is important to accruing this currency. If the most valuable Q&A exchanges occur on a descending linear timeline, i.e., earlier in the community's history, then there is a clear poverty trap that can occur just by a user having a later start date in the community.

There are multiple reasons why poverty traps in these communities, and communities like these, are troublesome, including the possibility of institutionalizing poor and privileged users and discouraging new users from joining. While individual users suffer, a poverty trap could very well help to define what experts in a community look like. For instance, in Stack Overflow's case, a small, interconnected community helps to develop and propagate the system [5]. The result is a tightly knit community where different users have specialized roles within the community [13]. The community has spread throughout the world, but there remains an entrenched group of users who have been part of the community since its origins.

2.2. Reputation Systems

There are multiple ways of implementing a reputation system. Wikimedia, a producer of online collaborative encyclopedias, has a content-based reputation system where approved users rate and edit content, which then indirectly creates the reputation status for the content originator [14,15]. A clear problem with this type of system is that it leads to calcification, where new users find it difficult to contribute and compete with more entrenched users [16]. Another method is to deploy a reputation system where contributions are under constant peer evaluation by the community. Users gain or lose reputation points through a voting system from their peers. Points indicate the ranking and status of the individual users [8] and can then be used to grant privileges to control community elements [5]. These types of reputation systems have proven to be effective in forming the *Stack Exchange* network of Q&A sites [5,6,17].

2.3. Stack Exchange Reputation System

Reputation points have a lot of meaning. Several studies show that reputation point aggregation is an accurate predictor of expertise in technical domains [1,3], although this is not always the case, as some users earn many reputation points from a single contribution [18]. This includes situations in which payment with money could be a confounding factor. Two analyses of Google Answers found no relationship between the price offered and the quality of the answer. Instead, there was a relationship between the quality and the reputation of the answerer [19,20]. The better the track record of the answerer, the better the answer. Peer evaluation is an effective way of identifying experts.

It is also established that expert users are motivated by earning reputation points [3], regardless of whether they are explicitly aware of the fact. They are also extremely adept at targeting and answering the highest reputation-earning questions available [1].

The reputation system in Stack Exchange consists of a voting system that moderates content and assigns points based on the quality of the submission. Most points are earned through material voted up by the rest of the community, especially answers [1]. Users can aggregate a maximum of 200 points per day via answer upvotes. This means that a popular answer or question may receive votes that will not raise the user's reputation score. This acts as both a cap on users benefiting from legacy exchanges and a barrier preventing new users from catching up to their older counterparts. Reputation in Stack Exchange can be used as a currency and as a way to gauge and manage a user's moderating abilities on Stack Exchange. Reputation is supposed to act as a measure of a user's ability to ask and answer, as well as their trustworthiness [21]. The forum implements a privilege scheme where users gain privileges by reaching certain point totals, as shown in Table 1. For instance, voting up and down requires 15 and 125 reputation points, respectively. As the moderating power of the privilege increases, the points needed to achieve the privilege increases. At 20,000 points, Stack Exchange gives the title of "Trusted User" to a member, thus conferring an elite moderator status on the member [22]. Stack Exchange also has rankings for smaller time periods, such as months or quarters. While this allows new users to compete with older users over a smaller time frame, older users can easily finish towards the top of the rankings without contributing within that period. If older Q&A interactions are still valued by the community, users can gain up to 200 reputation points per day from these legacy interactions.

Table 1. Stack Exchange privilege scheme [23].

Privilege	Reputation Points Needed to Earn Privilege
Create Posts	1
Participate in Meta	5
Create Wiki Posts	10
Remove New User Restrictions	10
Vote Up	15
Flag Posts	15
Talk in Chat	20
Comment Everywhere	50
Set Bounties	75
Create Chat Rooms	100
Edit Community Wiki	100
Vote Down	125
Reduce Ads	200
View Close Votes	250
Access Review Queues	500
Create Gallery Chat Rooms	1000
Established User	1000
Create Tags	1500
Edit Posts (Questions and Answers)	2000
Create Tag Synonyms	2500
Cast Close and Reopen Votes	3000
Approve Tag Wiki Edits	5000
Access Moderator Tools	10,000
Protect Questions	15,000
Trusted User	20,000
Access to Site Analytics	25,000

2.4. Types of Q&A on Stack Exchange

Most Q&A sessions on Stack Exchange are very straightforward, especially when considering registered users. If applicable, a user will ask questions and then choose the

best answer. Other eligible users can vote on the quality of the question and other answers. While users have agency over their votes, some basic content rules exist. For example, on Stack Overflow, all questions should adhere to the following:

- Questions should be unique to the programming profession (a specific programming problem, a software algorithm, software tools commonly used by programmers, etc.).
- Questions should not be based on opinion (e.g., what do you think is better: “x” or “y”).
- Questions should be reasonably scoped [24].

All sites on Stack Exchange communities follow similar rules. For instance, the “Mathematics” community allows for discussing the history of mathematics but limits questions that would have too many possible answers. A consistent feature across all Stack Exchange sites is the emphasis on questions that can be answered factually and concisely by the community.

