Two-Sided Long Tails on Blockchain-Based Crowdsourcing Platforms

Bingyi Wu
Univ. of Texas at San Antonio
bingyi.wu@utsa.edu

Charles Zhechao Liu
Univ. of Texas at San Antonio
charles.liu@utsa.edu

Kim-Kwang Raymond Choo
Univ. of Texas at San Antonio
raymond.choo@utsa.edu

Hongyi Zhu
Univ. of Texas at San Antonio
hongyi.zhu@utsa.edu

Shen Chang
Univ. of Texas at San Antonio
shen.chang@utsa.edu

Abstract

Motivated by the recent adoption of a blockchain-based token-weighted incentive mechanism in online crowdsourcing markets, which has sparked concerns over power and popularity concentration, this study investigates the interactions between the distributions on the two sides (voters and contributors) of a blockchain-based crowdsourcing platform (BCP). Built upon the theories of long tails and two-sided markets, we propose a two-sided long-tail framework to uncover the distributional dynamics on BCPs. Our weekly panel data analyses show that the concentration in token holdings distribution among voters negatively affects content growth, which in turn leads to a higher content popularity concentration. Such an adverse effect is more pronounced in niche content growth compared to popular content growth. The results also show that the distribution of voters' token holdings positively affects content diversity, which in turn results in a more evenly distributed content popularity. The theoretical contributions and practical implications of our findings are discussed.

Keywords: Long Tail, Two-Sided Markets, Online Crowdsourcing, Tokenomics, Blockchain Application

1. Introduction

The term "Long Tail," coined by former chief editor Chris Anderson in Wired magazine[1], describes how niche products gain significant market share, diverting market demand away from popular products and creating a relatively flatter demand distribution among product suppliers in online markets compared to brick-and-mortar markets (Brynjolfsson et al., 2006). While this long-tail demand distribution encourages suppliers to offer wider product selections that cater to diverse consumer tastes, it also leads to an overwhelming number of products competing for consumers' attention (Elberse, 2008).

The potential of having a long-tail distribution in demand or content popularity is crucial in shaping contributors' decisions to participate in monetized online crowdsourcing platforms (e.g., YouTube, MTurk) that are fueled by crowdsourced content from various contributors, who are monetarily rewarded based on their contribution performance (e.g., content popularity). In this context, contributors often compare their rewards with those of others, making their contribution decisions based on relative achievements and perceived fairness (Li et al., 2021; Liu & Feng, 2021; Tang et al., 2012), which are shaped by the long-tail content popularity distribution in the market. Yet, the "winner-takes-all" dynamic prevails on monetized online crowdsourcing platforms, as platform owners tend to subsidize top performers to exploit network effects (Sockin & Xiong, 2022). As a result, contributions from a few well-known contributors dominate consumer endorsement and platform rewards, limiting growth opportunities for lesser-known contributors and discouraging their contributions (Jin, 2020). This has fueled a demand for more democratic crowdsourcing mechanisms, with blockchain-based crowdsourcing platforms (BCPs) emerging as a potential solution (Barrera, 2018).In fact, expenditure on blockchain-based solutions by businesses worldwide has been undergoing exponential growth in recent years and is projected to reach 19 billion U.S. dollars in 2024[2].

BCPs differ from other centralized crowdsourcing platforms in that not only do they provide digital "shelf spaces" for crowdsourced contributions, but they also replace the centralized revenue-sharing scheme with a *token-weighted crowdsourcing (TWC) mechanism* (Tsoukalas & Falk, 2020) to incentivize effortful contributions and efficient information aggregation from contributors and evaluators (e.g., crowd voters) (Goldin, 2017). More precisely, the TWC mechanism consists of *voting* and *reward* mechanisms. The former aggregates crowd voters' preferences over

[1] https://www.wired.com/2004/10/tail/

[2] https://www.statista.com/statistics/800426/worldwide-blockchain-solutions-spending/

URI: https://hdl.handle.net/10125/106883
978-0-9981331-7-1

various content contributions submitted to the platform, where votes from users with more tokens are weighted more. The latter proportionally distributes token rewards to each contribution such that the contributor receives a portion of the rewards, and the rest is distributed among the voters based on their token holdings. Crowd voters who hope to enhance their reward power can obtain tokens through either direct purchase or their reward-eligible contributions, such as blog posts or comments.

With this seemingly democratic TWC mechanism, contributors are likely to join BCPs with expectations of less market concentration and a longer tail in content popularity distribution (Catalini & Gans, 2020). However, content popularity concentration persists on BCPs (Li & Palanisamy, 2019; Li et al., 2022), which could hinder the decentralized democracy of BCPs (Beck et al., 2018). For instance, Li et al. (2022) found a highly skewed distribution of token rewards on a BCP, resulting in a decline in users' knowledge contribution. Hence, it is crucial to investigate whether the TWC mechanism may shorten or contribute to the long tail in content popularity distribution on BCPs.

Existing literature on long tails in online markets has suggested that the long tail in demand distribution is often driven by technological drivers (Park et al., 2020), demand-side search behaviors (Brynjolfsson et al., 2011), and supply-side production incentives (Tan et al., 2017). Nevertheless, the *intertwined relationships* among these drivers have not yet been explicitly studied. Moreover, online crowdsourcing markets are two-sided, with contributors and consumers on the two sides, and thus subject to cross-side network effects, which arise when consumers (suppliers) derive values from each additional supplier (consumer) joining the market (Rochet & Tirole, 2006). Yet, no existing long-tail studies have considered how the *two-sidedness* of online crowdsourcing markets, including BCPs, may alter the contributors' contribution incentives, which could, in turn, influence content popularity distribution in the market. Hence, our study aims to fill these research gaps by addressing the following research questions:

RQ1: How does the distribution of crowd voters' token holdings influence content contribution outcomes such as content diversity and content growth on a two-sided BCP?
RQ2: How do the content contribution outcomes influence the content popularity distribution on a two-sided BCP?

