Understanding Sequential Decision-Making by Platform Participants: A Structural Analyses of Crowdsourcing and Cryptocurrencies

Vipul Aggarwal

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Yong Tan, Chair

Ming Fan

Hyeunjung (Elina) Hwang

Program Authorized to Offer Degree:
Foster School of Business

University of Washington

Abstract

Understanding Sequential Decision-Making by Platform Participants: A Structural Analyses of Crowdsourcing and Cryptocurrencies

Vipul Aggarwal

Chair of the Supervisory Committee:
Michael G. Foster Endowed Professor of Information Systems Yong Tan
Information Systems and Operations Management

In this dissertation, I study the decision-making process of participants on online platforms. I specifically focus my attention on the sequential aspects of such processes, i.e. my interest is in understanding the endogenous determinants of the decision-making processes in the contexts crowdsourcing and cryptocurrencies. First, I study the user participation in an open-innovation platform, specializing in solving socio-economic problems via crowdsourcing ideas and solutions. I hypothesize that platform participants are creatives who, with the right mix of knowledge and application, can find solutions to these problems requiring inter-disciplinary skills. I model their participation decisions as dependent on their entire history of knowledge acquisitions and applications i.e. every focal interaction, with the platform, is dependent on all the past interactions and all the future interactions are affected by the focal interactions. This allows me to uncover the nuanced usage of existing knowledge by platform participants in generating creative ideas that has implications for managers of knowledge-based crowdsourcing platforms. We also highlight the importance of feedback mechanisms that adds to platform participants' knowledge over time. Further, we also find a significant impact of learning-by-doing in improving users' *creative* ideation abilities. In my second study, I study the cryptocurrency mining industry, specifically focusing on the competitive equilibria in Bitcoin and Bitcoin-Cash around the time of Bitcoin fork (which led to the birth of Bitcoin-Cash). I also investigate the impact of emergency difficulty adjustment algorithm on the Bitcoin-Cash's equilibria. I model these mining pools (or miners) as forward-looking profit-maximizing firms whose strategic interactions lead to dynamic

changes in the underlying system's protocols. Our results indicate that competition among Bitcoin miners has a positive externality, improving the expected payoffs for every miner. Further, our investigation reveals the dynamic impact of this externality for the Bitcoin-Cash ecosystem owing to emergency difficulty adjustment algorithm. We also estimate the varying impact of the within-miner competition, from BTC mining, on the BCH mining rate. Our counterfactual simulations provide further insights into possible configurations for the emergency difficulty adjustment algorithm.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

BTC: Bitcoin

BCH: Bitcoin-Cash

EDA: Emergency Difficulty Adjustment

DAA: Difficulty Adjustment Algorithm

ACKNOWLEDGMENTS

I would like to take this opportunity to thank those individuals and organizations whose constant support and guidance has been very valuable in my doctoral studies. First and foremost, I would like to express my sincerest gratitude to my advisor, Prof. Yong Tan, for nurturing the scholar in me. His mentorship has empowered me to be an independent researcher, his very probing questions encouraged me to think deeply, his attention to detail motivated me to think about potential mistakes, and his off-the-cuff suggestions helped me explore and navigate the numerous and technically challenging methodologies. I am very thankful to him for his uninterrupted support and trust in me.

I would also like to express my sincere thanks to Prof. Hyeunjung (Elina) Hwang, whose assiduous and meticulous attitude constantly inspired me to be a better researcher and to focus on continuous improvements. I would also like to thank Prof. Ming Fan, Prof. Foad Iravani, and Prof. Manish Chalana for their insightful questions/comments and encouraging scholarly learning. I really appreciate their time and support.

I would be remiss to not express my deepest appreciation for the support provided by the essential employees of the Foster School of Business. My research life was thoroughly made better by the administrative and technical assistance provided by Shawna Reimers, Jaime Banaag, Beau Kirkeby, Jeffrey Balaz, and Brent. I would like like to thank CSDE Computing and its director, Matt Weatherford, for providing high-quality computing infrastructure to doctoral students.

I would also like to thank my friends and peers in UW doctoral programs. I was enriched by their knowledge and inspired by their dedication.

Finally, I want to acknowledge the constant support and unconditional love provided by my parents, parents-in-law, and brother. They inspire me to be a better and a thoughtful human being. My wife, Arunima, has made my life unimaginably better. Her love and emotional support has carried me through thick and

thin. I am eternally thankful to my family members for their love and care.

DEDICATION

to serendipitous opportunities,

&

to thousands of scholars on whose shoulders my work stands

Chapter 1

INTRODUCTION

The ability (or the inability) to correctly learn from our prior decisions, to guide future decisions, has shaped the course of human history. In an almost online world, billions of users make trillions of decisions (or choices) on digital platforms which facilitate their many needs, such as social connectivity (social media), buying and selling goods and services (e-commerce), sourcing knowledge and talents (crowdsourcing), etc. These decisions, however small or large, impact the users' future interactions with the platforms dynamically. Since a major part of our life now is spent navigating and interacting with these relatively young systems, it becomes imperative to analyze the sequences of choices and their dynamic inter-dependencies. Every interaction, such as an online click, comment submission, payment processing, etc. creates a digital trail or trajectory of rich user information. These trajectories can be utilized to understand the users' underlying dynamic decision-making processes.

Every user decision is made in a certain context and therefore, its analysis requires us to investigate it within the bounds of that context. While some research questions permit us to assume that these decisions (or interactions) do not affect the underlying context or that they are independent of the past, conditional on the present state of the context, some settings require us to understand and model these inter-dependent interactions with the environment or drop Markovian assumptions whenever necessary. I motivate my research from the vantage point of the econometrician: (a) abandon Markovian dynamics and assume that the entire past history of the focal user has an effect on his current choices, and (b) assume that the focal user is a forward-looking rational agent, i.e., his current choices are affected by his beliefs on the evolution of the context and his future decisions therein.

In Chapter 4, I analyze the decision made by users of an open innovation platform regarding common participation activities such as acquiring knowledge from their peers and submitting their ideas for challenges. Open innovation platforms connect knowledge-seeking firms with knowledge-providing users (also

referred to as crowdsourcing). Their success depends on finding the right set of creative and knowledgeable users and nurturing them to provide high-quality solutions. I focus on one of the key determinants of creativity, analogical learning, on the ability of these platforms to increase their creative throughput. Analogical learning is the phenomenon of using the knowledge acquired from one domain to solve problems of other domains. This phenomenon plays an even bigger role when dealing with inter-disciplinary problems, such as socio-economic challenges like eradicating poverty, reducing waste, etc. Our research context is an open innovation platform dealing with socio-economic challenges that require inter-disciplinary knowledge (and hence, the ability to connect knowledge from one domain to another). A user's current participation is affected, not only by his past activities, but also by the content of the said activities. In our context, this content represents knowledge about solving different types of problems (also referred to as analogies or stimulus ideas). We model the focal user's participation decisions through an endogenous framework of Mutually Exciting Point Processes and explicitly impose knowledge-based structure through the dependency of idea submissions on the interaction of acquired knowledge (from other ideas on the platform) with that of the required knowledge (for a particular challenge).

By utilizing the activity stream data of 13,028 participants from 2010 to 2016 from this open innovation platform, I uncover the synergistic effects of these participation activities on creative outcomes. I find that knowledge acquired from others' ideas (stimulus ideas) plays a vital role in the creative ideation process, but their effect is more nuanced than what we have known so far. In contrast to the prior belief that distant analogies, stimulus ideas outside of a problem domain, spur creativity, we find that distant analogies lead to failures. Yet, we further find that such failures are indispensable to the creative ideation process because failures motivate idea generators (1) to acquire more knowledge by increasing their future knowledge acquisition from other participants (learning-from-others), and (2) to persist in generating ideas that lead to improvements in their ability to apply the acquired knowledge and to identify innovation tasks that are relevant to their stock of acquired knowledge (learning-by-doing). Our results indicate that failures are a stronger driver of the learning activities than successes.

In Chapter 3, I empirically analyze the equilibrium behavior of miners involved in mining Bitcoin (BTC) and Bitcoin-Cash (BCH) cryptocurrencies during the period of April 2017 - February 2018. These miners

are responsible for processing transactions and creating new coins, and compete with each other to be the first one to do so. Hence, their strategic interactions with the underlying platform affect the commonly shared platform protocols, which in turn, affect their future decisions. Thus, we hypothesize that miners are forward-looking profit-maximizing agents whose aim is to maximize the current expected discounted profits from future mining (rather than profits from just the current mining operation). The shared evolutionary structure of the BTC (or BCH) protocols is affected by the collective mining decisions made by these miners. Our research interest lies in ascertaining the impact of competition on their expected payoffs.

BCH was born out of a contentious BTC fork in August 2017 and inherited BTC's characteristics such as its mining and difficulty adjustment algorithms. BCH developers introduced Emergency Difficulty Adjustment (EDA) algorithm to incentivize miners to process BCH's transactions over that of BTC's. Based on the strategic actions of the miners, difficulty changes brought about by EDA resulted in dynamic changes to the difficulty adjusted profitability of BCH. This manifested in wild swings in their mining rates. Consequently, it leads to two research questions: (a) effectiveness of such emergency mechanisms, and (b) impact of competition with EDA-modulated difficulty changes. Therefore, we construct an equilibrium model of cryptocurrency mining to uncover miners' profit motivations, with and without such mechanisms. We utilize a two-step structural approach to recover the model primitives.

Our results indicate the importance of competition to BTC (and BCH) mining. Competition for BTC blocks resulted in lower overall computational cost for everyone involved and therefore, competition positively impacted miners' per-period utilities. However, wild changes in difficulty levels can lead to conditions where competition has an erosive effect on payoffs. We find that during EDA-Active periods, BCH miners were impacted negatively by competition due to EDA-triggered difficulty reductions. This result is further corroborated by our theoretical model of mining in equilibrium, which derives the optimal level of competitors. Moreover, we also find strong fraternal competition between BTC and BCH miners for the shared computational hash power. As expected, competitive effect was strongest during EDA-active period as EDA temporarily made mining BCH more lucrative. Therefore, any diversion of hash power, from BCH to BTC, resulted in a much stronger negative effect on miners' payoff.

Chapter 2

LITERATURE REVIEW

My research encompasses two different research contexts: open innovation and cryptocurrency platforms. Since I study decision-making under different conditions, I cover the concerned conceptual literature on creativity and idea generation in Section 2.1 and cryptocurrency mining in Section 2.2.

2.1 Open Innovation Platform

Creativity research has a long history spanning diverse disciplines such as innovation (Von Hippel, 1988), psychology (Amabile, 1988), marketing (Toubia and Netzer, 2017), information systems (Huang et al., 2014; Hwang et al., 2019), and management science (Girotra et al., 2010). Across the diverse disciplines, scholars have identified several factors that lead to creative idea generation.

First, knowledge is an important component of creative idea generation. Creativity is defined as the production of novel and useful ideas that are built upon existing knowledge (Amabile, 1988; Toubia and Netzer, 2017; B. Ward et al., 1995). That is, creative ideas are not spawned in a vacuum but produced through the process of recombining knowledge in new ways (e.g. Schumpeter, 1934; Fleming, 2001). This definition of creativity indicates that the development of creative ideas is reliant upon access to and retrieval of relevant knowledge (Toubia and Netzer, 2017; A. Finke et al., 1992). In gastronomy, the ability of a chef to fit the right set of ingredients to the final taste (s)he desires is crucial. Similarly, generating a creative idea for a particular problem requires the access and retrieval of the right set of existing knowledge. In creativity parlance, this process of retrieving and applying knowledge to a new problem context is referred to as analogical reasoning (Dunbar, 1995; Gentner et al., 1997). Acknowledging the importance of analogical reasoning in creativity, Weisberg (1998) propositioned that the result of creative task is governed by an individual's ability to acquire knowledge and mold it into a novel idea that solves the given problem.

Scholars studying user innovation have also emphasized the importance of knowledge in creativity. In

particular, they state that users acquire relevant knowledge through interactions with other users. By sharing and providing feedback to each other's ideas, users get an opportunity to collect important needs information that can later be used to build their ideas (e.g. Von Hippel, 1988; Morrison et al., 2000). Especially when users are embedded in a community, the community serves as a reservoir of rich and relevant knowledge that is often non-replicable outside (Shah and Tripsas, 2007). As a result, several studies have documented evidence that innovations arise more frequently in industries where user communities are prevalent (e.g., computing, automobiles, and juvenile products industry) (Allen, 1983; Franke and Shah, 2003; Lüthje et al., 2005; Shah and Tripsas, 2007)

Likewise, user interaction is a critical feature in a community-based open innovation platform. In crowdsourcing platforms, users interact by sharing their ideas and reading & commenting on others' ideas. Comments may involve the sharing of pertinent information, critique of a specific aspect of an idea, suggestion of a different approach, questions, and compliments. Previous studies have shown that through the process of providing feedback, users collect new information represented in ideas and also recall relevant information from their memory (Shah and Tripsas, 2007; Hwang et al., 2019). In other words, user interactions enrich the creative intelligence of individuals by allowing them to discover and accumulate knowledge (Singh et al., 2011; Ransbotham et al., 2012). While the knowledge acquired from a community is a subset of individuals' entire stock of knowledge, this subset is likely to be activated and used as a source of new idea generation: Prior research has found that people have difficulty activating relevant information from their memory unless the information was recently accessed in a similar context (Higgins, 1996).