A potential problem for this type of system is that “valuable” questions are scarce. Treude et al. [11] found that there were two major types of questions: concrete and theoretical and that most of the answered questions had to do with practical problems. Compared to larger theoretical questions, most ask for solutions to task-based problems. Concrete questions were more likely to receive good and timely answers. This is most likely because these questions have an impact beyond just helping the questioner and fulfilling a desire to help a larger number of people [1,25]. Combined with the above rules, this would severely limit the type and number of good questions to answer, potentially leading to a self-perpetuating system. This can be problematic for a community since new question askers may leave if they do not receive motivation to stay, thus denying an essential element of the system [26].

2.5. Prior Findings and Reputation Systems

The reputation pyramid is thought to result from a specialized skill in which top users are adept at noticing interesting questions and answering them quickly [1]. In addition, researchers have found that users who create accounts with information filled in are more likely to aggregate reputation and be more efficient at it [27,28]. These findings indicate that good users who put effort into their accounts are likelier to climb the reputation pyramid. On the other hand, researchers found that the membership length had the biggest effect on reputation-earning ability, compared to other profile factors that would indicate effort put into the account [29]. This finding could be problematic since it indicates a structural issue with the reputation system. A possible explanation is that the most valuable information is created early on in a domain’s existence. This would make sense for narrow domains in which users are trying to solve concrete problems, where the first possible answer that solves their question is the most valuable [11,30]. In a case like this, we expect that many people have the same questions and, therefore, will refer to a previously asked question rather than post their questions. On the other hand, narrow domains in which questions require more open-ended answers may not be as affected by time.

2.6. Poverty Traps in Online Knowledge-Based Peer-Production Communities

To the best of our knowledge, no previous literature directly addresses the nature of poverty traps within online knowledge-based peer-production communities. Within the research community focusing on these types of communities, there is an understanding that power law is endemic [1,5,7]. This implies that a few users will have most of the reputation and that the reputation distribution will have a pyramid shape. The motivation for this research, in particular, comes from Vargo and Matsubara [29], where it was found that, when controlling for membership tenure, there was no difference in reputation-earning efficiency between different types of profile constructions. Previous work [27,28] had shown that having different profile features completed aligned with higher reputation earning but had not considered tenure as a confounding factor. That is, when comparing within similar groups, the length of tenure was the most important factor for earning reputation.

The result found by Vargo and Matsubara indicated that there might be a structural barrier within the knowledge-based peer-production system itself.

2.7. Theoretical Analysis

Based upon the previous research, it is known that online knowledge-based peer-production communities that use incentive systems typically produce a reputation-pyramid in which a few users have most of the points [1,7]. This is likely because there is a dedicated and talented group of users who provide most of the valuable contributions [1,3,12]. However, there is evidence that there is a built-in structural barrier based upon tenure within peer-production activity, which helps to produce a poverty trap [29]. This analysis seeks to understand the relationship between time and reputation-earning ability. In other words, it investigates whether longer tenure provides a relative benefit to a user.

3. Methodology

It is possible that the most valuable contributions from users in an online community are conducted early on. Therefore, we want to examine if there is a time bias in knowledge creation and if this influences the reputation system of the community. In order to do this, we set the following research questions:

1. Do communities exhibit value bias towards earlier Q&A submissions?
2. Does the strength and/or existence of this bias depend on the domain?
3. Does the existence of a time bias influence the amount of reputation a participant can earn?

We set out to investigate the effect of time on reputation earning. To do this, we looked at 32 narrow domain sites across four general categories, namely *Technology*, *Science*, *Life/Arts*, and *Culture/Recreation*. To ensure a base level of popularity in the community, we chose sites with at least 50,000 visits per day and conducted two exploratory analyses.

Analyzing a Q&A community for poverty traps is difficult due to a number of confounding factors. First, there needs to be a consideration, where each domain has its own cycle of information creation. That is, while the amount of possible questions that could be asked in a community likely increases over time, the rate at which it increases may be directly tied to instability within the domain. To understand each community, we look at the relationship between questions and when they are asked. This gives us an understanding of the shape of knowledge-creation within each community and whether there is a bias towards more valuable questions being asked at the beginning of a community's lifespan.

The next confounding factor is the skill possessed by each participant. Skill can be seen as a two-fold attribute. The first aspect is the ability to provide correct and helpful content to the community and information seekers. This aspect may require domain expertise, the ability to write clearly and concisely and to intuit what information seekers want. The other aspect of skill centers around the ability to master the nuances of the community's infrastructure. This includes being able to comply with the community's norms, use the system architecture efficiently, and identify valuable interaction opportunities.

To summarize, users could be domain experts. Still, if they cannot use the community's system architecture well, they will likely not be able to answer quickly enough to earn reputation. Likewise, users could be experts at using the platform but not have the domain expertise to provide popular or valuable contributions. To remove this influence, we analyzed a subset of users who have given 10–20 answers at least one-year-old and have a positive reputation. This allowed us to include users who have had some success on the site but were not power users or spam users.

The research was conducted as follows. First, we processed the data by scraping community data from the TSQL databases, which are provided by Stack Exchange (<https://data.stackexchange.com>, accessed on 9 May 2023). The specific queries used in this research can be found in this manuscript's *Data Availability Statement*. Secondly, data were then analyzed with the R language in *RStudio* (<https://posit.co>, accessed on 9 May 2023).