We review the related literature in the next section, followed by our theory and hypotheses development. We then describe our data and present the empirical methods and hypotheses testing results. Lastly, we conclude our findings and discuss the theoretical contributions and practical implications.

2. Literature Review

2.1. Tokenomics in BCPs

Prior studies on BCPs have pointed to the economic paradoxes induced by the TWC mechanism of BCPs (Ciriello et al., 2018).

On the one hand, blockchain-enabled TWC mechanisms can provide ample opportunities for content monetization and promote user engagement and knowledge contributions on BCPs (Zheng & Boh, 2021). For instance, the number of token holdings and token rewards is shown to influence user contribution in a blockchain-based content market positively (Liu et al., 2022). A positive association has also been found between token investment and users' participatory efforts (Zhang et al., 2019), suggesting that the investment behaviors motivated via the TWC mechanism may reinforce user commitment to BCPs.

On the other hand, such TWC mechanisms may dilute users' intrinsic participation motivation and reinforce wealth and power concentration (Li et al., 2022), threatening the democracy of BCPs (Beck et al., 2018). For example, receiving higher rewards for content contribution and voting are shown to crowd out users' motivation to craft good-quality content (Zhang et al., 2019). Prior literature has also suggested that the TWC mechanism on BCPs may encourage more strategic behaviors among users (Tsoukalas & Falk, 2020) and diminish users' prosocial participatory motivations. Li et al. (2022) found a highly-skewed token incentive distribution on a BCP, which consequently deteriorates users' knowledge contribution. Similarly, Li and Palanisamy (2019) showed that under the TWC mechanism, the governance power of current BCPs is concentrated at the top users, hindering decentralization.

No consensus has yet been reached in terms of how the current TWC mechanism benefits or hinders the development of BCPs. The introduction of the TWC mechanism turns BCPs into competitive marketplaces, highlighting the salient role of monetary reward and the relative comparison among contributors in driving knowledge contribution. Also, under the TWC mechanism, the contributors' reward is realized through voters' votes and the token weights they carry. Hence, the distribution of voters' wealth is likely to influence contributors' knowledge contribution decisions and outcomes, such as content growth and diversity, which may, in turn, impact the distribution of content popularity. Our study extends the current literature by investigating such distributional interactions induced by the TWC mechanism on BCPs.

2.2. Monetized Content Contribution

An extensive literature on content contribution in online crowdsourcing markets has primarily focused on voluntary contribution in public goods games where free-riding problems prevail. In this context, social incentives, such as peer feedback and support (Qiu & Kumar, 2017), reputation and recognition (Bhattacharyya et al., 2020), reciprocal norms (Chen & Hung, 2010), and sense of trust and belonging (Ren et al., 2012), have been shown to play critical roles in combating free-riding problems and driving voluntary content contribution.

Nonetheless, monetization of content contribution in online crowdsourcing markets is on the rise (Jin, 2020). For example, various content-based platforms, such as YouTube and Twitch, offer content contributors monetary rewards and incentives to appreciate the value their contributions bring to the platforms and to encourage their continuous contributions. Despite the prevalence of content monetization, relatively few studies have delved into knowledge contribution in the presence of monetary incentives (Liu & Feng, 2021; Tang et al., 2012). The context of monetized contribution is different from voluntary contribution in *two* important ways. First, even though social incentives still play important roles in motivating early-stage content contribution, those social capitals seem to matter to well-established contributors only when they can be translated into financial payoff (Tang et al., 2012). For example, Tang et al. (2012) suggest that while contributors contribute for exposure and reputation in addition to the revenue-sharing incentive on YouTube, they focus on attracting more subscribers instead of viewers as they climb up the status hierarchy.

Second, the presence of monetary incentives has turned content-based markets into competitive marketplaces rather than collaborative networks (e.g., IQ.Wiki[3] versus Wikipedia). Competition is a crucial factor influencing content contribution across monetized crowdsourcing markets since the audience becomes a scarce resource for which contributors must compete (Bhattacharyya et al., 2020). In a competitive setting, seeing other contributors' performance may induce social comparison, which could either motivate users to contribute more with greater effort to achieve a superior status position (Li et al., 2021) or drive them away when competition is perceived as unfair (Liu & Feng, 2021). For instance, using a game theoretical approach, Liu and Feng (2021) found that introducing monetary rewards in online communities exhibits competition crowding-out effects. Low-effectiveness contributors reduce their contributions due to the intensified competition induced by monetary rewards, whereas high-effectiveness contributors contribute even more. Moreover, a recent study of knowledge contribution in a blockchain-based knowledge community revealed that the misalignment between token incentives allocation and knowledge contribution efforts has an adverse impact on content contribution (Li et al., 2022). These findings suggest that contributors' contribution decisions in a monetized content community like BCPs are often shaped by their relative achievements, which are influenced by competition and the distribution of monetary incentives.

Our study extends this line of research by considering how the distribution of crowd voters' token holdings may influence the competitive landscape of a BCP, which could ultimately affect their contribution outcomes (i.e., content diversity and growth) and content popularity distribution.