To illustrate, we provide a few anecdotal evidence that suggest that users acquire knowledge through commenting activities. Although it is difficult to measure how much knowledge users acquire and how much relevant information they recall in the process of leaving comments, we observe that users put significant efforts in understanding the proposed ideas and formulating comments in our research context. The time and effort users put in understanding and providing feedback suggest that individuals are more likely to retain the knowledge in their memory and use it when they engage in creative tasks later (Higgins, 1996). Some comments do suggest that learning is happening. For example, we see that several comments explicitly mention that the idea an individual is commenting on "inspired" their own idea (e.g., "Your idea's emphasis

on using photographs for rewarding and reminding families about getting their child immunized is very insightful! Your idea inspired mine, Immunization Shield."). In addition, although not explicitly mentioning about direct impact on their own ideas, many comments reveal that individuals acquired knowledge from the idea they are commenting on (e.g., "I think this is a really interesting idea. It tackles a problem from a unique point of view, focusing specifically on making the whole experience a little more positive and more memorable. ... I am excited to see where this goes!", "This is a great way of simplifying an issue which has been overcomplicated..."). As the example indicates, the knowledge they acquire may involve facts but it may also involve the way an ideator approaches a problem.

Due to the importance of existing knowledge in creative ideation, researchers have paid specific attention to the characteristics of existing knowledge that influence the creative process. Researchers found that exposure to diverse knowledge aids creative idea generation. For instance, in an innovation experiment, Dahl and Moreau (2002) found that participants who are provided with multiple analogies were more likely to create original ideas than participants who are given a single analogy. In the setting of an innovation crowdsourcing community, Hwang et al. (2019) found that individuals who are exposed to diverse issues of customers are more likely to generate creative ideas. Diverse analogies aid creative idea generation because they enable easy access to and/or identification of relevant knowledge whose retrieval can aid idea generation (Nijstad and Stroebe, 2006).

Researchers have also examined how analogical distance, conceptual distance between the existing knowledge domains of analogies and the target problem domain, affects the creativity of new ideas. Based on the analogical distance, analogies can be subdivided into near and distant categories - near are drawn from the same or very similar domains as the target problem and distant analogies are drawn from very different domains. Researchers posit that distant analogies are valuable in generating creative ideas because they introduce a fresh approach to the current problem and also help individuals be less constrained by established solutions of the problem domain area (Gentner et al., 1997; B. Ward et al., 1995; Yu et al., 2014; Dahl and Moreau, 2002). However, the empirical evidence in support of their arguments is mixed. Dunbar (1995) found near analogies play a very significant role in scientific inquiry whereas distant analogies are useful to highlight or educate others about the salient features of the problems under discussion. On the

other hand, studies on product designs have found benefits of using distant analogies and detriments of near ones on the originality and creativeness of the outcome (Dahl and Moreau, 2002; Christensen and Schunn, 2007). Yet, others such as Fu et al. (2013), Enkel and Gassmann (2010) have found no or a negative effect of analogical distance in ideation of creative solutions. This study aims to advance our understanding of the impact of analogical distance on creativity by applying a novel empirical approach.

Another important factor in generating creative outputs is learning. Arrow (1962) posited that learning is dependent on the agent's (here, ideator's) experience in solving problems. Moreover, these attempts at problem-solving, whether successful or not, lead to better performance when these problems are steadily evolving. Similarly, there is a wide acceptance among creativity scholars on the role of learning through deliberate practice in creative ideation. In a qualitative study of famous musicians involving The Beatles, Mozart, etc., it has been found that their fame was preceded by years of deep immersion to their craft (Weisberg, 2009). They utilized this extended period of time to practice and improve their skills after which their true creativity started to blossom. Similarly, chess players need to master over 50,000 chess-move patterns for at least ten years to reach the level of grandmaster (Chase and Simon, 1973). These studies underscore the importance of learning-by-doing in creativity.

A corollary of deliberate practice is the role of failure or failed prior attempts in spurring creative intelligence (Hayes et al., 1987; Wishbow, 1988; Sims, 2012; Thorley, 2018). The Geneplore model (A. Finke et al., 1992), built on the creative cognition approach to understanding human creativity, asserts that creative idea generation is an iterative process where existing knowledge is recalled and recombined in novel fashions to generate new ideas. This exploration continues until the created value best fits the goal of the innovation task. Although this role of failures and persistence are theoretically justified, there is scant empirical evidence on the positive role of failures in creativity outcomes (Hayes, 1989).

Based on the review of existing literature, we build a comprehensive model of idea generation focusing on the following determinants of creative ideation that are observable in an open innovation platform: knowledge acquisition and learning through deliberate practice. In contrast to previous empirical frameworks, our empirical approach allows for endogenous interplay of these processes. We explain our measure of knowledge acquisition and learning through deliberate practice in the next section. Besides the above factors,

there is unobservable individual heterogeneity that is expected to influence the creative outcome. For example, asymmetries in individuals' knowledge bases and unique approaches to framing problems may make some individuals more likely to be successful than others in their creative tasks (Lüthje et al., 2005; von Hippel, 1986). Our model incorporates such effect as unobserved heterogeneity in ideators' propensities to acquire knowledge and submit ideas for different problems.

2.2 Cryptocurrencies & Mining

2.2.1 Related Research

Our analysis of the dynamic mining game is closely related to the research on the economics of Bitcoin transaction fees and the mining competition. Ma et al. (2018) study the dynamics of the mining Bitcoin blocks in a dynamic setting with exogenous fees and rewards. They prove the existence of a unique Nash equilibrium in hash power allocations of predetermined number of miners for every difficulty level. Further, they find that higher difficulty levels lead to higher hash power allocations which leads to erosion in miner's profits as difficulty adjusts to maintain constant block generation times. Dimitri (2017) analyze a similar game of mining with heterogeneous cost structure to find the improbability of the existence of monopoly in Bitcoin mining.

While increases in mining power can be viewed as a wasteful expenditure, their true utility lies in securing the underlying blockchain and ensuring trust in the system (Houy, 2014). As such, competition helps maintain the robustness of the Bitcoin blockchain. Moreover, the collective rate of mining is also responsible for transaction processing, which in turn affects the levels of transaction fees (Easley et al., 2019). Such fees plays a major role in incentivizing miners to process transactions during times of high demand or high difficulty levels. Our empirical model, with endogenous determination of transaction fees and difficulty levels, aims to further increase our understanding of the competition within and between Bitcoin and Bitcoin-Cash mining.

2.2.2 Dynamic Game Estimation

Our dynamic model of mining assumes that miners' strategic behavior is in line with Markov perfect equilibrium (MPE) i.e. miners' actions are only dependent on the values of their current states and their private shocks (Maskin and Tirole, 2001). These actions impact the commonly observed state variables such as, difficulty levels and fees, which affect future strategic interactions. Using models based on MPE is fraught with computational difficulties owing to the difficulty in computing the allowed equilibria in the theoretical model and then, matching those equilibria to the ones in the data. This difficulty prohibits using Nested Fixed-Point types of estimators (Rust, 1987). We follow the econometric literature in estimating dynamic games using two-step estimators, based on forward simulation procedure of Hotz et al. (1994), which make it possible to recover the model parameters without solving for equilibrium even once (e.g. Bajari et al., 2007; Aguirregabiria and Mira, 2007; Pesendorfer and Schmidt-Dengler, 2008; Ryan, 2012; Blevins et al., 2017).

The two-step estimators are built on the assumption that in equilibrium, agents exhibit rational behavior conditional on their beliefs about competitors and environment. Therefore, recovering the probability distributions governing what agents do during equilibrium is akin to recovering their equilibrium beliefs. Once the equilibrium beliefs are recovered, econometrician can simulate many plays of the underlying game while imposing the restrictions of the theoretical model on these beliefs. This allows us to recover the underlying structural parameters that have generated the observed data.

2.2.3 Conceptual Background: Cryptocurrencies

In their seminal paper, Haber and Stornetta (1990) outlined the principles of blockchain technology, which is a chronologically-chained collection of data blocks. Their distributed (multiple copies reside on different nodes), decentralized (participating agents have equal voting rights), and append-only property (data can be added but not removed) made them the preferred underlying technology for BTC (Conley, 2017; Nakamoto, 2008). BTC's consensus mechanism, Proof-of-Work, enabled agents to agree on the state of the underlying blockchain without requiring trust in any central authority.

Using public key cryptography, BTC (and BCH) miners can verify whether the transactions on the

network are legitimate and their initiators own the coins being transacted. After verifying whether the coins being sent are owned by the sender and whether those coins have been spent already or not, transacted amount is associated with the recipient's account. Miners detect these initiated transaction requests from users, verify their information contents, and add a block of validated transactions to the blockchain (Peck, 2017). Usually, every transaction includes transaction fees which incentivizes the miner to include it into the block being processed. This entire process of verification and block generation is referred to as mining.

The process of block generation results in creation of new coins (known as coinbase) which act as a major incentive for participation in mining. During the period of our study, 12.5 coins of their respective currencies were created for every valid block of BTC and BCH. After all the verified transaction are bundled into a block, the current block's header (which includes information about the transactions in the block) and hash of the most recently generated block are hashed using the hashcash proof of work algorithm[1]. The goal of this mining exercise is to find a hash value that is less than the publicly-known difficulty target that all the BTC/BCH clients share. The winner (first miner to find the valid hash value) transmits its proof or proof-of-work on the network for every other miner to verify and update their copies of the underlying blockchains.

The difficulty target is a 256-bit number which is a measure of how difficult it is to find a hash value below a given target. It is revised after every 2016 blocks in BTC to keep the rate of coin generation at roughly a block every 10 minutes. If 2016 blocks are mined in less than two weeks, target is lowered to increase the rate of difficulty for the next set of 2016 blocks (Figure 2.1c). The difficulty never changes by more than a factor of 4. Since it is a append-only ledger, transactions, once buried under enough number of blocks (6 in BTC's case), are considered confirmed and cannot be reversed. This irreversibility is also a result of immense difficulty in computing valid proof-of-works (hash value below the target) which deters malicious actors from attempting to modify the existing records[2].

In the early days of Bitcoin mining, individual enthusiasts would use their personal computing resources

---

[1]A cryptographic hash function, like hashcash-SHA256[2], is a complex mathematical function which generates a fixed length output for any data of arbitrary size and is designed to be a one-way function (very hard to invert).

[2]Theoretically, a group of malicious actors can capture 51% of the total network mining power and reverse the recorded transactions. This is referred to as 51% attack and almost impossible to carry out.

such as central processing units or graphics processing units to mine. However, with growing demand and popularity, the level of difficulty has also risen by many orders of magnitude. It has led to establishment of mining pools and usage of Application-Specific Integrated Circuits (ASIC) mining equipment (Pilkington, 2016). A mining pool is a collection of individual miners who pool in their computing resources to reduce the volatility of their returns. Antpool, ViaBTC, BTC.TOP, BTC.com, etc. are some of the prominent mining pools. Energy-intensive nature of BTC mining adds significant costs to its production. BTC.com, Antpool, BTC.TOP, ViaBTC, Bitcoin.com, and F2Pool possess more than the half of the global Bitcoin mining power[3]. Therefore, we focus our analysis on their actions. Mining pool with more powerful ASIC mining equipment can compute trials of the proof-of-work faster than its rivals and thus, has a better chance of winning. However, such a powerful mining pool does not guarantee success in block generation due to inherent randomness owing to brute-force calculations in proof creation.

The mining process for BCH is very similar to BTC's as BCH inherited BTC's characteristics such as the SHA-256^2 based hashcash proof-of-work mechanism. BCH's developers did modify the difficulty adjustment algorithm for BCH. We elaborate on the new algorithm and its properties in the next subsection.

Bitcoin Fork

Nakamoto (2008) conceived BTC as a distributed, temper-proof, and decentralized peer-to-peer cryptocurrency. It did not require trust in third-party to perform verification. However, BTC was limited to processing 7 transactions per second. This was a resultant of BTC's technical architecture that restricted the sizes of its blocks to 1 MB and effectively, limited the number of transactions that could be accommodated in a single block to roughly 1500-2200. In comparison, Visa could handle up to 25K transactions per second (Raul, 2018). During higher network activity periods, block size limitation lead to slower confirmation times and higher transaction fees.