4. Data Set

Stack Exchange has many communities, both small and large. We sampled 32 out of 68 English language communities with over 50,000 registered users. This was conducted to make sure that the analyzed communities had reached a critical mass of content production and community interaction, where the community was stable and self-sustaining. In particular, the threshold was chosen to assure lively communities in which there were not only power users but users who made smaller contributions that were vital to the survivability of a community [31]. The most visited site by far was *Stack Overflow*, with over 20 million registered users. With *Stack Overflow* removed, the average number of users was 335,000, with a standard deviation of 361,000 users.

All communities in Stack Exchange belong to a category of similar domains. In our data set, the communities belong to *Technology*, *Science*, *Culture/Recreation*, and *Life/Arts*. For instance, *Stack Overflow* belongs to the *Technology* category, while *Travel* belongs to *Culture/Recreation*. The categories and basic data for all 32 communities are presented in Table 2.

Table 2. Stack Exchange community information [32].

Site	Category	Users (In Thousands)	Site Age in Months
Android Enthusiasts	Technology	279	151
Arqade (Video games)	Culture/Recreation	202	153
Ask Difference	Technology	391	152
Ask Ubuntu	Technology	1389	153
Chemistry	Science	99	132
Code Review	Technology	234	147
Cross-Validated (Statistics)	Science	324	153
Database Administrators	Technology	233	147
Electrical Engineering	Technology	269	151
English Language and Usage	Culture/Recreation	378	152
English Language Learners	Culture/Recreation	141	123
Geographic Information Systems (GIS)	Technology	182	153
Graphic Design	Life/Arts	133	147
Home Improvement	Life/Arts	119	153
Information Security	Technology	239	149
Mathematics	Science	1018	153
Motor Vehicle Maintenance and Repair	Culture/Recreation	55	145
Movies and TV	Life/Arts	72	137
Physics	Science	287	149
Role-Playing Games	Culture/Recreation	62	152
Salesforce	Technology	102	129
Science Fiction and Fantasy	Life/Arts	127	147
Seasoned Advice (Cooking)	Life/Arts	74	153
Server Fault	Technology	874	168
Software Engineering	Technology	363	151
Stack Overflow	Technology	20,173	177
Super User	Technology	1464	165

Table 2. Cont.

Site	Category	Users (In Thousands)	Site Age in Months
TeX–LaTeX	Technology	250	153
Travel	Culture/Recreation	100	142
Unix and Linux	Technology	516	152
Web Applications	Technology	228	154
WordPress Development	Technology	188	152

5. Results

In this project, we examine three research questions:

1. Do communities exhibit value bias towards earlier Q&A submissions?
2. Does the strength and/or existence of this bias depend on the domain?
3. Does the existence of a time bias influence the amount of reputation a participant can earn?

We first examine whether there is any bias towards earlier Q&A submissions. Our findings show that a community's tendency toward value bias is dependent on the domain. The degrees of strength of such a value bias vary across categories, meaning that different domains within similar categories (i.e., *Technology*, etc.) value time and the age of a contribution differently. Upon finding this bias to occur in these communities, we then examine whether this bias influences user reputation scores that can lead to poverty traps.

5.1. The Value of Early Q&A Submissions

To understand the relationship between each community and the potential bias towards popular questions being asked early, we sampled questions from each community, which had at least one aggregate upvote and was at least 365 days old. We then ran a non-parametric correlation between the age of the question (in months) and the number of upvotes it had received. The reason for removing questions with negative or zero aggregate upvotes was to remove spam and redundant questions. Thus, we are only considering questions that have some value according to the community after one year of existence. The results, as shown in Table 3, reveal that all communities have a significant correlation between the age of the question and the number of upvotes. However, there are clear differences between communities. For instance, *Motor Vehicle Maintenance and Repair* exhibits a strong relationship, with a coefficient near 0.29, while *Stack Overflow* has a coefficient of 0.0925.

Table 3. Correlation results between the time when a question was asked and the aggregate number of upvotes received. The table is ordered by descending correlation coefficients.

Site	Category	Sample Size	Correlation Coefficient
Motor Vehicle Maintenance and Repair	Culture/Recreation	19,539	0.289 *
Seasoned Advice (Cooking)	Life/Arts	21,665	0.2274 *
Chemistry	Science	26,217	0.2226 *
Geographic Information Systems (GIS)	Technology	40,147	0.2188 *
TeX–LaTeX	Technology	43,220	0.2125 *
English Language Learners	Culture/Recreation	34,563	0.2106 *
Arqade (Video games)	Culture/Recreation	28,379	0.2104 *
Science Fiction and Fantasy	Life/Arts	36,634	0.2067 *
Electrical Engineering	Technology	37,276	0.1948 *

Table 3. Cont.