3. Theories and Hypotheses Development

In this section, we first introduce our two-sided long-tail research framework (see Figure 1) for investigating the distributional interactions induced by the TWC mechanism on BCPs. Then, we present theories and hypotheses development.

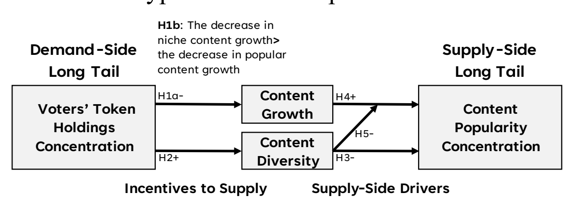

Figure 1. Two-sided long tail framework

3.1. Theories of Long Tails

Prior literature on long tails in online markets has suggested that the long tail in demand distribution is driven by technological drivers, demand-side search behaviors, and supply-side production incentives (Brynjolfsson et al., 2006, 2010; Hinz et al., 2011).

Several studies have delved into the *technological* drivers and explored the role of recommender systems. They found that passive recommender systems based on historical demand and popularity of the products or services (e.g., favorite lists) reduce the tail length in aggregate sales distribution on e-commerce platforms,

[3] IQ.Wiki, formerly known as Everipedia, is a blockchain-based online encyclopedia that crowdsources knowledge contribution from various contributors and offer them crypto incentives for their contributions.

whereas personalized recommenders and active search tools (e.g., search filters) facilitate the long tail phenomenon (Hinz et al., 2011; Li et al., 2021).

Prior studies have also explored the *supply-side* drivers, including product pricing strategies (Oh et al., 2016) and the size of product assortments (Hinz et al., 2011; Tan et al., 2017), that sit behind the long-tail distribution in online markets. Specifically, scholars have shown that implementing a paywall on the content goods (e.g., news content) that were previously free could drive the long tail in the distribution of word-of-mouth for niche and popular content on social media. This is because the originally heavy users who consume mixed niche and popular content are likely to continue their consumption with a fee, while light users who prefer popular content are likely to reduce or even discontinue their news consumption (Oh et al., 2016). Furthermore, studies on the impacts of assortment size on the long tail in demand distribution have produced mixed findings. While Hinz et al. (2011) have found that the increasing assortment size stretches the tail length in the sales distribution of a video-on-demand operator, Tan et al. (2017) have shown that the expansion of the number of distinct movie DVDs diversifies more demand away from niche movie titles than from popular titles, thus increasing demand concentration. These mixed findings call for further investigation of the impacts of product characteristics on the long tail phenomenon.

Neither technological nor supply-side drivers can be isolated from the *demand-side* drivers that ultimately shape demand distribution. It is not the mere shift in technologies that alters the product demand; rather, it is the changes in consumers' search behaviors and preferences, induced by technological advancements, that reshape the demand distribution (Brynjolfsson et al., 2011; Hinz et al., 2011; Park et al., 2020). Likewise, the demand distribution is not solely driven by the changes in product assortment and price but rather by the interplay among consumers' evolving tastes (Brynjolfsson et al., 2006), the composition of consumer segments (e.g., heavy and light users), their responses to the shifts in product supplies (Hinz et al., 2011; Oh et al., 2016) and their collective effort that shifts the demand. Moreover, the strategic changes on the supply side are also likely to be the results of suppliers' responses to the observed market demand. Therefore, accounting for the intertwined relationships among technological, supply-side, and demand-side drivers is necessary when investigating the long tail in online markets.

3.2. Demand-Side Search Behaviors on BCPs

In online markets, including BCPs, consumers' search behaviors and product preferences can be influenced by their engagement or investment in that market (Hinz et al., 2011; Oh et al., 2016). This can be explained from multiple theoretical perspectives. For example, according to equity theory (Huseman et al., 1987), individuals assess the ratio of their inputs and outcomes and work to restore equity between the two. Hence, the more time or money consumers invest in a market, the more they would want to get out of it. They may expend their search efforts to discover better deals or unique products, thereby providing advantages to the niche product segments in the market. Alternatively, cognitive dissonance theory (Festinger, 1957) suggests that when individuals' attitudes or beliefs clash with their behaviors, they undergo a state of dissonance, which motivates them to modify their attitudes or beliefs in order to alleviate the discomfort. Thus, as consumers invest more time or money in a market, they would adapt to believe the market is worth their efforts. This adapted belief would, in turn, motivate them to actively engage and intensify their search efforts, increasing their likelihood of discovering and consuming niche products in the market. The exposure theory (McPhee, 1963) further supports the notion that heavy consumers are more inclined to consume niche products than light consumers. This is because their intensive engagement and exploration in a market expose them to many alternatives, while light consumers who are only familiar with limited alternatives are more likely to stick with the blockbusters (Oh et al., 2016).

On a BCP powered by the TWC mechanism, "powerful" crowd voters who possess substantial token holdings are more likely to be heavy consumers on the platform who obtain tokens through direct purchase and active engagement in incentive-eligible activities (i.e., voting, posting, and commenting). , As a result, they are not only exposed to more content products (Elberse, 2008; McPhee, 1963) but also are more likely to appreciate niche and fresh topics rather than mainstream and redundant information, compared to other powerless voters. This is consistent with McPhee's exposure theory. Moreover, powerful voters are also the token holders who not only stand to benefit more from an increase in the market value of the platform but also bear the loss from a decrease (Tsoukalas & Falk, 2020). Consequently, they are more likely to appreciate unique and valuable content that helps establish a competitive advantage of the platform. In contrast, "powerless" crowd voters with few token holdings are more likely to be light consumers on the platform who stick to the popular content (see Figure 2). In addition, we also plot the Kolmogorov-Smirnov (KS) test results in Figure 3 to support our arguments.