BTC stakeholders could not agree on the means and methods to scale BTC's throughput. Some of them felt the need for larger block sizes (8 MB) to process greater number of transactions. Though it partially solved the scaling issue, it created another problem with respect to centralization. Large blocks require

[3]https://hackernoon.com/why-mining-pool-concentration-is-the-achilles-heel-of\\-bitcoin-ce91089ce1f

mining nodes to possess expensive storage and bandwidth facilities that was expected to disproportionately affect the smaller miners and result in consolidation of mining power in the hands of big players. The core developers were strictly opposed to such centralization of mining power in lieu of higher throughput. They proposed their own scaling solution, however, the two groups could not reconcile. With the proposition of faster transaction processing times and lower fees, Bitcoin Cash was forked from the Bitcoin chain on 1^{st} August 2017.

Since BTC and BCH share the same codebase, BCH also inherited BTC's hashing algorithm (hashcash SHA-256^2), its difficulty adjustment algorithm (DAA), and its transaction history. Initially, they shared the same difficulty levels but BTC was priced much higher in comparison to BCH (Figure 2.1d,2.1c). To incentivize mining on BCH chain, its developers proposed Emergency Difficulty Adjustment (EDA) algorithm, which would lower the difficulty of BCH chain by 20% if the time difference between the 12^{th} and the 6^{th} blocks leading up to the current block was greater than 12 hours (Song, 2017a). This ensured that profit-maximizing miners would allocate a majority chunk of their mining power to BCH when difficulty adjusted profitability was in its favor (Song, 2017c). EDA helped the newly born minority chain, BCH, to survive by garnering the support of mining pools (Menegerian, 2017). The inherited BTC difficulty adjustment algorithm (adjustment after every 2016 blocks) was responsible for any upward revisions in BCH difficulty. BCH deprecated EDA on November 13^{th} 2017 and switched to a rolling DAA (BitcoinABC, 2017). This algorithm was based on a 144-period simple moving average. The difficulty was adjusted every block, based on the amount of work done for the past 144 blocks.

Wild fluctuations in computation or mining hash power were observed between BTC and BCH during the EDA-active periods (Figure 2.1b). Miners measure profitability in terms of difficulty adjusted rewards index or DARI (Equation 2.1). Triggering of EDA would result in reduction of BCH difficulty Haywood, 2017; Song, 2017b,c. This would increase BCH's DARI and incentivize miners to allocate more hash power to BCH at the cost of BTC (Figure 2.1e). Influx of mining power resulted in higher competition among miners for the limited low-difficulty BCH blocks. Such shifts, inadvertently, led to higher confirmation times on BTC network and pushed up BTC's transaction fees as BTC users competed for the attention of reduced number of miners (Figure 2.1f). Therefore, miners make repeated choices between the two, i.e.,

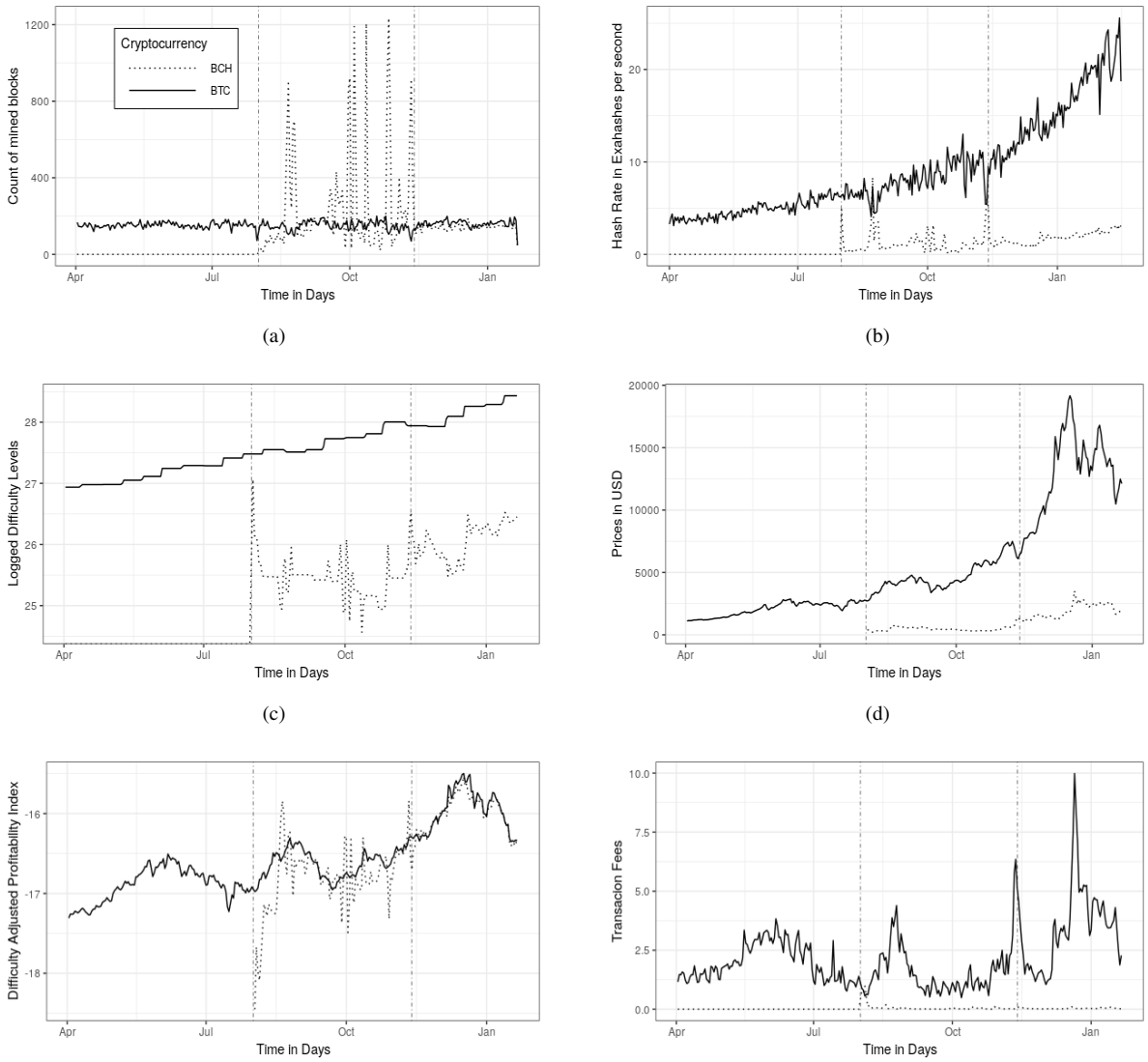


Figure 2.1: We present the plots for different state variables during the period of our study, April 2017 - February 2018. Vertical lines mark the time of BTC fork and the deactivation of Emergency Difficulty Adjustment algorithm on BCH. (a) Daily count of mined Bitcoin and Bitcoin Cash blocks. Before EDA was deprecated in mid-November, BCH chain witnessed extreme mining where 2016 blocks would be mined within 2-3 days. (b) Hash-power allocations for Bitcoin and Bitcoin Cash. (c) Difficulty Levels. (d) Price for Bitcoin and Bitcoin Cash. (e) Difficulty Adjusted Reward Index (2.1) for Bitcoin and Bitcoin Cash. (f) Average fees per transaction for Bitcoin and Bitcoin Cash.

mine highly-competitive low-priced low-difficulty BCH blocks) or mine high-priced high-difficulty BTC blocks with lower competition and higher transaction fees). Their dynamic decisions, in turn, affect the evolving difficulty and fees levels.

$$\text{DARI} = \frac{\text{block reward in USD + transaction fees in USD}}{\text{block difficulty}} \tag{2.1}$$

Chapter 3

LEARNING TO BE CREATIVE: A MUTUALLY EXCITING SPATIO-TEMPORAL POINT PROCESS MODEL FOR IDEA GENERATION IN OPEN INNOVATION

3.1 Empirical Methodology

3.1.1 Research Setting & Data

Our empirical context is a large open innovation platform that originated in the U.S (the open innovation platform hereafter). The open innovation platform is an online idea crowdsourcing platform, which invites crowds to solve socio-economic challenges such as food wastage, sanitation issues in India, e-waste recycling, combating viral diseases, etc. Individuals with different backgrounds and expertise participate to find solutions to these challenges. Challenges are sponsored by organizations such as United States Agency for International Development, Amplify Program, American Association of Retired Persons, Coca-Cola Enterprises, Bill and Melinda Gates foundation etc. The management team of the open innovation platform, in conjunction with the sponsors, creates a challenge brief that describes the challenge statement, and sets the expectations about their requirements.

A challenge usually lasts for three months and goes through the following life cycle. Majority of challenges begin with the research phase. During this phase, participants post potential reference ideas and online articles that could help create a cogent solution for the challenge statement. The research phase is followed by the ideation phase where participants create detailed solutions that can potentially solve the target problem. Once the solution idea is uploaded to the platform, other platform participants can vote upon these ideas. In the crowd selection phase, creative ideas are shortlisted based on the crowd voting results and the sponsors' feedback. Finally, in the expert selection phase, the shortlisted creative ideas are further scrutinized and refined for their technical and financial feasibility by an expert panel.

Although experts' evaluation is often used in an open innovation setting, a recent study found that there is a high likelihood of experts' unintended negligence in identification of creativity because of their overem-

phasis on feasibility of solutions (Piezunka and Dahlander, 2015). Howe (2006) finds crowd voting, where active community members evaluate the crowd-produced content, to be highly efficient in filtering and selecting creative outcomes accurately by tapping into the collective wisdom of the crowd. Consequently, we use the judgments derived from crowd voting supplemented with the sponsors' feedback in the crowd selection phase as an indicator of whether the submitted idea is a creative solution or not. This focus on creative ideas stems from our limited accessibility to implementation constraints of each challenge. These constraints (which can be of technological or financial nature) are set by the sponsors and can vary with problem types and sponsors' appetite.

We obtained activity stream data of 13,028 participants from 2010-2016. The activity streams of these platform participants consist of their timestamped idea submission and commenting activities. In our research setting, in addition to submitting ideas, platform participants can interact with others through casting votes and/or writing comments. Each comment represents the time and effort spent by a commenting participant to understand and analyze the corresponding idea. These ideas serve as sources of new knowledge and inspirations for the commenting participant's future ideas. The 13,028 participants have submitted a total of 7,510 ideas to 39 challenges, and 798 of these ideas were selected as creative ideas in the crowd selection phase. For convenience, we refer to the 798 ideas as successful ideas and the remaining as unsuccessful ideas. The 7,510 ideas received a total of 81,893 comments. We also have detailed information on the text contents of the submitted ideas and the challenge descriptions. Out of 13,028 participants, 11,384 have made at least one comment on some idea(s) on the platform, 4,418 have submitted at least one unsuccessful idea, and 649 have submitted at least one successful idea. Out of these 649 participants with at least one successful idea, 348 have experienced at least one failure and one success each. In these activity streams, we find 289 participants had at least one success after an initial failure whereas 59 participants failed at least once after an initial success.

Table 3.1 presents user frequency in terms of their knowledge acquisition, unsuccessful ideation, and successful ideation activities. For example, 1,644 users have never acquired knowledge from ideas posted on the platform, 5,228 have acquired knowledge once, 1,702 have acquired knowledge twice, etc. Similarly, Table 3.2 presents the distribution of users in terms of their participation in challenges. 11,211 users par-

ticipated, i.e. either submitted an idea or acquired knowledge, in only one challenge, 1,192 participated in two challenges, and 338 participated in three challenges, etc. Likewise, 4,323 users contributed ideas to a single challenge, 305 contributed ideas to two challenges, etc.

Table 3.1: Distribution of ideators' participation levels across different activity types

Activity Type	User Frequency						
	0	1	2	3	4	5	> 5
Knowledge Acquisition	1,644	5,228	1,702	891	573	408	2,582
Unsuccessful Ideation	8,610	3,453	546	202	88	38	91
Successful Ideation	12,379	572	51	11	5	3	7

Note. This table presents the user frequency of participation for different type of activities.

Table 3.2: Distribution of ideators' participation in unique challenges

	User Frequency					
	1	2	3	4	5	> 5
Unique challenges an ideator participated in	11,211	1,192	338	135	59	93
Unique challenges an ideator contributed to	4,322	305	55	14	9	23

Note. This table presents the user frequency for the number of unique challenges ideators participated in or contributed to. For participation, we include challenges that they acquired knowledge from or submitted an idea to. For contribution, we only include those challenges that they submitted an idea to.