Site	Category	Sample Size	Correlation Coefficient
Salesforce	Technology	36,227	0.1934 *
Home Improvement	Life/ Arts	29,771	0.1882 *
Graphic Design	Life/ Arts	20,605	0.1855 *
Web Applications	Technology	23,700	0.183 *
Android Enthusiasts	Technology	43,230	0.1786 *
Cross-Validated (Statistics)	Science	40,823	0.176 *
Physics	Science	39,885	0.1738 *
Movies and TV	Life/ Arts	20,249	0.1696 *
Code Review	Technology	36,243	0.1634 *
WordPress Development	Technology	32,729	0.1627 *
Unix and Linux	Technology	42,060	0.1551 *
Travel	Culture/Recreation	29,461	0.1541 *
Software Engineering	Technology	32,603	0.141 *
Database Administrators	Technology	35,419	0.1394 *
English Language and Usage	Culture/Recreation	38,419	0.1388 *
Mathematics	Science	46,034	0.1257 *
IT Security	Technology	32,574	0.1171 *
Role-Playing Games (RPG)	Culture/Recreation	30,778	0.1158 *
Ask Ubuntu	Technology	44,707	0.1122 *
Ask Different	Technology	38,594	0.1082 *
Super User	Technology	43,343	0.1074 *
Server Fault	Technology	44,145	0.0948 *
Stack Overflow	Technology	48,109	0.0925 *

* $p < 0.05$.

Overall, the correlation coefficients for the communities are relatively small; however, this does not mean that the results are not meaningful. We would expect that the quality of the question would be the most important factor for reputation aggregation, and in fact, should render time as a non-significant factor, unless high-quality questions are more likely to be asked earlier in a community's lifespan.

5.2. The Value of Early Questions in Each Community

The results, as shown in Figure 1, reveal that the community's Q&A structures are non-monolithic. We can see that *Motor Vehicle Maintenance and Repair* is significantly different from all other categories, suggesting that there is a stronger prevalence for valuable questions to have been asked earlier in the community's life compared to other communities. On the other hand, *Stack Overflow* has a lower relationship with time, with the exception for the four communities from the *Technology* category.

We analyze the differences between communities by comparing the correlation results between all of the communities. We compare the results using *cocor* [33] to compare the correlations with Fisher's Z. We account for multiple comparisons between communities by applying a Bonferroni Correction. Zou's confidence interval is also used to confirm a significant result.

What is most surprising, however, is that communities that seem to have the same format have significantly different correlations. For instance, we might expect that *Android Enthusiasts* and *Ask Different* might have no difference due to the similarity of community

focus on the popular operating system. However, we can see that *Ask Android* has a significantly higher correlative value. There are a number of reasons why this might be: the development cycles of the technology itself, the norms of the community, or the type of users drawn to the community (e.g., information-searching behavior). In addition, we see that, overall, there is little relationship between the category of the community and the result. This may be because the type of information that is sought in each community is more important than the overall genre of the community. While the exact reasons for the differences between the communities are interesting, they go beyond the scope of this paper.