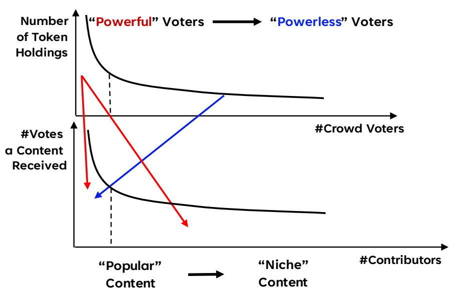

Figure 2. The inverse relationship between content popularity and crowd voters' token holdings on BCPs (adapted from Oh et al. 2016)

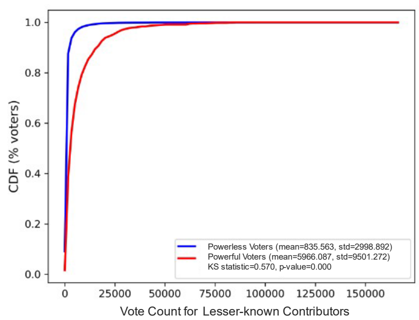

Figure 3. Kolmogorov-Smirnov (KS) test results

Notes: The Y axis shows the cumulative probability of voters, and the X axis shows the vote count for content contributed by lesser-known contributors on the HIVE Blog. The KS test result suggests that the powerful voters cast, on average, significantly more votes for lesser-known content contributors than the rest of the voters.

3.3. Two-Sidedness and Contribution Outcomes on BCPs

Online crowdsourcing markets are *two-sided markets* characterized by *cross-side network effects (CNEs)*, which arise when consumers (suppliers) derive values from each additional supplier (consumer) joining the market (Rochet & Tirole, 2006). In our research context, CNEs from voters to contributors arise when the contribution values to contributors increase with each additional voter casting votes on a BCP. Conversely, CNEs from voters to contributors diminish when the contribution values to contributors decrease with each additional voter leaving the platform. Moreover, the differential token holdings and thus reward power of crowd voters suggest that the different segments of voters (i.e., powerful and powerless voters) may carry distinct values to contributors. Crowd voters with more token holdings bring more contribution value and, therefore, greater CNEs to contributors than those with fewer tokens. Hence, the change in the distribution of crowd voter segments will likely alter the degrees of CNEs that arise from their participation, which is important in incentivizing contributors' contributions.

When holding the total number of crowd voters constant, shifting towards a more concentrated distribution of voters' token holdings will create an imbalanced voter base with a smaller group of powerful voters and a larger group of relatively powerless voters (see Figure 4). As contributors' reward incomes largely rely on the votes from powerful voters, the reduction in the number of powerful voters would diminish the contribution values of contributors more than the value the increased number of powerless voters would bring to them. Therefore, a more concentrated distribution of crowd voters' token holdings would reduce the overall degree of CNEs, leading to a slower growth of content contribution on a BCP.

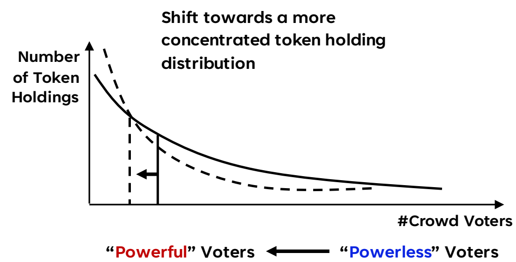

Figure 4. Distributional shift in crowd voter base

Notes: Solid line shows the original distribution; dash line shows the distribution after the shift.

The CNEs can often be highly *local* (Sundararajan, 2007). A subset of consumers (suppliers) may enter the market due to the increase in a subset of suppliers (consumers) whom they interact mostly with. For example, some consumers may join a crowdsourcing platform just because it has a growing number of offerings from a specific content category that the consumers are interested in. More contributors may join to contribute more niche content as more consumers consume them on the platform.

As suggested previously, powerful voters are more likely to consume a mix of niche and popular content, while powerless voters are more likely to stick to popular content. Therefore, the greater concentration in crowd voters' token holdings due to a reduced number of powerful voters and an increased number of powerless voters would reduce the community value for contributors who produce niche content to a greater extent than for contributors who produce popular content. With these logics, we hypothesize that:

H1: *The greater the concentration in the distribution of crowd voters' token holdings, the slower the growth of content contribution on a BCP (H1a). This negative effect is more salient for niche content than for popular content (H1b).*

Moreover, uniqueness theories predict that individuals' needs to establish distinct identities to shape their self-identity and manage their impressions on others through their content contributions (Snyder

& Fromkin, 2012). The desire for uniqueness may become more pronounced in a competitive marketplace, where idiosyncratic and non-imitable ideas or strategies are crucial for sustaining individuals' competitive advantages (Ulrich & Lake, 1991). On a competitive BCP, the greater token-holding concentration among crowd voters may induce intensified competition among contributors to compete for the limited attention from the less powerful voters. Thus, contributors may deliberately contribute unique and distinctive content in order to stand out from the crowd. In doing so, they signal crowd voters their true efforts and qualities, which attracts votes from powerful voters whose incentives are tied to the community wellbeing. Thus, we hypothesize that:

H2: *The greater the concentration in the distribution of crowd voters' token holdings, the more diverse the content on a BCP.*

3.4. The Impacts of Contribution Outcomes

The prior literature establishes the relationships between the size of product assortments and the long tail distribution in demand based on two *seemingly* contradictory theoretical perspectives. On the one hand, the classical variety-seeking theories suggest that providing a greater assortment size (i.e., more products) helps consumers to find the product that meets their diverse and changing needs (Huffman & Kahn, 1998; Lancaster, 1990). Variety-seeking behaviors are more commonly observed for experiential goods (e.g., books) compared to search goods (e.g., cameras), as the former tend to have more subjective qualities that vary across consumers (Park et al., 2020; Tan et al., 2017). This line of theory would predict that offering more experiential products contributes to the long tail phenomenon in online markets (Hinz et al., 2011), where consumers would benefit from having more product choices that are previously unknown.