3.1.2 Model Development

To enhance our understanding of how participants channelize the knowledge acquired through interactions into their idea generation process and generate creative ideas, this study applies the mutually exciting spatio-temporal point process model. In section 3.1.2, we explain our modelling strategy through an example. In section 3.1.2, we develop our empirical model that captures the dynamic interactions across participants' interactions and their idea generation activities. In section 3.1.2, we extend our model to incorporate conceptual distances between stimulus ideas and a target problem through the spatial extension of mutually exciting point processes.

Model Overview.

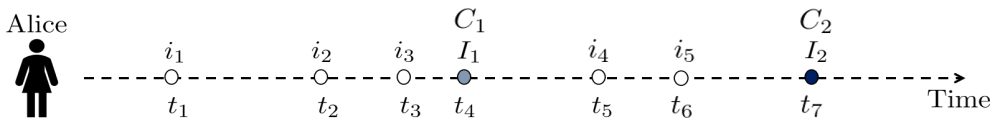


Figure 3.1: Illustrative example of Alice's interaction and idea generation activity.

To illustrate our modeling strategy, in Figure 3.1, Alice acquires knowledge from an idea i_1 at time t_1. Later, she acquires knowledge from other ideas, i_2 at t_2 and i_3 at t_3. She submits an idea, I_1 at t_4 for innovation challenge C_1, which is not chosen as a creative idea later. She, again, acquires knowledge from other ideas, i_4 at t_5 and i_5 at t_6. Later, she submits an idea I_2 at t_7 to another innovation challenge C_2, which is chosen as a creative idea. In such a scenario, how do we disentangle the effect of Alice's stimulus ideas (ideas of others that she acquired knowledge from) on her own idea's likelihood of success? For example, how do we apportion the effect of i_1, i_2 and i_3 on I_1? Existing empirical frameworks weigh the contribution of i_1, i_2, and i_3 equally on the likelihood of success of I_1 because participants' knowledge acquisition activities are assumed to be exogenous. However, knowledge acquisition from i_2 might have not occurred without her prior knowledge acquisition from i_1. Similarly, knowledge acquisition from i_3 might be a result of prior knowledge acquisitions from i_1 and i_2. As such, exogenous knowledge acquisition assumption will assign values that are not commensurate with their actual contribution. In other words, in Alice's example, existing empirical frameworks are unable to account for the dissipation of the past knowledge acquisitions' effects over other intermediate interactions and assign them fully to the submitted idea I_1. This dynamic calls to account for self-exciting effects among the same type of activities (e.g., the effects of occurrences of past knowledge acquisitions on the likelihood of occurrences of future knowledge acquisitions), before we ascertain their collective effect on idea generation.

In addition, apart from the self-exciting effects, we also need to account for the exciting effects among different activity types. For instance, in Alice's example, the generation of creative idea, I_2 at t_7 may be attributed to all prior stimulus ideas (i_1,i_2,i_3,i_4,i_5) and prior failure (I_1). To estimate the impact of all prior activities on I_2, existing empirical frameworks treat prior idea generation activities and stimulus ideas as

exogenously given. However, in reality, prior idea generation efforts influence the stimulus ideas individuals acquire knowledge from and vice versa. For instance, Alice may have acquired knowledge from stimulus ideas i_4 and i_5 because of her earlier failure at t_4 (submitted idea I_1 was unsuccessful). Without including such dynamic interactions among different activity types, our ability to draw conclusions regarding the exact role of stimulus ideas on idea generation will be quite limited.

Moreover, individuals' decisions to participate in a certain innovation task are also dependent on all prior stimulus ideas and idea generation experience because idea generation is very much dependent on availability of knowledge (Weisberg, 1998). That is, the kinds of knowledge that have been accumulated from other ideas is expected to influence the future problems an individual will participate in. For example, Alice's participation in innovation challenge, C_1 at t_4, and C_2 at t_7, is dictated by her accumulated stock of knowledge through her knowledge acquisitions from other ideas. If the contents of ideas i_1, i_2, and i_3 were different, it is quite possible that Alice would have not participated in C_1 but in some other challenge as her acquired knowledge might not be relevant for C_1 anymore. Existing studies did not model this endogenous task selection and take it as predetermined. In order to account for this crucial aspect of endogenous task selection, we extend the mutually exciting spatio-temporal point process model to high-dimensional text spaces: We include the machine-learned semantic understanding of the stimulus ideas and innovation tasks by representing their textual contents through high-dimensional numerical vectors.

Lastly, the key strength of the mutually exciting point process model is in its ability to infer the causal network of the data generating process (Bacry et al., 2015). Since a focal event can only cause those events which occur after it, our modeling framework allows us to disentangle the causal relationships between and within activity types. In addition, this framework allows us to calculate the expected number of direct descendants of every event type i.e. number of events of one type causally generated by another type. For instance, we can calculate the number of successful creative ideas generated as a result of unsuccessful idea submissions in the past or as a consequence of an individual's prior knowledge acquisition from other ideas. This capability gives us detailed insights into the mechanism of idea generation process.

A Mutually Exciting Point Process Model for Idea Generation.

The backbone of our modeling strategy is premised on multivariate point processes (also known as counting processes). A counting process, $N(\cdot)$ is a right-continuous non-negative integer-valued process which measures the cumulative number of occurrences of the events of the phenomenon under study. This process can also be understood as the sequence of random arrival times, $\mathbf{T} = \{T_1, T_2, ...\}$, at which the counting process $N(\cdot)$ has jumped or incremented in value (Laub et al., 2015). It is used to model the random occurrences of events in time and through its spatial extensions, in n-dimensional space. Point processes are characterized by the conditional intensity function or the hazard function,

$$\lambda(t|\mathcal{H}_t) = \lim_{\Delta t \to 0} \frac{\Pr\{N(t + \Delta t) - N(t) > 0|\mathcal{H}_t\}}{\Delta t} \tag{3.1}$$

$\mathcal{H}_t$ is the history of the arrivals up to but not including t. $\mathcal{H}_t$ includes the timestamps of the occurrences of the events along with other metadata associated with the points. $\lambda(t|\mathcal{H}_t)\Delta t$ is the instantaneous probability of an event occurrence in the interval $(t, t + \Delta t]$. The conditional intensity function, therefore, gives the rate of expected number of arrivals in the same interval conditioned on history $\mathcal{H}_t$.

In many natural phenomenon such as earthquake occurrences, we observe clustering of events, for example, one earthquake is followed by many aftershocks in the vicinity of the epicenter (Ogata, 1998). This clumpiness of events renders the commonly used Poisson process unusable to model such phenomenon due to its memoryless property: in Poisson process, past events do not affect the probability of future events. In 1971, A.G. Hawkes (Hawkes, 1971a,b) published the seminal paper on 'self-exciting processes' where the occurrence of a past event 'excites' the process to generate future event up until some time period, thereby accounting for the clumpiness in the data. These processes have found successful applications in various fields to model neuron interactions in neurosciences (Reynaud-Bouret et al., 2013), earthquake occurrences in seismology (Ogata, 1988; Veen and Schoenberg, 2008), re-tweet behavior on twitter (Mishra et al., 2016; Srijith et al., 2017; Yang and Zha, 2013), advertisement clicks in marketing (Xu et al., 2014), and market microstructure in finance (Bacry et al., 2015). Multivariate extension of the Hawkes processes allowed for mutually exciting behavior where events of one type could excite the events of other types. A D-dimensional

multivariate Hawkes process is a collection of D marginal counting processes, taking values in N^D, and is characterized by $N = (N_1, N_2, ..., N_D)$ whose individual intensities, $(\lambda_1, \lambda_2, ...\lambda_D)$, are given by

$$\begin{aligned}\lambda_k(t|\mathcal{H}_t) &= \mu_k + \sum_{j=1}^{D} \int_{-\infty}^{t} \phi_{kj}\left(t-u\right) N_j(du) \\ &= \mu_k + \sum_{j=1}^{D} \int_{-\infty}^{t} \alpha_{kj} f_{kj}\left(t-u\right) N_j(du) \\ &= \mu_k + \sum_{j=1}^{D} \sum_{m:t_m^j<t} \alpha_{kj} f_{kj}(t-t_m^j) \end{aligned} \tag{3.2}$$

where $k, j \in \{1, 2, ..., D\}$ and $N_j(du) = 1$ if an infinitesimal element (du) includes an event and 0 otherwise (Daley and Vere-Jones, 2003).

$\lambda_k(t|\mathcal{H}_t)$ is the risk of an occurrence of a k-type process's event at time t conditional on the history of the process, $\mathcal{H}_t$. Equation (3.2) can be decomposed into two parts: exogenous and endogenous parts. $\mu_k \in \mathbb{R}^+$ is the exogenous risk of occurrence of the k-type process i.e. an event occurrence of a homogenous Poisson process with the rate parameter μ_k. The summation of the stochastic integral over all the process types represents the endogenous event occurrences. $\phi_{kj}(t-u)$ is the excitation kernel or function which quantifies the self-excitation effects ($k = j$) and the mutual-excitation effects ($k \neq j$). It can be decomposed as a product of the excitation coefficient and the decay kernel. $f_{kj}(t-u)$ is the decay kernel which measures the decaying effect of past occurrences based on their time difference with current time. α_{kj} are the excitation coefficients which quantify the increment in the instantaneous intensity of k-type process due to the occurrence of a j-type process's event. Stability conditions dictate that $\int_0^\infty f_{kj}(t)dt = 1$ and $0 \leq \alpha_{kj} < 1 \; \forall \; k, j$ (Karabash, 2012). Hence, a probability density function can be used as the decay kernel. Our choice of kernel is the two-parameter gamma density function for its flexibility over more commonly preferred exponential density function (Halpin and De Boeck, 2013).

This process formulation can also be viewed as a branching process i.e. as an immigration-birth process where each event is either an immigrant or a descendant of the earlier immigrants. The immigrant ideas arrive as a homogeneous Poisson process and each immigrant arrival can give birth to second-generation ideas through the excitation function $\phi_{kj}(t-u)$. Hence, the first event is always the immigrant and events

that follow after can either be its children or other immigrants and their children. This ties in neatly with our conceptualization of endogenous idea generation. Thus, α_{kj} are also referred to as the branching coefficients and are strictly less than 1 lest the process explodes. The benefits of the branching process representation are discussed in the estimation subsection along with the expectation-maximization (EM) procedure.

In the context of idea generation in the open innovation platform, idea generation occurs either due to exogenous reasons or endogenous reasons. Exogenous reasons may include reaction to word-of-mouth publicity for a challenge or advertisement by the platform or general propensity of the ideators to participate. Endogenous influence is through participants' prior participation on the platform which influences them to participate further. Statistically, the process can be viewed as the superposition of a memoryless homogeneous Poisson process (representing the exogenous occurrences) and an inhomogeneous Cluster process which is dependent on its history (representing the endogenous process). Our model captures both exogenous and endogenous forces.

In our setting, there are three main processes interacting with each other: knowledge acquisition process, unsuccessful ideation process, and successful ideation process. As mentioned earlier, participants interact with other ideas on the platform through comments and this commenting activity signifies knowledge acquisition. This process enables a participant to understand different ideas which can be later recalled as analogies to solve new problems[1]. The second process is the unsuccessful ideation process. This process includes idea submissions which fail to make the cut in the crowd selection phase. Finally, the successful ideation process includes idea submissions which make the cut in the crowd selection phase. In our model, the knowledge acquisition, the unsuccessful ideation, and the successful ideas are the three different types of process ($k, j \in \{0, 1, 2\}$), respectively. A participant's history of activities consists of orderly timestamps belonging to these three processes along with the content data associated with the ideas he/she submitted and read. We measure time difference between processes in months (for example: if time difference between two events is 45 days, we would label the time difference as 1.5 months).

An example branching structure for our research setting is illustrated in Figure 3.2. Here, t_{kn} represents the n^{th} occurrence of the k^{th} process. The vertical orientation of events is arbitrary. Solid arrows represent

[1]Although users may also accumulate knowledge through passively browsing others' ideas without leaving any comments, the complete history of community users' viewing log data is not available.

self-excitation and dashed represents mutual-excitation. For example: t_{01} is a knowledge acquisition immigrant which leads to t_{11}, a failed idea submission. t_{11} excites another failed idea submission, t_{12}, which in turn leads to a successful idea submission, t_{21}. t_{02} is a knowledge acquisition immigrant which does not excite any other activity.