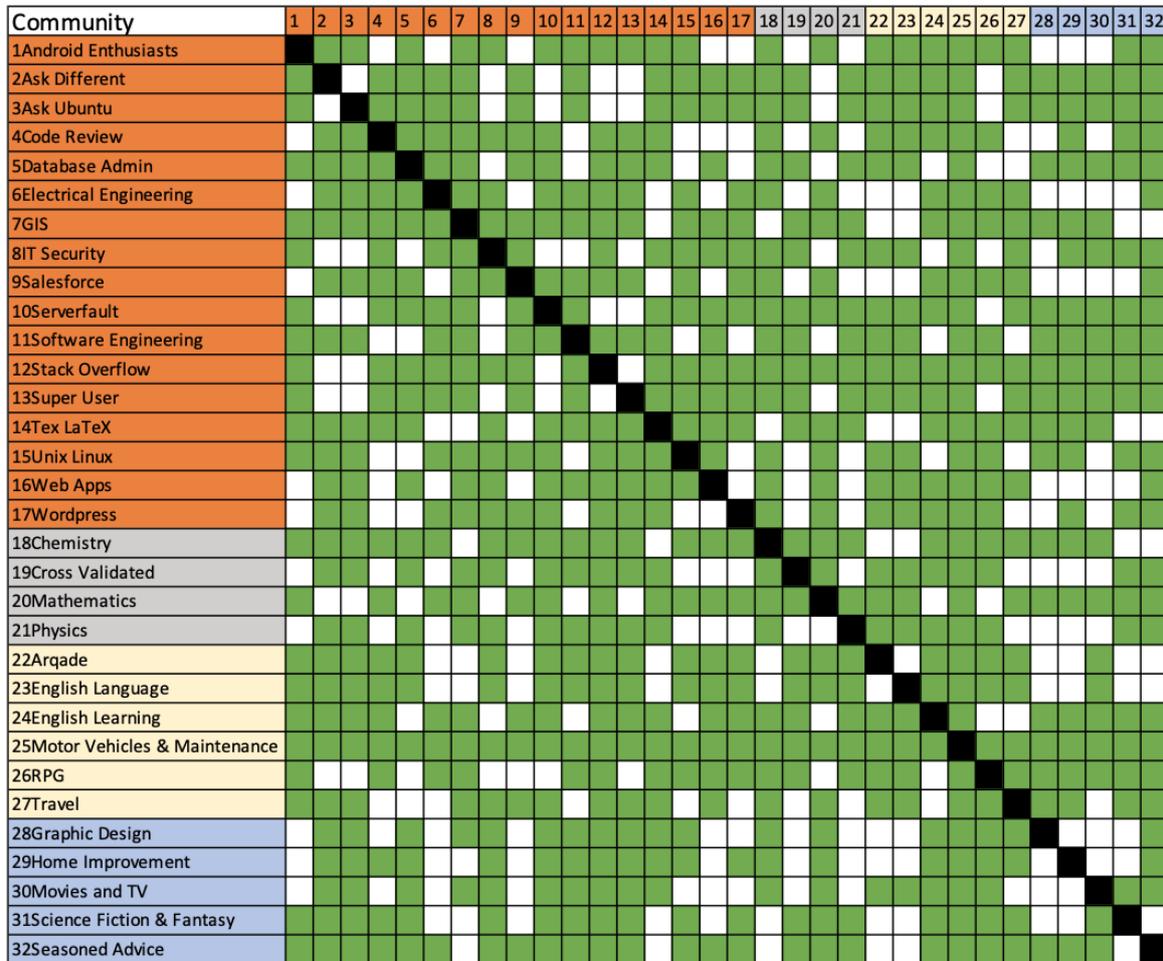


Figure 1. Correlation matrix for all communities. Communities are organized by category. orange—technology, grey—science, yellow—culture/recreation, blue—life/arts. Green indicates that there is a significant difference at $p < 0.05$ between the communities, whereas white indicates no significant difference.

5.3. User Ability to Earn Reputation Based on Tenure

Based on the results of the correlation analysis, we next evaluate whether a poverty trap effect could be seen in the user base for each community. Due to the power law [7], sampling from the entire community includes users who are part of either end of the distribution tail. We expect that the top users would earn the most reputation from their answers. Therefore, in order to understand how tenure impacts a contributor who is not a power user, we sampled from the set of users who contributed between 10 and 20 answers and had a positive aggregate answer score. Full details of each community’s sample size can be found in Table 4. We did not control for when answers were given, as it is impossible

to control for the quality of answering opportunities. This likely diminishes the impact of tenure and reputation earning for this analysis.

Table 4. Stack Exchange community information.

Site	Sample Size	Average Answer Score	Standard Deviation
Android Enthusiasts	374	1.68	1.63
Arqade (Video games)	1109	3.39	2.52
Ask Different	816	2.69	3.08
Ask Ubuntu	2966	3.25	5.71
Chemistry	205	3.35	2.86
Code Review	542	3.07	1.51
Cross-Validated (Statistics)	952	3.08	3.69
Database Administrators	488	2.32	2.98
Electrical Engineering	836	2.11	1.60
English Language and Usage	1340	3.09	2.81
English Language Learners	610	2.40	2.29
Geographic Information Systems (GIS)	822	2.20	1.45
Graphic Design	250	2.54	2.31
Home Improvement	316	2.35	1.89
Information Security	631	4.06	3.97
Mathematics	4791	1.79	1.95
Motor Vehicle Maintenance and Repair	169	2.41	1.83
Movies and TV	169	8.28	5.24
Physics	1304	2.33	2.34
Role-Playing Games	543	7.26	4.95
Salesforce	696	1.39	1.36
Science Fiction and Fantasy	596	9.05	6.73
Seasoned Advice (Cooking)	308	3.73	2.67
Server Fault	2846	2.14	2.57
Software Engineering	947	5.17	5.31
Stack Overflow	48,373	2.6	6.81
Super User	4056	2.98	4.39
TeX–LaTeX	607	4.72	5.93
Travel	273	6.27	4.29
Unix and Linux	1472	4.13	6.43
Web Applications	223	2.88	2.27
WordPress Development	888	1.35	1.28

Table 4 shows the details of the data sets for the analysis. Because reputation is not normally distributed for any of the communities, we ran a generalized linear model (GLM) with a Gamma distribution and log-link function for each of the communities. For each community, the dependent variable is the average answer score, and the independent variable is the creation date of the community. The results are shown in Table 5. Pseudo R^2 was calculated with McFadden's R^2 . We use this value as a rough estimation of how much the user performance can be explained by tenure, rather than other factors, such as skill.

Table 5. GLM results for the user's average answer score and the number of months of membership. Ordered by pseudo R^2 .

Site	Estimate	Error	R^2	t-Value
TeX–LaTeX	0.0160	0.0010	0.384	16.0 ***
Cross-Validated (Statistics)	0.0139	0.0008	0.319	17.2 ***
Motor Vehicle Maintenance and Repair	0.0156	0.0018	0.296	8.7 ***
Database Administrators	0.0135	0.0014	0.264	9.9 ***
Graphic Design	0.0130	0.0016	0.257	8.0 ***
Unix and Linux	0.0171	0.0011	0.257	16.0 ***
Android Enthusiasts	0.0144	0.0018	0.222	8.0 ***
Mathematics	0.0091	0.0004	0.205	25.7 ***
Chemistry	0.0099	0.0016	0.204	6.2 ***
Web Applications	0.0117	0.0015	0.202	7.8 ***
Electrical Engineering	0.0087	0.0007	0.192	12.5 ***
WordPress Development	0.0115	0.0009	0.180	12.6 ***
Salesforce	0.0120	0.0011	0.166	10.5 ***
Stack Overflow	0.0132	0.0003	0.153	51.3 ***
Arqade (Video games)	0.0085	0.0006	0.150	13.6 ***
Super User	0.0128	0.0006	0.149	19.9 ***
Ask Ubuntu	0.0136	0.0009	0.140	15.2 ***
English Language Learners	0.0106	0.0012	0.140	8.6 ***
Physics	0.0086	0.0007	0.139	11.6 ***
English Language and Usage	0.0092	0.0008	0.105	11.4 ***
Geographic Information Systems (GIS)	0.0061	0.0007	0.105	9.0 ***
Server Fault	0.0085	0.0006	0.100	14.2 ***
Ask Different	0.0080	0.0012	0.086	6.6 ***
Software Engineering	0.0089	0.0012	0.080	7.6 ***
Science Fiction and Fantasy	0.0065	0.0009	0.078	7.0 ***
Code Review	0.0038	0.0006	0.077	6.2 ***
Information Security	0.0077	0.0014	0.064	5.5 ***
Home Improvement	0.0047	0.0012	0.058	3.8 ***
Movies and TV	0.0059	0.0019	0.058	3.2 ***
Seasoned Advice (Cooking)	0.0035	0.0011	0.036	3.1 ***
Travel	0.0043	0.0017	0.026	2.5 *
Role-Playing Games	−0.0005	0.0009	0.001	−0.6