On the other hand, too many choices can also exhaust the product selection (Gourville & Soman, 2005; Huffman & Kahn, 1998). This is because making choices from a large size of assortments induces intensive evaluations that require more cognitive capacity, which is limited for each consumer in a given timeframe (Huffman & Kahn, 1998). As evaluating experiential goods whose qualities may not be fully revealed upfront often demands more effort (Tan et al., 2017), having more choices of these goods would only complicate the decision-making process further. Hence, to avoid exhaustive product selection, consumers may prefer to stick to popular experiential products when facing excessive product choices. Under this line of theory, the greater assortment size would inhibit the long tail phenomenon in online markets (Tan et al., 2017), contrasting the previous prediction.

We attribute the conflicting predictions to their underlying assumptions regarding the diversity of product assortments. The line of variety-seeking theories assumes that a larger product assortment size brings greater diversity in the assortment, thereby satisfying consumers' heterogeneous tastes for experiential goods (Hinz et al., 2011). The second line of theory assumes that as product assortment size gets larger, the differences among the alternatives become smaller (Iyengar & Lepper, 2000), making them substitutable by each other. These underlying assumptions predict that, even with excessive choices for experiential goods such as the content on BCPs, having more diverse selections may help diversify consumers' preferences across the selections, leading to a more evenly distributed demand in the market. This argument is also supported by the notion of "hyperdifferentiation" (Clemons et al., 2006), which suggests that providing highly differentiated products helps firms grow sales. Hence, we hypothesize that:

H3: *Increasing content diversity reduces concentration in content popularity distribution on a BCP.*

While the size of product assortments alone may not directly impact the demand distribution in a market, it is important to consider the effect of continuous expansion of product selections at increasing rates. As the number of available choices expands *rapidly*, consumers' cognitive resources become strained, depleting their decision-making capacity. This cognitive overload can result in consumers resorting to simpler heuristic cues when evaluating and selecting products (Tan et al., 2017). For example, on a content-based BCP, voters may look for peripheral quality signals, such as the number of likes or upvotes, to economize on their cognitive costs during their content evaluations (Brynjolfsson et al., 2010). Hence, we hypothesize that:

H4: *Increasing content growth rate concentrates content popularity distribution on a BCP.*

Lastly, we also propose that if content growth is accompanied by an increase in content diversity, the undesirable impact of content growth on content popularity concentration can be alleviated. This is because as the content growth rate is distributed across a broader range of content selections, it leads to a more digestible growth pattern for each group of content. We hypothesize that:

H5: *Content diversity mitigates the positive impact of content growth rate on content popularity concentration on a BCP.*

4. Research Setting and Data

To test our hypotheses, we traced and collected data for 42 weeks (01/01/2022 to 10/25/2022) from HIVE Blog, a leading BCP launched in March 2020 with over 548,000 active users. HIVE Blog crowdsources various content with topics ranging from photography, cryptocurrency, and art to gaming and daily life. It also crowdsources content evaluations through voting from its users and is powered by the TWC mechanism to incentivize effortful content contribution and voting behaviors. The data documents users' daily activities along with users' weekly account attributes for 184 active content communities on the platform. A community is considered active when it has at least one post each week. We excluded voting records where the voter and the contributor of the content share the same user identity. The remaining dataset contains over 68 million votes from 37,666 crowd voters for 831,930 posts contributed by 23,238 contributors. We construct a weekly panel for our empirical analyses.

We used the number of upvotes a content received each week to measure the content popularity. We then followed Brynjolfsson et al. (2011) and Tan et al. (2017) to calculate the Gini coefficients of content popularity across all the content in each week for each community as the measures of the weekly content popularity concentration (ContPopGini). The Gini coefficient is a relative measure that ranges from 0 to 1, with 0 indicating complete equality in distribution and 1 for complete inequality. By calculating the Gini coefficient, we hold the long-tail measure scale invariant. Similarly, we calculated the Gini coefficients of voters' token holdings across all the crowd voters in each week for each community to measure the weekly voters' token holding concentration (VTHGini). To measure the weekly content diversity (ContDvst) for each community, we first tokenized and lemmatized all the content titles and computed the Jaccard similarity scores between each pair of content titles in each community in a week. We then calculated diversity using the average Jaccard similarity score of all pairs of content titles, subtracted from one, in each community in that week. To measure the weekly content growth (ContGrow), we calculated the percentage changes in the content post counts between two adjacent weeks for each community each week. Additionally, we calculated the weekly content growth for niche (NicheGrow) and popular content (PopGrow) each week for each community. We categorized content as a niche when the number of upvotes is less than the average number of upvotes and popular when it is more than the average. The average value is calculated after removing outliers in the samples. We also performed log transformations for growth and diversity variables to alleviate within-level and between-level skewness.