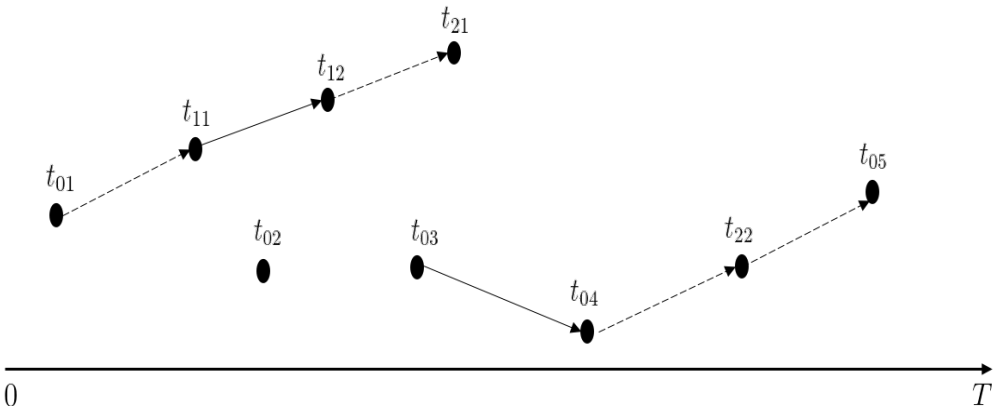


Figure 3.2: An example of the branching structure (i.e. interactions between different processes) over the time period $[0, T]$.

Spatial Extension of Hawkes Process

Hawkes processes can also be defined over a spatial distribution along with a temporal one i.e. the occurrence of events is also dependent on the exact spatial coordinates (user-defined abstract space) of events in the past along with their location on the temporal dimension. This implies that an occurrence of an event will affect the future occurrences of events in time as well as space. For instance, occurrence of an earthquake triggers more earthquakes or aftershocks in the aftershock region (clustered around the epicenter). To account for this spatio-temporal clustering of earthquakes, scholars developed the spatio-temporal Hawkes models (Ogata, 1998; Veen and Schoenberg, 2008). Similar to the earthquake occurrence processes, idea generation process can also be clustered along the spatial dimension of topics. That is, individuals' decisions to participate in innovation tasks are dependent on the interaction between the content (textual) properties of the prior analogies and the content or knowledge requirements of the innovation task itself (Weisberg, 1998). Therefore, this implies that as user accumulates knowledge through his or her interactions with oth-

ers' ideas, this acquired knowledge will also decide the kind of problem he/she will participate in the future. This endogenous selection of problems by the ideator is essential in understanding the analogical reasoning phenomenon in idea generation as individuals interact and participate based on their inherent preferences over topics. We model this interaction as a function of the conceptual distance between the contents of the target problem and prior analogies.

The conceptual distance is measured in the following way. Ideas are embedded in the vocabularies used by the ideators, and thus, any conceptual understanding of the idea can be inferred from the textual analysis of the idea (Kuhn, 2012). This necessitates the need to develop a numerical representation of these ideas that captures the underlying semantics involved in exposition. One of the most popularly used numerical representations is bag-of-words and text models built upon that representation. Every idea is represented as a count vector of words present in the global vocabulary. Addition of the tf-idf scheme (term frequency-inverse document frequency) increases the usability of such representations in identification of discriminative terms for classification but falls short of providing any information about semantic relationships between texts. These inadequacies led us to unsupervised machine learning technique of paragraph vectors, which take order of the words in a text document into account to deliver a fixed length semantic representation (Le and Mikolov, 2014).

Paragraph vectors are quite efficient in discovering the latent semantic structures in the text body. This method has been found to be the state-of-the-art in many natural language processing (NLP) tasks such as sentiment analysis and document clustering (Dai et al., 2015). Each paragraph vector is an output of the training process whose goal is to predict words in the documents. Training process samples words from a moving contextual window, thereby, capturing the semantics of the words in the order they are written. We train a paragraph vector model to infer one vector of length 300 for every submitted idea[2].

Cosine similarity between vectors has been used in the majority of NLP literature to quantify similarity between text vectors. However, angular distance has been found to have better discriminative properties at smaller angles of separation (Cer et al., 2018). Therefore, we measure analogical distance as the angular distance between a stimulus idea and the target challenge. This angle captures the conceptual dissimilarity

[2]Algorithmic details for the paragraph vector generation process are elaborated in the Appendix .2

between the underlying concepts for any pair of text vectors in the high-dimensional text vector space.

$$\text{angular distance} = \frac{\cos^{-1}\left(\text{cosine similarity}\right)}{\pi}$$

The spatial distribution is helpful in understanding that idea generation for a specific challenge is also a function of the accumulated knowledge (Weisberg, 1998). The rate of idea generation is, hence, affected by the spatial distribution of analogies with respect to the upcoming challenges. This allows us to capture the effect of analogical distance on creative output where the selected problem is itself a realization of the stochastic process rather than exogenously determined.

Let us assume that ideas and problems are points in some high-dimensional knowledge space. We represent them as such to accommodate the paragraph vector representation of ideas and problems in the knowledge space. In the case of earthquake occurrences, spatial distribution of events is defined over the $\mathbb{R}^2$ euclidean space. Assume $\mathcal{H} = \{(t_1, x_1, c_1), (t_2, x_2, c_2), ..., (t_n, x_n, c_n)\}$ is a realization of the orderly process i.e. $0 \leq t_1 < t_2 < ... < t_n$, where each x_i and c_i is a coordinate in high-dimensional vector space. Here, x_i is an idea generated for the corresponding c_i problem. Now, the endogenous part of the conditional intensity function for each marginal k^{th} process can be formulated as the multiplicative interaction of the spatial density function with the temporal density function (Ogata, 1998).

$$\begin{aligned}
\lambda_k(t|\mathcal{H}_t) &= \mu_k + \sum_{j=1}^{D} \int_{-\infty}^{t} \int_{\mathbf{X}} \phi_{kj}\left(t-u, \mathbf{c}_t - \mathbf{x}\right) .N_j(du, d\mathbf{x}) \\
&= \mu_k + \sum_{j=1}^{D} \int_{-\infty}^{t} \int_{\mathbf{X}} \alpha_{kj} \mathrm{g}_{kj}\left(t-u, \mathbf{c}_t - \mathbf{x}\right) .N_j(du, d\mathbf{x}) \\
&= \mu_k + \sum_{j=1}^{D} \sum_{m:t_m^j<t} \alpha_{kj} \mathrm{g}_{kj}\left(t - t_m^j, \mathbf{c}_t - \mathbf{x}_m^j\right) \\
&= \mu_k + \sum_{j=1}^{D} \sum_{m:t_m^j<t} \alpha_{kj} f_{kj}^{temporal}\left(t - t_m^j\right) f_{kj}^{spatial}\left(|\mathbf{c}_t - \mathbf{x}_m^j|_\theta\right)
\end{aligned} \tag{3.3}$$

In this formulation of the intensity function, the instantaneous rate of occurrence is also dependent on the angular distance (conceptual distance) between the challenge of the k^{th} process's event at time t, $\mathbf{c}_t$, and

the ideas that the user interacted with prior to time t. $\mathbf{c}_t$ is the numerical representation of the text of the challenge at time t. $\phi_{kj}(t-t_m^j, \mathbf{c}_t - \mathbf{x}_m^j)$ can be decomposed into the multiplicative form, $\alpha_{kj} \times f_{kj}^{temporal}(t-t_m^j) \times f_{kj}^{spatial}(|\mathbf{c}_t - \mathbf{x}_m^j|_\theta)$, where α_{kj} is the branching coefficient, $f_{kj}^{temporal}(\cdot)$ is the temporal decay kernel, $f_{kj}^{spatial}(\cdot)$ is the spatial density function, and $|\cdot|_\theta$ represents the angular distance. While usage of gamma density function to model temporal decay has been preferred in prior literature (Halpin and De Boeck, 2013), we select the density function for spatial effects empirically. Specifically, we experiment with different distributions that have support over $\mathbf{R}^+$ and select the one with the lowest AIC value (see the Appendix .1 for details). This leads to the usage of weibull density function for $f_{kj}^{spatial}(\cdot)$. The above formulation assumes that the effect of time and conceptual distance are independent of each other. Apart from reasons of computational tractability, we believe this is a reasonable assumption since the propensity to contribute to any challenge is decided by the past activities and the specific challenge to which the submission is made is decided by the spatial distribution of the topics of prior analogies.

Since our focus is on the spatial effects of acquired knowledge's topics on ideation outcomes, we only estimate $f_{kj}^{spatial}(\cdot)$ for $j \in \{$*knowledge acquisition*$\}$ and $k \in \{$*unsuccessful ideation*, *successful ideation*$\}$. For all other cases, such as effect of *unsuccessful ideation* on *successful ideation*, we do not include them in our model. In other words, we only estimate $f_{20}^{spatial}(\cdot)$ and $f_{10}^{spatial}(\cdot)$, and we do not estimate other spatial effects, for example, $f_{21}^{spatial}(\cdot)$, etc. These other effects, if present, are subsumed in their corresponding excitation parameters. Also for a fixed k, we assume $f_{kj}^{temporal}(\cdot)$ is same for all j i.e. on a given process, the effect of all processes decays at the same rate. Therefore, we estimate $f_{k}^{temporal}(\cdot)\ \forall\ k \in \{0, 1, 2\}$, to model the decaying effect of past activities on the rate of occurrence of the events of k^{th} process.

Platform users could also differ in their baseline rates of participation in different activities. This could be a result of individual participant's innate proclivity to submit more or fewer ideas, to acquire knowledge from others' ideas in varying degrees, and their exposure to platform's publicity efforts. Therefore, we allow for individual heterogeneity in the baseline rates, μ, for event occurrences. Though we model unobserved user heterogeneity in the baseline rates, the modelling structure of multivariate Hawkes process leads to these heterogeneities also being manifested in the excitation parameters leading to multi-dimensional heterogeneity structure. Our proposed model for the rate of occurrence of an event of the k^{th} process at time t

for user i is as follows:

$$\lambda_k^i(t|\mathcal{H}_t^i) = \mu_k^i + \sum_{j=1}^{D} \sum_{m:t_m^j<t} \phi_{kj}^i(t - t_m^j, \mathbf{c}_t - \mathbf{x}_m^j) \tag{3.4}$$

$$\phi_{kj}^i(t - t_m^j, \mathbf{c}_t - \mathbf{x}_m^j) = \begin{cases} \alpha_{kj}^i f_k^{temporal}\left(t - t_m^j\right) f_{kj}^{spatial}\left(|\mathbf{c}_t - \mathbf{x}_m^j|_\theta\right), & \text{if } t \geq t_m^j \text{ and } (k,j) \in \{(1,0), \\ & (2,0)\} \\ \alpha_{kj}^i f_k^{temporal}\left(t - t_m^j\right), & \text{if } t \geq t_m^j \text{ and } (k,j) \in \{(0,0), \\ & (0,1),(0,2),(1,1),(1,2), \\ & (2,1),(2,2)\} \\ 0, & \text{otherwise} \end{cases} \tag{3.5}$$

A summary of the key model parameters is provided in Table 3.3.

Table 3.3: Summary Description of Model Parameters

Model Parameter	Parameter Name	Description	
α_{kj}	Excitation Coefficient	Instantaneous increment in the conditional intensity of the k-type process due to the occurrence of a j-type process's event	
μ_k	Baseline Rate	Poissonian rate of arrival of events of the k^{th} process	
$f_k^{temporal}$	Temporal Density function	Measures the rate of decay of past events on the occurrence of the k^{th} process	
$f_{k0}^{spatial}$	Spatial Density function	Measures the spatial effect of conceptual distance between the target challenge of the unsuccessful/successful idea submission and prior knowledge acquisition	

3.1.3 Estimation

Ozaki (1979) derived the log-likelihood expression for the self-exciting Hawkes processes. The extension to multivariate case is relatively straightforward. Asymptotic properties of the maximum likelihood estimator,

such as consistency and efficiency, are discussed in Ogata (1978). The log-likelihood is given by

$$l(\theta) = \sum_{i=1}^{N} \left(\sum_{k=1}^{D} \left(\sum_{m=1}^{N_T^k} \log(\lambda_k^i(t^m|\mathcal{H}^i)) - \int_0^T \lambda_k^i(t|\mathcal{H}^i)dt \right) \right) \tag{3.6}$$

Equation (3.6) is also known as the incomplete data log-likelihood as the likelihood does not take branching information into account. N is the number of users and D is the number of marginal processes under consideration. In the context of Hawkes processes, the incomplete data log-likelihood has been found to be very flat near the optimal values (Veen and Schoenberg, 2008) leading to convergence issues and can be susceptible to the choice of the starting values of the parameters. The problem of flatness is further exacerbated in multivariate processes due to proliferation of parameters. In particular, the problem of numerical instability arises due to the inclusion of logarithm of the weighted sum of probability density functions (Halpin and De Boeck, 2013) in Equation (3.6).