* $p < 0.05$, *** $p < 0.001$.

The results, as seen in Table 5, show that most communities have tenure as a significant predictor of the average answer score. One community, *Role-Playing Games*, does not have tenure as a predicting factor. Overall, this would indicate that there is a bias towards earning reputation points earlier in the life cycle of almost all communities. In some cases, the results show a strong bias towards tenure, such as in the cases of *TeX–LaTeX* and *Cross-Validated*. In general, it would appear that certain communities that cover stable technical and scientific fields have a stronger bias towards tenure than other communities. On the other hand, communities that cover topics such as art and leisure tend to have lower

relationships with tenure and reputation earning. Some technological communities are also included in the latter group.

It is also interesting to note that there is no consistent relationship between the question values, as found in Tables 3 and 5. For instance, *Seasoned Advice* has a strong relationship between the age of the question and upvotes, but does not have a strong relationship with reputation earning and tenure. Meanwhile, *Stack Overflow* shows the lowest correlation with the age of the question, but finishes in the middle of the community results when considering answer score and tenure.

6. Discussion

There appear to be real changes facing Q&A communities regarding the ability of users to earn new reputation points from newer material. That is, Q&A sessions from the past are essentially worth more than sessions today. There is an initial suspicion that the number of questions in each sub-domain is decreasing while new areas are opening up. It is becoming harder for users to aggregate reputation points across the whole community. The other problem is that any dry period of activity, no matter the reason, is more likely to have a detrimental effect on newer users. They will stagnate as their tenured peers continue to gain reputation.

Some communities do not exhibit a strong bias. This may be due to the nature of the communities themselves, where the question asking and answering allows for more redundancy or nuanced questions, or where there is an opportunity from knowledge-creation cycles (for instance, a software platform that periodically makes large, critical changes).

Let us consider Jain's hypothesis about the questioner's motivation as a basis for analyzing the data set [30]. Jain supposes that two types of questions seek factual answers: (1) questions in which the fastest satisfactory answer is the most desirable, and (2) questions in which cumulative answers increase the value for the questioner.

An example of the former is the question, "Where do I find directions for the conference submission?" Once a satisfactory answer has been proffered (for instance, a hyperlink to the directions), additional answers have less value for the questioner. The latter question would be, "What is the most efficient programming language for Data Science". In this case, successive answers can complement and add value to the questioner. The current setup for Stack Exchange may be better suited for the latter type of questioning if the system's goal is to have lower barriers to entry.

The results of this study show that there is some bias towards tenure in almost all communities, but that it is not consistent. Therefore, it would be ideal if communities have reputation systems that are tailored to their specific information creation needs.

7. Conclusions

This study's results can help inform both Q&A sites and other knowledge-based peer-production communities focused on information sharing and aggregation. We recommend the following:

- Community administrators should understand the scarcity of valuable Q&A interactions as a community ages. A progressive community should seek to include inflationary measures that allow new users to overcome the advantages had by more tenured users. In addition, administrators could relax policies toward duplicated questions, as they are often asked by novice users and may actually be a source of new information [34].
- Community administrators should reward users for improving existing information by allowing them to share in the reputation revenue. Instead of merely offering small reputation rewards for improving existing material, a progressive scheme would allow for reputation sharing to improve information interaction.
- Community administrators should tailor their reputation scheme based on the domain in which the community exists and modify the reputation system based on the severity of the poverty trap.

Finally, it might seem that chatbots may end the need for peer production and reputation systems. While chatbots will certainly make information querying simpler, it may have trouble with information creation, from both question and answering perspectives. That is, a bot may be very good in certain information spaces, but not in spaces where knowledge creation is uncertain or rapidly changing. In this climate, it is especially important that peer-production systems increase equity within their reputation systems and increasingly reward incremental knowledge creation as a community ages.

Online knowledge-based peer-production communities, like collaborative Q&A communities, are susceptible to bias towards early Q&A contributions. This means that these systems can effectively have “poverty traps”, which prevent users from competing on equal footing with their peers. Identifying which types of domains suffer from these barriers and which do not is an important step in building more equitable communities. We find that there is a consistent bias that may be exacerbated by the type of information needed and the creation cycle established within each community. The overall implication of this study is that the ability to add and include new users to these types of knowledge-based peer-production communities is threatened by the existence of poverty traps.