Control Variables. We control for various factors that may influence the distribution of content popularity, content growth, and content diversity. This includes the average content length, the average content payout, the average contributor and voter reputation score, the share of dual-role users who contributed and voted, the number of reblogs, the average number of comments, the average number of downvotes, and the token price in each week. We show summary statistics in Table 1.

Table 1. Summary Statistics
(T = 42 Weeks, N = 184 Active Communities, Obs = 7728)

Variable	Mean	Std. dev.	Min	Max
ContPopGini	0.469	0.156	0.000	0.865
ContGrow	0.150	0.822	-0.962	21.500
PopGrow	0.163	0.954	-1.000	18.000
NicheGrow	0.181	1.004	-1.000	29.000
ContDvst	0.926	0.129	0.000	1.000
VTHGini	0.922	0.038	0.167	0.984
Controls				
AvgPostLen	4567.812	3568.741	110.000	47888.100
AvgPayOut	4.635	5.076	0.000	99.701
AvgCRP	63.216	5.018	26.820	81.410
AvgVRP	57.273	2.885	34.023	65.130
DualRole	0.273	0.064	0.034	1.000
BSpan	5.302	3.710	0.000	59.167
TokenPrice	0.740	0.290	0.382	1.545
AvgCom	7.016	7.686	0.000	234.667
AvgDVs	1.390	1.455	0.000	46.000
Reblogs	151.893	286.084	0.000	5454.000

5. Model Specifications and Results

5.1. Model Specifications

Given that our dependent variables are likely to be influenced by their past values, we adopt the dynamic Arellano-Bover/Blundell-Bond panel-data model with the two-step generalized method of moments (GMM) estimators (Arellano & Bover, 1995; Blundell & Bond, 1998), which is designed for short panel datasets (N>T). The method takes the first differences of variables to remove the panel-level effects and uses lagged levels as instruments for the first-differencing equation, which makes the model highly robust to heteroscedasticity and autocorrelation. Moreover, the method does not require strict assumptions of exogeneity, which is ideal for data collected from two-sided markets in which the two market sides influence each other through CNEs.

Before specifying our models, we first conduct a series of Hausman tests to determine whether fixed-effect specifications are appropriate. The test results reject the null hypothesis of non-systematically different coefficients and confirm the appropriateness of the fixed-effect models. Then, we conduct a series of unit root tests using the Levin-Lin-Chu approach, and the results confirm the stationarity of all the variables involved in our models. We treat our independent variables as predetermined variables without strict exogeneity, and we include up to three lags of dependent variables for use as instruments. We specify our models as follows:

$$\Delta lnY_{it} = \alpha_1 \Delta lnY_{i,t-1} + \beta_1 \Delta X_{i,t-1} + \sum_{j=1}^{n} \beta_j \, \Delta Control_{in,t-1} + \epsilon_{i,t} + C$$

where lnY_{it} is our logged dependent variables of community i in week t, which are the functions of their own past values $lnY_{i,t-1}$ and the values of independent variables $X_{i,t-1}$ and control variables $Control_{in,t-1}$ in week $t-1$ for community i. $\epsilon_{i,t}$ is the error term assumed to be serially uncorrelated, and C is the constant. We conduct Arellano-Bond autocorrelation tests and Sargan overidentification tests for our models. The test results confirm that the models are appropriately specified without autocorrelation and overidentification concerns. We present our model results in the next section.

Table 1. Content Growth and Diversity Model Results

DVs		lnContGrow$_t$		lnPopGrow$_t$		lnNicheGrow$_t$		lnContDvst$_t$	
		Coeff.	Std. err.	Coeff.	Std. err.	Coeff.	Std. err.	Coeff.	Std. err.
lnContGrow$_{t-1}$		-0.299***	0.001	/	/	/	/	/	/
lnPopGrow$_{t-1}$		/	/	-0.205***	0.000	/	/	/	/
lnNicheGrow$_{t-1}$		/	/	/	/	-0.219***	0.000	/	/
lnContDvst$_{t-1}$		/	/	/	/	/	/	-0.002***	0.000
VTHGini$_{t-1}$		**-3.352*****	0.056	**-1.725*****	0.050	**-3.223*****	0.057	**0.749*****	0.001
Constant		-1.007***	0.047	-0.943***	0.046	-2.185***	0.049	-1.402***	0.001
Controls		Included		Included		Included		Included	
Wald chi2		3.20E+05 (0.000)		2.72e+06 (0.000)		2.03e+06 (0.000)		1.30E+08 (0.000)	
Arellano-Bond	**Order1**	z = -9.950 (0.000)		z = -9.950 (0.000)		z = -6.337 (0.000)		z = -1.521 (0.028)	
	Order2	z = 0.268 (0.789)		z = 0.268 (0.229)		z = 0.048 (0.962)		z = 1.032 (0.302)	
Sargan Tests		Chi2 = 183.808 (0.573)		Chi2 = 183.617 (0.577)		Chi2 = 183.867 (0.572)		Chi2 = 183.363 (0.679)	

Notes: N = 184, T = 40, Obs. = 7360. ***p < 0.001; **p < 0.01; *p < 0.05. Bolded results are hypothesized in our study.

5.2. Model Results

Content Growth and Diversity Model Results. Table 1 shows the results of the effects of the voters' token-holding concentration on content growth and content diversity on a BCP. The results show that the effect of VTHGini on the logged content growth is negative and significant, suggesting that, in general, the increase in the crowd voters' token-holding concentration hinders the content growth (H1a is supported). Consistent with the overall effect of VTHGini, we also observe significant and negative effects of VTHGini on both the logged niche content growth and the logged popular content growth. The negative effect on the logged popular content growth is less than that on the logged niche content growth (H1b is supported). Moreover, consistent with our hypothesis 2, we find a significant and positive relationship between VTHGini in week $t-1$ and the logged content diversity in week t, suggesting that a more concentrated distribution of voters' token holdings encourages more diverse content production on a BCP.