The proposed solution to this problem involves the explicit usage of the branching structure of the Hawkes process (Rasmussen, 2013) which includes probabilistic information about the immigration and descendant status of events. Since event occurrences are either a resultant of baseline behavior or causal effect of any of the past events whose influence is proportional to the excitation provided by those events (Bacry et al., 2015), we can compute the posterior probabilities of the branching structure. This allows us to rewrite the Hawkes process as a mixture of Poisson processes which can be estimated by expectation-maximization algorithm. While the EM algorithm for temporal model of Hawkes process (3.2) can be found in Halpin (2013) and Veen and Schoenberg (2008), we present the modified estimation algorithm for the spatial extension of Hawkes process with unobserved user heterogeneity (Equations 3.4,3.5).

Our extension of the mutually exciting spatio-temporal point process model allows for non-parametric random effects in the specification of the baseline rates. In other words, the ideator-specific baseline rates are denoted by $\mu^i = \left(\mu_0^i, \mu_1^i, \mu_2^i\right)$ where i denotes the ideator. The mixing distribution is approximated by a discrete distribution whose masses and mass points (i.e. μ^i) are estimated along with the model parameters.

The mixing distribution is given by

$$\mathbb{P}\left(\mu^i = \mu^h\right) = \pi^h \quad \text{for } h = 1, 2, 3, ..., H; \quad \sum_{h=1}^{H} \pi^h = 1$$
$$\text{where } \mu^h = \left(\mu_0^h, \mu_1^h, \mu_2^h\right) \tag{3.7}$$

Let us define $\lambda_{kjm}^h(t)$ as the excitation effect of the m^{th} instance of the j^{th} process on the occurrence of an event of k^{th} process at time t when the baseline rate, $\mu^i = \mu^h$. Under this notation, $\lambda_{k00}^h = \mu_k^h$ is the baseline rate of k^{th} process.

$$\lambda_{kjm}^h(t) = \begin{cases} \phi_{kj}^h(t - t_m^j, \mathbf{c}_t - \mathbf{x}_m^j), & \text{if } t \geq t_m^j \text{ and } (k,j) \in \{(1,0),(2,0)\} \text{ and } \mu^i = \mu^h \\ \phi_{kj}^h(t - t_m^j), & \text{if } t \geq t_m^j \text{ and } (k,j) \in \{(0,0),(0,1),(0,2),(1,1),(1,2), \\ & (2,1),(2,2)\} \text{ and } \mu^i = \mu^h \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

In the modified EM algorithm, the random effects and the branching structure conditional on the random effect, are treated as the missing variables. We drop the i superscript, denoting users, for the ease of exposition. Let z_{kjm}^h be the process of k^{th} process's responses to the event t_{jm} i.e. events of the k^{th} process generated due to the event t_{jm} (m^{th} event of the j^{th} process) when the baseline intensity is μ^h. If any event t_{kn} is a response to event t_{jm}, then $t_{kn} \in z_{kjm}^h$ with $\mu^i = \mu^h$. If t_{kn} is a spontaneous occurrence, then $t_{kn} \in z_{k00}^h$. We define the missing data variables for the branching structure conditional on the random effect, as $Z_k^h = (Z_{k1}^h, ..., Z_{kN_k}^h)$. Z_{kn}^h indicates the z^h process to which the event t_{kn} belongs to i.e. we write $Z_{kn}^h = z^h$ if $t_{kn} \in z^h$ for $z^h \in z_k^h$ when the baseline risk is μ^h. Let $\theta^{(r,h)}$ denote the r^{th} iteration parameters for the h^{th} random effect i.e. $\mu^h \in \theta^{(r,h)}$. The conditional posterior probabilities for the branching structure are given by:

$$\text{Prob}(Z_{kn}^h = z^h \mid \mathcal{H}_t, \theta^{(r,h)}) = \frac{\lambda_z^{(r,h)}(t_{kn})}{\sum_{z' \in z_k^h} \lambda_{z'}^{(r,h)}(t_{kn})} \tag{3.9}$$

We use these conditional probabilities to calculate the likelihood for observing the data, $f(\theta|\theta^{(r,h)}, \mathcal{H}_t)$

when $\mu^i = \mu^h$:

$$\log f(\theta|\theta^{(r,h)}, \mathcal{H}_t) = \sum_{users} \Bigg(\sum_{k=1}^{D} \Bigg(\sum_{z \in z_k^h} \Bigg(\sum_{n=1}^{N_k} \log(\lambda_z^h(t_{kn})) \times \text{Prob}(Z_{kn}^h = z^h \mid \mathcal{H}_t, \theta^{(r,h)}) - \int_0^T \lambda_z^h dt \Bigg) \Bigg) \Bigg) \tag{3.10}$$

Finally, the posterior probabilities for the random effects are given by,

$$\mathbb{P}\left(\mu = \mu^h | Z^h, \theta^{(r,h)}, \mathcal{H}_t\right) = \frac{\pi^h \times f(\theta|\theta^{(r,h)}, \mathcal{H}_t)}{\sum_{h'=1}^{H} \pi^{h'} \times f(\theta|\theta^{(r,h')}, \mathcal{H}_t)} \tag{3.11}$$

In the iterative maximization step, we will maximize the expected complete data log-likelihood, $Q(\theta|\theta^{(r)}, \mathcal{H}_t)$:

$$Q(\theta|\theta^{(r)}, \mathcal{H}_t) = \sum_{users} \Bigg(\sum_{h=1}^{H} \mathbb{P}\left(\mu = \mu^h | Z^h, \theta^{(r,h)}, \mathcal{H}_t\right) \times \Bigg(\sum_{k=1}^{D} \Bigg(\sum_{z \in z_k^h} \Bigg(\sum_{n=1}^{N_k} \log(\lambda_z^h(t_{kn})) \times \text{Prob}(Z_{kn}^h = z^h \mid \mathcal{H}_t, \theta^{(r,h)}) - \int_0^T \lambda_z^h dt \Bigg) \Bigg) \Bigg) \Bigg) \tag{3.12}$$

The posterior probabilities, derived from equations (3.9) and (3.11), can be used to estimate the expected count of immigrants and descendants for all event types. For example, expected count of k^{th} process events generated due to j^{th} process will be given by $\sum_{users} \left(\sum_{h=1}^{H} \mathbb{P}\left(\mu = \mu^h | Z^h, \theta^{(r,h)}, \mathcal{H}_t\right) \left(\sum_n \text{Prob}(Z_{kn}^h = z^h \mid \mathcal{H}_t, \theta^{(r,h)}) \right) \right) \forall\ z^h \in \left\{ z_{kj1}^h, z_{kj2}^h, ..., z_{kjN_j}^h \right\}$. The asymptotic standard errors were computed by using the hessian of Q evaluated at the MLEs (Jamshidian and Jennrich, 2000).

A key problem in calculating the complete data log-likelihood arises due to the integral over $\lambda_z^h(t)$ in Equation (3.12). Since the text vectors lie in a high-dimensional space, we are left with two options: either to employ dimensionality reduction techniques, such as PCA or t-SNE, to reduce the text embedding space to two dimensions ideally or to solve this integral using numerical approximation. Conceptual distances calculated using low-dimensional embeddings have a very low coefficient of correlation with the conceptual distances calculated using original vectors. Therefore, we numerically approximate the integral as shown

below:

$$\begin{aligned}
\int_0^T \lambda_z^h dt &= \int_0^T \int_{\theta \in \Theta_m^j} \phi_{kj}^h (t - t_m^j, \mathbf{c}_t - \mathbf{x}_m^j) dt d\theta \\
&= \int_0^T \int_{\theta \in \Theta_m^j} \alpha_{kj} f_k^{temporal}(t - t_m^j) f_{kj}^{spatial}(|\mathbf{c}_t - \mathbf{x}_m^j|_\theta) dt d\theta \\
&= \int_0^T \alpha_{kj} f_k^{temporal}(t - t_m^j) \left(\int_{\theta \in \Theta_m^j} f_{kj}^{spatial}(|\mathbf{c} - \mathbf{x}_m^j|_\theta) d\theta \right) dt \\
&= \alpha_{kj} F_k^{temporal}(T - t_m^j; \Phi) F_{kj}^{spatial}(\max(\Theta_m^j) - \min(\Theta_m^j); \Phi) \qquad (3.13)
\end{aligned}$$

The set Θ_m^j is the set of analogical distances, $\mathbf{x}_m^j$, of analogy at time t_m^j from all the challenges that the ideator has participated in or could have participated in during his or her tenure on the platform. Therefore, we integrate over all the challenges that were active during the focal ideator's activity stream. Φ is the set of parameters governing the spatial and temporal distributions. $F(\cdot; \Phi)$ is the corresponding cumulative distribution function. Please see Appendix .3 for more details.

3.1.4 Identification Strategy

Our identification strategy is premised on (1) the exogeneity of challenge arrivals on the platform, (2) the exogeneity of others' idea arrivals from which an ideator can acquire knowledge, and (3) the capability of the mutual-exciting framework to explicitly model endogenous process of individuals' decisions (to choose specific challenges to participate in and ideas to acquire knowledge from).

First of all, the exogenous sequencing of challenges on the platform occurs independently of the ideators' actions. The content and requirements for all challenges are decided by the platform (in conjunction with the sponsor) and taken as given by the platform users. As such, any ideator has very limited control over which challenges to participate in (i.e. s/he can only participate in those challenges that are active at any point of time). In our framework, the decision to submit an idea to any challenge is dependent on the entire history of the ideator's actions and the content properties of the posted challenge. These exogenous occurrences of challenges provide us with the required variation in content properties of the data to identify the structural parameters.

Second, the existing ideas on the platform, which the ideator can acquire knowledge from, occur exogenously. At any point in time, a platform participant chooses to acquire knowledge from these given ideas (that have been submitted by other ideators in the past) i.e. the participant takes the presence of these ideas on the platform as given. These exogeneity conditions combined with the modelling structure imposed using the inter-temporal ordering of event occurrences aids us in identification.

Third, the mutually-exciting framework implicitly accounts for non-systematic unobservables and idiosyncratic shocks (Xu et al., 2014). For example, a platform participant could be affected by discussions on specific topics outside of the platform which could impact his/her knowledge acquisition and ideation behavior on the platform. As long as this exposure to outside discussions is not systematic, the expected effect of all such discussions is included in the model parameters that enter the stochastic conditional intensity function ($\lambda^i(t)$). An example of systematic exposure could be correlation of challenge arrivals with contemporaneous topics of discussion in the wider world. Given the general and universal nature of topics of challenges on the platform, like food waste reduction or climate change or women safety, etc., there is no reason to suspect such a contemporaneous correlation.

Lastly, we have also incorporated individual heterogeneity into our model, which further controls for unobserved ideator characteristics that could bias our results, and strengthens our identification strategy.

3.2 Results

3.2.1 Results for the Mutually Exciting Spatio-Temporal Point Process Model

The goal of this study is to investigate the effect of stimulus ideas on the creative outcome in an open innovation platform. Specifically, we are interested in quantifying the effect of conceptual distance of stimulus ideas from the target problem in spurring creative intelligence. Our model framework incorporates key components of idea generation process, namely, knowledge acquisition, learning from (successful or unsuccessful) experiences, and unobserved individual heterogeneity. Our model endogenizes occurrences of knowledge acquisition and idea generation, as well as the selection of the specific challenges for which the ideas are generated.

We present our results in Table 3.4. As explained in the previous section, the excitation coefficients,

α_{kj}, is interpreted as the instantaneous increment in the probability of occurrence of k^{th} process event due to an event of j^{th} process. We contrast the excitation coefficients, α_{kj}, with its corresponding μ_k values (the baseline intensity of process k) which serves as a benchmark. In our model, the three main processes, knowledge acquisition, unsuccessful ideation, and successful ideation, are denoted as process 0, 1, and 2 respectively. Each column of Table 3.4 presents the results for the k^{th} marginal process. To ease the comparison with the baseline rates, we use the weighted average of the $\mu^h \ \forall \ h$ as the reference value, denoted as $\mu^{avg} = (\mu_0^{avg}, \mu_1^{avg}, \mu_2^{avg})$. Using the posterior probabilities from Equation (3.11) as the weights, we obtain $\mu^{avg} = (1.1353, 0.2336, 0.0246)$.