Limitations

A limitation of this study is the absence of interviews or surveys of the users and non-users of these communities. There is much to gain by exploring and understanding how system rules impact different users. It is also important to understand whether the shape of the community acts as a firm barrier to entry for prospective non-users. Future studies should include these findings.

Author Contributions: Conceptualization, A.V., B.T. and K.K.; methodology, A.V. and B.T.; formal analysis, A.V. and C.B.; writing—original draft preparation, A.V. and B.T.; writing—review and editing, A.V., B.T., C.B. and K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by a grant from the Japan Science and Technology Agency (JST) (grant no. JPMJCR20G3).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data generated for this study were obtained from Stack Exchange’s Data Explorer. The following queries can be used and edited for reproduction and further exploration. Question posts and votes obtained: <https://data.stackexchange.com/stackoverflow/query/1724457> (accessed on 23 March 2023) Reputation earned by users: <https://data.stackexchange.com/stackoverflow/query/1720915> (accessed on 16 March 2023).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Q&A Question and Answer
GLM Generalized Linear Model

References

1. Anderson, A.; Huttenlocher, D.; Kleinberg, J.; Leskovec, J. Discovering Value from Community Activity on Focused Question Answering Sites: A Case Study of Stack Overflow. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, 12–16 August 2012; ACM: New York, NY, USA, 2012; pp. 850–858. [[CrossRef](#)]
2. Tausczik, Y.R.; Kittur, A.; Kraut, R.E. Collaborative Problem Solving: A Study of MathOverflow. In Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, Baltimore, MD, USA, 15–19 February 2014; ACM: New York, NY, USA, 2014; pp. 355–367. [[CrossRef](#)]

3. Tausczik, Y.R.; Pennebaker, J.W. Participation in an Online Mathematics Community: Differentiating Motivations to Add. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, Seattle, WA, USA, 11–15 February 2012; ACM: New York, NY, USA, 2012; pp. 207–216. [[CrossRef](#)]
4. Li, G.; Zhu, H.; Lu, T.; Ding, X.; Gu, N. Is It Good to Be Like Wikipedia? Exploring the Trade-offs of Introducing Collaborative Editing Model to Q&A Sites. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Vancouver, BC, Canada, 14–18 March 2015; ACM: New York, NY, USA, 2015; pp. 1080–1091. [[CrossRef](#)]
5. Mamykina, L.; Manoim, B.; Mittal, M.; Hripcsak, G.; Hartmann, B. Design Lessons from the Fastest Q&a Site in the West. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; ACM: New York, NY, USA, 2011; pp. 2857–2866. [[CrossRef](#)]
6. MacLeod, L. Reputation on Stack Exchange: Tag, You're It! In Proceedings of the 2014 28th International Conference on Advanced Information Networking and Applications Workshops, Victoria, BC, Canada, 13–16 May 2014; pp. 670–674. [[CrossRef](#)]
7. Adamic, L.A.; Zhang, J.; Bakshy, E.; Ackerman, M.S. Knowledge Sharing and Yahoo Answers: Everyone Knows Something. In Proceedings of the 17th International Conference on World Wide Web, Beijing, China, 21–25 April 2008; ACM: New York, NY, USA, 2008; pp. 665–674. [[CrossRef](#)]
8. Furtado, A.; Andrade, N.; Oliveira, N.; Brasileiro, F. Contributor Profiles, Their Dynamics, and Their Importance in Five Q&A Sites. In Proceedings of the 2013 Conference on Computer Supported Cooperative Work, San Antonio, TX, USA, 23–27 February 2013; ACM: New York, NY, USA, 2013; pp. 1237–1252. [[CrossRef](#)]
9. Hardin, C.D.; Berland, M. Learning to Program Using Online Forums: A Comparison of Links Posted on Reddit and Stack Overflow (Abstract Only). In Proceedings of the 47th ACM Technical Symposium on Computing Science Education, Memphis, TN, USA, 2–5 March 2016; ACM: New York, NY, USA, 2016; p. 723. [[CrossRef](#)]
10. Ford, H. Online Reputation: It's Contextual, 2012. Available online: <http://ethnographymatters.net/blog/2012/02/24/online-reputation-its-contextual/> (accessed on 27 April 2023).
11. Treude, C.; Barzilay, O.; Storey, M.A. How Do Programmers Ask and Answer Questions on the Web? (NIER Track). In Proceedings of the 33rd International Conference on Software Engineering, Honolulu, HI, USA, 21–28 May 2011; ACM: New York, NY, USA, 2011; pp. 804–807. [[CrossRef](#)]
12. Vargo, A.W.; Matsubara, S. Editing Unfit Questions in Q&A. In Proceedings of the 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), Kumamoto, Japan, 10–14 July 2016; pp. 107–112. [[CrossRef](#)]
13. Livan, G.; Pappalardo, G.; Mantegna, R.N. Quantifying the relationship between specialisation and reputation in an online platform. *Sci. Rep.* **2022**, *12*, 16699. [[CrossRef](#)] [[PubMed](#)]
14. Adler, B.T.; de Alfaro, L. A Content-driven Reputation System for the Wikipedia. In Proceedings of the 16th International Conference on World Wide Web, Banff, AB, Canada, 8–12 May 2007; ACM: New York, NY, USA, 2007; pp. 261–270. [[CrossRef](#)]
15. De Alfaro, L.; Kulshreshtha, A.; Pye, I.; Adler, B.T. Reputation Systems for Open Collaboration. *Commun. ACM* **2011**, *54*, 81–87. [[CrossRef](#)] [[PubMed](#)]
16. Halfaker, A.; Geiger, R.S.; Morgan, J.T.; Riedl, J. The Rise and Decline of an Open Collaboration System How Wikipedia's Reaction to Popularity Is Causing Its Decline. *Am. Behav. Sci.* **2013**, *57*, 664–688. [[CrossRef](#)]
17. Wei, X.; Chen, W.; Zhu, K. Motivating User Contributions in Online Knowledge Communities: Virtual Rewards and Reputation. In Proceedings of the 2015 48th Hawaii International Conference on System Sciences, Kauai, HI, USA, 5–8 January 2015; pp. 3760–3769. . [[CrossRef](#)]
18. Wang, S.; German, D.M.; Chen, T.H.; Tian, Y.; Hassan, A.E. Is reputation on Stack Overflow always a good indicator for users' expertise? No! In Proceedings of the 2021 IEEE International Conference on Software Maintenance and Evolution (ICSME), Luxembourg, 27 September–1 October 2021; pp. 614–618. . [[CrossRef](#)]
19. Chen, Y.; Ho, T.H.; Kim, Y.M. Knowledge Market Design: A Field Experiment at Google Answers. *J. Public Econ. Theory* **2010**, *12*, 641–664. [[CrossRef](#)]
20. Jeon, G.Y.; Kim, Y.M.; Chen, Y. Re-examining Price As a Predictor of Answer Quality in an Online Q&A Site. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; ACM: New York, NY, USA, 2010; pp. 325–328. [[CrossRef](#)]
21. What Is Reputation? How Do I Earn (and Lose) It?—Help Center—Stack Overflow. 2015. Available online: <http://stackoverflow.com/help/whats-reputation> (accessed on 25 June 2023).
22. Vargo, A.W.; Matsubara, S. Corrective or critical? Commenting on bad questions in Q&A. In Proceedings of the iConference 2016, iSchools, Philadelphia, PA, USA, 20–23 March 2016. [[CrossRef](#)]
23. Answer to “What Are the Reputation Requirements for Privileges on Sites, and How Do They Differ per Site?”. 2023. Available online: <https://meta.stackexchange.com/a/160292> (accessed on 25 June 2023).
24. Help Center—What types of questions should I avoid asking?—Stack Overflow, 2019. Available online: <https://stackoverflow.com/help/dont-ask> (accessed on 20 September 2019).
25. Nam, K.K.; Ackerman, M.S.; Adamic, L.A. Questions in, Knowledge in?: A Study of Naver's Question Answering Community. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4–9 April 2009; ACM: New York, NY, USA, 2009; pp. 779–788. [[CrossRef](#)]
26. Kang, M. Motivational affordances and survival of new askers on social Q&A sites: The case of Stack Exchange network. *J. Assoc. Inf. Sci. Technol.* **2022**, *73*, 90–103. [[CrossRef](#)]