Content Popularity Concentration Model Results. We show the results of the effects of contribution outcomes (i.e., content growth and content diversity) on content popularity concentration on a BCP in Table 2. The results show that the content growth in week $t-1$ has a significant and positive impact on content popularity concentration in week t, while the content diversity in week $t-1$ has a significant and negative impact on content popularity concentration in week t. The results suggest that faster content growth leads to a more concentrated popularity distribution among content, while more diverse content selections lead to a less concentrated content popularity distribution. Hence, both our hypotheses 3 and 4 are supported.

Moreover, we find a significant and negative moderation effect of content diversity on the relationship between content growth and content popularity concentration, implying that increasing content diversity can alleviate the negative impact of content growth on content popularity concentration (H5 is supported). Overall, our model results show strong support for our hypotheses. We discuss the theoretical and practical implications in the following section.

Table 2. Content Popularity Concentration Model Results

DV	ContPopGini$_t$	
	Coeff.	Std. err.
ContPopGini$_{t-1}$	0.066***	0.003
ContGrow$_{t-1}$	**0.417*****	0.044
ContDvst$_{t-1}$	**-0.016***	0.007
ContGrow*ContDvst$_{t-1}$	**-0.628*****	0.046

	Constant	0.469***	0.010
	Controls	Included	
	Wald chi2	8973.770 (0.000)	
Arellano-Bond	**Order1**	z = -7.189 (0.000)	
	Order2	z = 1.827 (0.068)	
Sargan Tests		Chi2 = 171.504 (0.173)	

6. Conclusion and Discussion

Accompanied by the recent adoption of a token-weighted incentive mechanism powered by blockchain technology in online crowdsourcing markets, there are increasing concerns over the power and popularity concentration that could compromise the decentralized democracy of the blockchain (Beck et al., 2018). Motivated as such, we investigate the interactions between the distributions on the two market sides (crowd voters and content contributors) through the two-sided long-tail framework. We find concentration in crowd voters' token holdings to have a negative effect on content growth, which leads to a higher content popularity concentration. Such a negative effect is greater for niche content growth than for popular content growth. Interestingly, we also find the distribution of voters' token holdings to have a positive effect on content diversity, which results in a more evenly distributed content popularity. Moreover, we find that the presence of content diversity can alleviate the adverse effect of content growth on content popularity concentration.

The contributions of our study are two-fold. First, building on theories of long tails and two-sided markets, we propose a novel research framework that investigates the two-sided long tails. The extensive literature on long tails suggests that the long tail in supply-side distribution is driven by technological drivers, demand-side search behaviors, and supply-side production incentives. We extend these theoretical perspectives by integrating the intertwined relationships among the three drivers. Specifically, we shed light on how the demand-side wealth distribution induced by the blockchain-based TWC mechanism alters supply-side contribution incentives through two-sided market dynamics, consequentially influencing the supply-side popularity distribution. Second, our study contributes to the literature on monetized online content contribution by identifying the effects of the distribution of crowd evaluators' reward power (in our case, the crowd voters' token holdings) on knowledge contributors' contribution incentives, including their incentives to contribute and to craft unique content.

Our study also yields important practical implications for developers of BCPs. Specifically, our findings suggest that if the primary goal of a BCP is to incentivize more content contributions, having a lower concentration level of voters' token holdings is desirable. This can be done, for example, by introducing a redistribution mechanism that allows users to benefit from redistributing their token holdings to others. However, if a platform's goal is to build a repository for novel and unique content contributions, then having a high concentration in voters' token holdings could be helpful as it induces competition among contributors and their need to establish unique identities through their contributions.

Nevertheless, the long-term impacts of these strategies deserve further investigation in future research. Moreover, some voters are likely to be more experienced and influential than others. Future studies may consider such multi-dimensional heterogeneity in voters, in terms of their voting power, investment experience, and influence when examining the interplay of the two-sided long tails.

7. References

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29–51.

Barrera, C. (2018, March 29). The Blockchain Effect: Network Effects without Market Power Costs. MIT Cryptoeconomics Lab.

Beck, R., Müller-Bloch, C., & King, J. L. (2018). Governance in the Blockchain Economy: A Framework and Research Agenda. Journal of the Association of Information Systems, 19(10), 1020–1034.

Bhattacharyya, S., Banerjee, S., Bose, I., & Kankanhalli, A. (2020). Temporal Effects of Repeated Recognition and Lack of Recognition on Online Community Contributions. Journal of Management Information Systems, 37(2), 536–562.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1), 115–143.

Brynjolfsson, E., Hu, Y. (Jeffrey), & Simester, D. (2011). Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. Management Science, 57(8), 1373–1386.

Brynjolfsson, E., Hu, Y. (Jeffrey), & Smith, M. D. (2010). Research Commentary— Long Tails vs. Superstars: The Effect of Information Technology on Product Variety and Sales Concentration Patterns. Information Systems Research, 21(4), 736–747.

Brynjolfsson, E., Yu, Hu, & Smith, M. D. (2006). From Niches to Riches: Anatomy of the Long Tail. MIT Sloan Management Review.

Catalini, C., & Gans, J. S. (2020). Some simple economics of the blockchain. Communications of the ACM, 63(7), 80–90.