We first discuss the excitation effects of knowledge acquisition. Based on our results, the baseline intensity of the knowledge acquisition activity ($\mu_0^{avg} = 1.1353$) is almost 59% higher than its self-exciting effect ($\alpha_{00} = 0.7146$). This means that individuals acquire knowledge from the platform, at a rate of approximately 1 per month, and through self-excitation, they increase their knowledge acquisition rate by 0.7146. This self-excitation can be interpreted as the explicit effect of prior knowledge acquisition on individuals' inquisitiveness to acquire more knowledge through other ideas. Further, our results indicate that knowledge acquisition activity does not directly excite a successful idea generation in the future: α_{20} is statistically insignificant. Even though knowledge acquisition through interactions does not directly influence successful idea generation, our model uncovers second-order effects (i.e. child of a child of an immigrant/child): We find that knowledge acquisition activities lead individuals to generate more ideas, even though those ideas are unsuccessful ($\alpha_{10} = 0.0575$), but these unsuccessful ideas boost the intensity of successful idea generation in the future ($\alpha_{21} = 0.0277$). This increase in the probability of successful idea generation is equivalent to the baseline intensity of successful idea generation ($\mu_2^{avg} = 0.0246$).

Based on the branching process representation, the excitation coefficients can also be interpreted as branching coefficients i.e. α_{kj} is the expected number of k^{th} process's child events spawned by an occurrence of an event of j^{th} process. Hence, using our estimates, we compute the expected count of every activity that is generated as a result of baseline intensities and excitation effects. Table 3.5 presents the expected counts. We calculate their standard errors using the Delta method[3]. We find that, on average, 68% of

[3]Asymptotic Normality of the MLE is discussed in Ogata (1978) which makes the usage of Delta method possible.

Table 3.4: Parameter Estimates of the Mutually Exciting Spatio-Temporal Point Process Model with Random Effects

Process Types	Knowledge Acquisition	Unsuccessful Ideation	Successful Ideation
	Baseline Intensities		
π^1	μ_0^1	μ_1^1	μ_2^1
0.7256	0.9686^{***}	0.3076^{***}	0.0256^{***}
	(0.0090)	(0.0051)	(0.0015)
π^2	μ_0^2	μ_1^2	μ_2^2
0.0405	0.4067^{***}	0.0296^{***}	0.0037^{***}
	(0.0106)	(0.0029)	(0.0010)
π^3	μ_0^3	μ_1^3	μ_2^3
0.0923	1.1383^{***}	0.0083^{***}	0.0001
	(0.0207)	(0.0018)	(0.0001)
π^4	μ_0^4	μ_1^4	μ_2^4
0.0979	3.1197^{***}	0.0689^{***}	0.0586^{***}
	(0.0375)	(0.0056)	(0.0051)
π^5	μ_0^5	μ_1^5	μ_2^5
0.0436	0.1234^{***}	0.0378^{***}	0.0018^{***}
	(0.0036)	(0.0020)	(0.0004)
	Exciting Effects		
Knowledge Acquisition	α_{00}	α_{10}	α_{20}
	0.7146^{***}	0.0575^{***}	0.0768
	(0.0030)	(0.0032)	(0.0780)
Unsuccessful Ideation	α_{01}	α_{11}	α_{21}
	0.2463^{***}	0.1791^{***}	0.0277^{***}
	(0.0061)	(0.0056)	(0.0023)
Successful Ideation	α_{02}	α_{12}	α_{22}
	0.2106^{***}	0.0000	0.0912^{***}
	(0.0175)	(0.0001)	(0.0122)
	Temporal Distribution		
Shape	k_0	k_1	k_2
	0.2624^{***}	0.1564^{***}	0.1213^{***}
	(0.0012)	(0.0034)	(0.0068)
Scale	θ_0	θ_1	θ_2
	0.0745^{***}	2.3016^{***}	4.3432^{***}
	(0.0007)	(0.1919)	(1.2205)
	Spatial Distribution		
Shape	k_0'	k_1'	k_2'
	n.a.	5.1683^{***}	1.5089^{*}
	n.a.	(0.2337)	(0.7273)
Scale	θ_0'	θ_1'	θ_2'
	n.a.	0.4005^{***}	0.1609^{*}
	n.a.	(0.0046)	(0.0771)

Note: Standard Errors are reported under the parameter estimates in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.5: Expected Counts of comments and ideas generated through self and mutual excitation.

Process Types	Knowledge Acquisition	Unsuccessful Ideation	Successful Ideation
	Baseline Intensities		
	$24,235.5893^{***}$	4342.1881^{***}	468.5578^{***}
	(155.6778)	(65.8953)	(21.6462)
	Exciting Effects		
Knowledge Acquisition	$55,905.7629^{***}$	1315.5455^{***}	112.2697
	(236.4669)	(73.8648)	(114.0499)
Unsuccessful Ideation	1606.4982^{***}	1054.2663^{***}	159.8303^{***}
	(40.0812)	(33.2340)	(13.1070)
Successful Ideation	145.1495^{***}	0.0000	57.3423^{***}
	(12.0478)	(0.0574)	(7.6453)

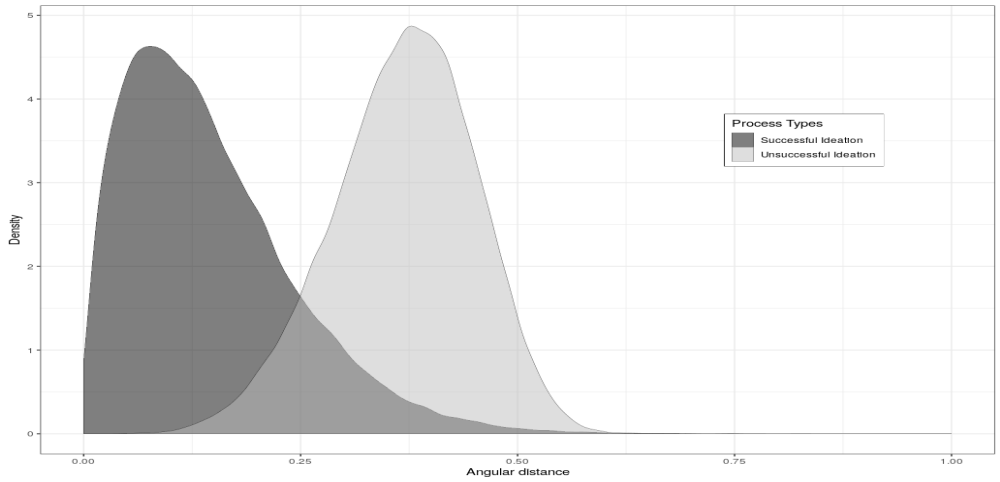
Note: Standard Errors are reported under the parameter estimates in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the knowledge acquisition activity (55,906 out of a total of 81,893 knowledge acquisitions) is a consequence of the prior knowledge acquisition activities (self-exciting effect) of the ideators. Moreover, the expected count of unsuccessful ideas generated due to knowledge acquisition (mutually exciting effect) is 1,316 (out of 6,712 unsuccessful ideas submissions). This result stresses the important role of knowledge acquisition activities in stimulating future participation.

The interpretation of how knowledge acquisition activities affect creative outcomes can further be enriched by incorporating the effect of conceptual distances between stimulus ideas and the target problems. As we described in section 3.1.2, ideas are clustered along the spatial dimension of conceptual topics just as earthquake occurrences are spatially clustered in the aftershock regions. Our model captures such spatial effects through the angular distance between the contents of stimulus ideas and target problem. Low value of the angular distance would indicate that the knowledge acquired through stimulus ideas is conceptually closer to the chosen target problem while high value would indicate that the chosen target problem is conceptually distant (or unrelated) to the knowledge in the stimulus ideas.

Figure 3.3 depicts the density plots illustrating the varying impact of the angular distance on the rate of unsuccessful and successful ideation. The distribution of the unsuccessful ideation density plot indicates that the stimulus ideas that are driving the statistically significant excitation effect of α_{10} are conceptually

Figure 3.3: Density plots for the Spatial Distributions (Weibull).



distant from the problem contexts for which the focal ideas are generated: As the stimulus ideas become more and more uncorrelated with the target problem (i.e. their angular distance increases from 0.25 towards 0.5), the spatial density function for the unsuccessful ideation process increases many folds than it does for the successful ideation process. Because the contribution of every prior knowledge acquisition activity is included as a product of the excitation coefficient, temporal decay, and spatial effect in the conditional intensity function (3.5), therefore, for a fixed temporal decay, the spatial effect of distant analogies increases the risk of occurrence of an unsuccessful idea ($\lambda_1(t)$) more than it does for a successful idea ($\lambda_2(t)$). In a similar fashion, we observe that the stimulus ideas, which are conceptually closer to the target problem, are driving the statistically insignificant excitation effect of α_{20} as the mean of the spatial distribution governing successful ideation is 0.1452 (compared to the mean value of 0.3684 for unsuccessful ideation). These low-distance stimuli are, however, unable to excite any successful idea as indicated by statistical insignificance of α_{20}.

Next, we turn our attention to the excitation effects of idea submissions. The self-exciting coefficients of unsuccessful ideas ($\alpha_{11} = 0.1791$) and successful ideas ($\alpha_{22} = 0.0912$) are statistically significant and almost 0.77x and 3.7x of their baseline intensities ($\mu_1^{avg} = 0.2336$ and $\mu_2^{avg} = 0.0246$). These results indicate that every unsuccessful and successful idea submission leads to instantaneous increment in the probability of

future submissions of the same kind by 0.1791 and 0.0912 respectively. Furthermore, the result indicates that past failure leads to future success: Mutual excitation effect from unsuccessful ideation efforts to successful ideation ($\alpha_{21} = 0.0277$) is comparable with the $\mu_2^{avg} = 0.0246$. This result suggests that through deliberate practice and failures, an individual learns how to apply his or her accumulated knowledge correctly, which leads to future successes. In contrast, we do not find any evidence of past success leading to future failures ($\alpha_{12} = 0.0$). These excitation effects offer evidence of the presence of learning-by-doing (Arrow, 1962). Prior ideation efforts have a direct impact on the occurrence of future idea submissions and these efforts can lead to better performance as illustrated by α_{22} and α_{21} (160 and 57 successful ideas, out of a total of 798 successful ideas, were expected to be generated due to excitation from past successful and unsuccessful idea submissions respectively). In other words, ideators can learn through their failures and successes to create successful ideas in the future.

The statistically significant mutual excitation effects from idea submissions to knowledge acquisition activities (i.e. α_{01}, α_{02}) suggest that every idea submission, unsuccessful or successful, leads to more knowledge acquisition by the focal ideator. The rate of increment is $\alpha_{01} = 0.2463$ for unsuccessful ideation and $\alpha_{02} = 0.2106$ for successful ideation. This result indicates that individuals are more likely to acquire knowledge from others' ideas after a failure rather than a successful submission. Through these knowledge-seeking activities, individuals not only collect new knowledge, but also learn how that knowledge has been applied by their corresponding ideators. Comparatively higher rate of knowledge acquisition after failure suggests the presence of a "learning drive" i.e. ideator's failure has a direct influence on him or her to seek more knowledge from others' ideas.

Lastly, the estimates in the temporal distribution section present the decay effect on the rate of event generation of different processes. We find that the excitation effect decays rapidly for knowledge acquisition activities i.e. ability of past platform activities to spur future knowledge acquisitions diminishes speedily in comparison to other two processes. Effect of past activities declines less rapidly for unsuccessful idea generation and even more slowly for successful idea generation. This implies that the rate of occurrence of successful idea submissions is affected by events way more in the distant past when compared with rates of occurrences for unsuccessful idea submissions and knowledge acquisitions.

3.2.2 Benchmark Models Analysis

Our mutually exciting framework allows for a variety of model nesting strategies to validate the incremental impact of the underlying theoretical assumptions (Halpin and De Boeck, 2013; Xu et al., 2014). In this subsection, we build a baseline benchmark model (with maximum restrictions on model parameters) and then, incrementally relax these restrictions in subsequent models for model comparison purposes. The restrictions were designed to aid our understanding of how the impact of conceptual distance varies as we build up our model with just (a) knowledge acquisition without explicit spatial effects (Knowledge Acquisition Model without Spatial Effects), (b) knowledge acquisition with spatial effects (Knowledge Acquisition Model), (c) knowledge acquisition with spatial effects and learning-by-doing through self-excitation (Self-Excited Knowledge Acquisition Model), and (d) knowledge acquisition with spatial effects and learning-by-doing through self- and mutual-excitation (Fully-Excited Knowledge Acquisition Model). The results of these benchmark models help us in understanding the relative impact of knowledge acquisition and learning-by-doing activities on creative ideation outcome. All benchmark models include the effects of unobserved individual heterogeneity. We focus our attention on the main parameters of interest, namely, the spatial parameters ($k_1', k_2', \theta_1', \theta_2'$) and the excitation coefficients (α_{kj}).