27. Adaji, I.; Vassileva, J. Towards Understanding User Participation in Stack Overflow Using Profile Data. In Proceedings of the Social Informatics, Bellevue, WA, USA, 11–14 November 2016; Springer: Cham, Switzerland, 2016; pp. 3–13. [[CrossRef](#)]
28. Ginsca, A.L.; Popescu, A. User Profiling for Answer Quality Assessment in Q&A Communities. In Proceedings of the 2013 Workshop on Data-Driven User Behavioral Modelling and Mining from Social Media, San Francisco, CA, USA, 28 October 2013; ACM: New York, NY, USA, 2013; pp. 25–28. [[CrossRef](#)]
29. Vargo, A.W.; Matsubara, S. Identity and performance in technical Q&A. *Behav. Inf. Technol.* **2018**, *37*, 658–674. [[CrossRef](#)]
30. Jain, S.; Chen, Y.; Parkes, D.C. Designing Incentives for Online Question and Answer Forums. In Proceedings of the 10th ACM Conference on Electronic Commerce, Stanford, CA, USA, 6–10 July 2009; ACM: New York, NY, USA, 2009; pp. 129–138. [[CrossRef](#)]
31. Solomon, J.; Wash, R. Critical Mass of What? Exploring Community Growth in WikiProjects. In Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, Ann Arbor, MI, USA, 1–4 June 2014.
32. All Sites—Stack Exchange. 2023. Available online: <https://stackexchange.com/sites?view=list#traffic> (accessed on 28 March 2023).
33. Diedenhofen, B.; Musch, J. cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS ONE* **2015**, *10*, e0121945. [[CrossRef](#)] [[PubMed](#)]
34. Abric, D.; Clark, O.E.; Caminiti, M.; Gallaba, K.; McIntosh, S. Can Duplicate Questions on Stack Overflow Benefit the Software Development Community? In Proceedings of the 2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR), Montreal, QC, Canada, 25–31 May 2019; pp. 230–234. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.