Chen, C.-J., & Hung, S.-W. (2010). To give or to receive? Factors influencing members’ knowledge sharing and community promotion in professional virtual communities. Information & Management, 47(4), 226–236.

Ciriello, R. F., Beck, R., & Thatcher, J. B. (2018). The Paradoxical Effects of Blockchain Technology on Social Networking Practices. 39th International Conference on Information Systems, 19.
Clemons, E. K., Gao, G. G., & Hitt, L. M. (2006). When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry. Journal of Management Information Systems, 23(2), 149–171.
Elberse, A. (2008). Should You Invest in the Long Tail? Harvard Business Review, 11.
Festinger, L. (1957). A Theory of Cognitive Dissonance. Stanford University Press.
Goldin, M. (2017, September 14). Token-Curated Registries 1.0. Medium.
Gourville, J. T., & Soman, D. (2005). Overchoice and Assortment Type: When and Why Variety Backfires. Marketing Science, 24(3), 382–395.
Hinz, O., Eckert, J., & Skiera, B. (2011). Drivers of the Long Tail Phenomenon: An Empirical Analysis. Journal of Management Information Systems, 27(4), 43–70.
Huffman, C., & Kahn, B. E. (1998). Variety for sale: Mass customization or mass confusion? Journal of Retailing, 74(4), 491–513.
Huseman, R. C., Hatfield, J. D., & Miles, E. W. (1987). A New Perspective on Equity Theory: The Equity Sensitivity Construct. Academy of Management Review, 12(2), 222–234.
Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? Journal of Personality and Social Psychology, 79, 995–1006.
Jin, L. (2020, December 17). The Creator Economy Needs a Middle Class. Harvard Business Review.
Lancaster, K. (1990). The Economics of Product Variety: A Survey. Marketing Science, 9(3), 189–206.
Li, C., & Palanisamy, B. (2019). Incentivized Blockchain-based Social Media Platforms: A Case Study of Steemit. Proceedings of the 10th ACM Conference on Web Science - WebSci '19, 145–154.
Li, X., Grahl, J., & Hinz, O. (2021). How Do Recommender Systems Lead to Consumer Purchases? A Causal Mediation Analysis of a Field Experiment. Information Systems Research. 33(2), 620-637.
Li, Z., Wang, G., & Wang, H. J. (2021). Peer Effects in Competitive Environments: Field Experiments on Information Provision and Interventions. MIS Quarterly, 45(1a), 163–191.
Li, Z., Xie, Y., & Xu, X. (2022). An Analysis of the Influence of Token Incentive Allocation Monopoly on User Knowledge Contribution in Blockchain-Based Knowledge Communities. Journal of Systems Science and Mathematical Sciences, 42(6), 1362.
Liu, Y., & Feng, J. (2021). Does Money Talk? The Impact of Monetary Incentives on User-Generated Content Contributions. Information Systems Research, 32(2), 394–409.
Liu, Z., Li, Y., Min, Q., & Chang, M. (2022). User Incentive Mechanism in Blockchain-based Online Community: An Empirical Study of Steemit. Information and Management, 103596.
McPhee, W. N. (1963). Formal theories of mass behavior. Free Press of Glencoe.
Oh, H., Animesh, A., & Pinsonneault, A. (2016). Free Versus For-a-Fee: The Impact of a Paywall on the Pattern and Effectiveness of Word-of-Mouth via Social Media. MIS Quarterly, 40(1), 31–56.
Park, Y., Bang, Y., & Ahn, J.-H. (2020). How Does the Mobile Channel Reshape the Sales Distribution in E-Commerce? Information Systems Research, 31(4), 1164–1182.
Qiu, L., & Kumar, S. (2017). Understanding Voluntary Knowledge Provision and Content Contribution Through a Social-Media-Based Prediction Market: A Field Experiment. Information Systems Research, 28(3), 529–546.
Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut, R. E. (2012). Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds. MIS Quarterly, 36(3), 841–864. JSTOR.
Rochet, J.-C., & Tirole, J. (2006). Two-sided markets: A progress report. The RAND Journal of Economics, 37(3), 645–667.
Snyder, C. R., & Fromkin, H. L. (2012). Uniqueness: The Human Pursuit of Difference. Springer Science & Business Media.
Sockin, M., & Xiong, W. (2022). Decentralization Through Tokenization. NBER, 73.
Sundararajan, A. (2007). Local Network Effects and Complex Network Structure. The BE Journal of Theoretical Economics, 7(1), 46.
Tan, T. F., Netessine, S., & Hitt, L. (2017). Is Tom Cruise Threatened? An Empirical Study of the Impact of Product Variety on Demand Concentration. Information Systems Research, 28(3), 643–660.
Tang, Q., Gu, B., & Whinston, A. B. (2012). Content Contribution for Revenue Sharing and Reputation in Social Media: A Dynamic Structural Model. Journal of Management Information Systems, 29(2), 41–76.
Tsoukalas, G., & Falk, B. H. (2020). Token-Weighted Crowdsourcing. Management Science, 66(9), 3843–3859.
Ulrich, D., & Lake, D. (1991). Organizational capability: Creating competitive advantage. Academy of Management Perspectives, 5(1), 77–92.
Zhang, R., Park, J., & Ciriello, R. F. (2019). The Differential Effects of Cryptocurrency Incentives in Blockchain Social Networks. SIGBPS 2019 Workshop on Blockchain and Smart Contract.
Zheng, Y., & Boh, W. F. (2021). Value drivers of blockchain technology: A case study of blockchain-enabled online community. Telematics and Informatics, 58, 101563.