The baseline benchmark model, the Knowledge Acquisition model without spatial effects, captures the excitation effect of knowledge acquisition on idea generation while ignoring self-excitation among all processes and all other mutual-excitation effects ($\alpha_{00} = \alpha_{01} = \alpha_{02} = \alpha_{11} = \alpha_{12} = \alpha_{21} = \alpha_{22} = 0$), and $f_{kj}^{spatial}(\cdot) = 1$. In this model, we allow only past knowledge acquisition activities (without explicitly modelling their spatial effects) to affect future ideation outcomes. In the next benchmark model, the Knowledge Acquisition model, we relax the restrictions on $f_{kj}^{spatial}(\cdot)$ and estimate the spatial effects of knowledge acquisition activities. This model still restricts all other excitation effects.

The third benchmark model is the Self-Excited Knowledge Acquisition model where the occurrences of successful and unsuccessful idea submissions are allowed to increase the probability of future occurrences of successful and unsuccessful idea submissions respectively: i.e. α_{11} and α_{22} are no longer set to zero. Thus, only $\alpha_{00} = \alpha_{01} = \alpha_{02} = \alpha_{21} = \alpha_{12} = 0$. In our final benchmark model, referred to as the Fully-Excited Knowledge Acquisition model, we further extend the Self-Excited Knowledge Acquisition model by

relaxing the restrictions on α_{21} and α_{12}: i.e. we now also allow for mutual-excitation between the successful and unsuccessful ideation processes. This model comes closest to our main model. However, in comparison to the main model, it ignores the excitation in the generative process for the knowledge acquisition activities ($\alpha_{00} = \alpha_{01} = \alpha_{02} = 0$): i.e. knowledge acquisition process is restricted to follow a homogeneous Poisson process (we employ this restriction in all the benchmark models).

Table 3.6: Model Comparison: Goodness-of-fit

Models	AIC
Knowledge Acquisition	-95553.43
Self-Excited Knowledge Acquisition	-109692.60
Fully-Excited Knowledge Acquisition	-111872.80
Main Model	-253079.00

Table 3.6 presents the goodness-of-fit scores for the benchmark models. The results suggest that our proposed main model (3.4) outperforms all the benchmark models by a very wide margin based on the AIC scores[4] (see Table 3.6). The relative likelihood of our second best model in comparison to our main model is $\approx 2.9058 \times 10^{-30663}$. Through the Knowledge Acquisition model without spatial effects, we find that participants' knowledge acquisition activities on the platform lead to future occurrences of idea generation. In the second model, the Knowledge Acquisition model, we further dissect this result by explicitly including the effect of conceptual distances. We find that the prior analogies increase the probability of occurrence of unsuccessful ideation more than that of successful ideation ($\alpha_{10} = 0.0870, \alpha_{20} = 0.0152$), but there is no discernible difference between their spatial properties (Figure 3.4). In the Self-Excited Knowledge Acquisition model, where we allow for self-excitation in the ideation processes (i.e. we also estimate α_{11}, α_{22}), we find similar effect of knowledge acquisition activities ($\alpha_{10} = 0.0569, \alpha_{20} = 0.0154$) on ideation outcomes. However, we do notice slight difference in the spatial distributions $f_{10}^{spatial}(\cdot)$ and $f_{20}^{spatial}(\cdot)$ (Figure 3.4).

In the Fully-Excited Knowledge Acquisition model, we further relax the constraints on mutual-excitation between ideation processes i.e. α_{21} and α_{12} are estimated along with α_{10}, α_{11}, α_{20}, α_{22}, $f_{10}^{spatial}(\cdot)$, $f_{20}^{spatial}(\cdot)$. After the inclusion of the mutual-excitation effects, we find that prior analogies no longer incre-

[4]Since the random data for the Knowledge Acquisition model without spatial effects does not include the spatial data (unlike all other models), we can not compare its goodness-of-fit measure with the other models.

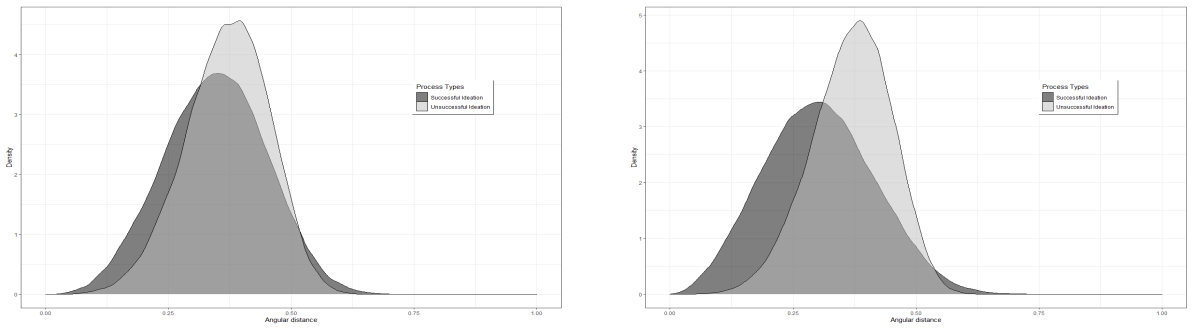
Table 3.7: Model Comparison: Parameter Comparison

Parameter	Models				
	Knowledge Acquisition Model w\o Spatial Effects	Knowledge Acquisition Model	Self-Excited Knowledge Acquisition Model	Fully-Excited Knowledge Acquisition Model	Main Model
α_{10}	0.0061^{***}	0.0870^{***}	0.0569^{***}	0.0568^{***}	0.0575^{***}
	(0.0003)	(0.0041)	(0.0031)	(0.0031)	$0.0032)$
α_{20}	0.0003^{***}	0.0152^{***}	0.0154^{**}	0.1184	0.0768
	(0.0001)	(0.0028)	(0.0049)	(0.2618)	(0.0780)
α_{11}	$n.a$	$n.a$	0.1825^{***}	0.1826^{***}	0.1791^{***}
			(0.0057)	(0.0056)	(0.0056)
α_{22}	$n.a$	$n.a.$	0.1072^{***}	0.0930^{***}	0.0912^{***}
			(0.0136)	(0.0124)	(0.0122)
α_{12}	$n.a$	$n.a.$	$n.a.$	0.0000	0.0000
				(0.0000)	(0.0001)
α_{21}	$n.a$	$n.a.$	$n.a.$	0.0292^{***}	0.0277^{***}
				(0.0024)	(0.0023)
k_1'	$n.a$	4.9443^{***}	5.2021^{***}	5.2060^{***}	5.1683^{***}
		(0.1881)	(0.2307)	(0.2307)	(0.2337)
k_2'	$n.a$	3.7090^{***}	2.9613^{***}	1.1907	1.5089^{*}
		(0.5224)	(0.5848)	(0.6898)	(0.7273)
θ_1'	$n.a$	0.4024^{***}	0.4006^{***}	0.4007^{***}	0.4005^{***}
		(0.0040)	(0.0045)	(0.0045)	(0.0046)
θ_2'	$n.a$	0.3818^{***}	0.3397^{***}	0.1164	0.1609^{*}
		(0.0191)	(0.0335)	(0.1319)	(0.0771)

Note: Standard Errors are reported under the parameter estimates in parenthesis. * p < 0.05, ** p < 0.01, *** p < 0.001

ment the probability of occurrence of successful ideas (α_{20} is statistically insignificant). Moreover, we find a clear difference in the spatial distributions for successful and unsuccessful ideation. However, the spatial parameters for successful ideation process become statistically insignificant. This is most likely an outcome of also assigning the parentage of successful idea submissions to prior unsuccessful submissions (through α_{21}). That is, after controlling for generation of successful ideas through prior failed attempts, the spatial effect of prior analogies on successful idea generation process is rendered statistically insignificant.

Our main model, the Mutually Exciting Spatio-Temporal Point Process model, includes the generative process for knowledge acquisition activity i.e. we also estimate $\alpha_{00}, \alpha_{01}, \alpha_{02}$ in addition to parameters in the Fully-Excited Knowledge Acquisition model. Through this model, we are able to observe the nuanced effect of prior analogies, in relation to their conceptual distance from the target problem, on the ideation outcomes. This model differs significantly in its estimated $\mu^{avg} = (1.1352, 0.2335, 0.0246)$ when compared



(a) (b)

Figure 3.4: We present the density plots for the effect of spatial distance on ideation efforts for the following baseline models: (a) Knowledge Acquisition model. (b) Self-Excited Knowledge Acquisition model.

with the Fully-Excited Knowledge Acquisition model's $\mu^{avg} = (2.9862, 0.1831, 0.0190)$. This difference is responsible for the statistical significance of $f_{20}^{spatial}(\cdot)$ parameters in the main model because it leads to significant changes in the parentage assignments (i.e. branching probabilities, Equation (3.9)) of the idea submission activities after modelling the generative process for knowledge acquisitions. Our ability in correctly inferring the true statistical parentage is essential in understanding the spatial properties of analogies which lead to successful and unsuccessful outcomes. Therefore, inclusion of all the observed processes of knowledge acquisition and its applications via idea submissions are necessary to understand the nuanced effects of conceptual distances on creativity outcomes.

In addition to the above benchmark analyses, we further compare our mutually exciting model framework with the conventional models belonging to the reduced-form framework. For the reduced-form analysis, similar to prior studies (e.g. Chan et al., 2015), we created a dataset using all the ideas whose authors had acquired knowledge through interactions on the platform before submitting their ideas. We measured the conceptual distance as the mean of angular distances between the stimulus ideas and the challenge of the focal idea submission. We ran a logistic regression model with fixed effects at the author level and used controls for author's tenure on the platform, number of prior stimulus ideas, success in earlier submissions, number of challenges participated in, and number of ideas submitted by the author prior to focal idea submission. The result of this analysis is presented in Table 3.8.

Table 3.8: Logistic Regression with Fixed Effects for Analogical Learning.

Covariate Name	Point Estimate
Intercept	-20.2618
	(1773.0360)
Conceptual Distance	0.8620
	(0.9046)
Time on Platform	-0.7845
	(0.4798)
Number of prior analogies	-0.7255^{*}
	(0.2835)
Number of prior submissions	-1.2586^{*}
	(0.5098)
I(Prior Success>0)	-1.3655^{***}
	(0.2588)
Prior Challenges Participated In	1.7683^{**}
	(0.5381)

Note: Standard Errors are reported under the parameter estimates in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In contrast to the nuanced results from our main model, the result from the logistic regression model indicates that there is no effect of conceptual distance on the likelihood of creative idea generation. The statistically insignificant effect in the reduced-form model partially supports our main results (as α_{20} is also statistically insignificant). However, it does not provide us with a deeper insight into the differentiating effects of conceptual distance on creative and non-creative outcomes. Moreover, reduced-form analysis focuses on statistical correlations in the data to showcase negative impact of prior activities (e.g. as user's time on the platform increases, number of future successes would reduce in finite data samples, which would result in estimation of spurious correlations). On the other hand, in the mutually exciting framework, the comparison of excitation coefficients with exogenous rates clearly distinguishes between ideas that are generated independently of the platform stimuli and ideas that are a resultant of it. This allows us to understand the generative processes behind event occurrences. Moreover, modelling structure imposed through the addition of spatial density functions elicits the nuanced effects of conceptual distance for different submission types which is not possible through the reduced-form framework.

3.3 Robustness Checks

We perform a variety of robustness checks to validate the different underlying assumptions and assess their impact on our final results.

3.3.1 Minimum User Tenure

In our main analysis, we set the minimum duration for the tenure of the user on the platform to one month. We use one month because it is the mean value for the duration of ideation phase across all challenges. We check the robustness of our results across different thresholds for minimum user tenure (i.e., 0.5 months and 1.5 months). We find that the estimated parameters remain consistent in their direction and statistical significance. The results of this robustness analysis are reported in Table 3.9.

Table 3.9: Estimates with Different Minimum User Tenure

Parameters	1 month	0.5 month	1.5 month
	Baseline Intensities		
μ_0^{avg}	1.1352^{***}	2.0236^{***}	0.7813^{***}
	(0.0077)	(0.0140)	(0.0054)
μ_1^{avg}	0.2335^{***}	0.4254^{***}	0.1612^{***}
	(0.0037)	(0.0068)	(0.0026)
μ_2^{avg}	0.0246^{***}	0.0425^{***}	0.0163^{***}
	(0.0012)	(0.0021)	(0.0008)
	Exciting Effects		
α_{00}	0.7146^{***}	0.7129^{***}	0.7206^{***}
	(0.0030)	(0.0030)	(0.0030)
α_{10}	0.0575^{***}	0.0550^{***}	0.0593^{***}
	(0.0032)	(0.0033)	(0.0032)
α_{20}	0.0768	0.1078	0.0642
	(0.0780)	(0.9250)	(1.0152)
α_{01}	0.2463^{***}	0.1932^{***}	0.2881^{***}
	(0.0061)	(0.0055)	(0.0066)
α_{11}	0.1791^{***}	0.1864^{***}	0.1753^{***}
	(0.0056)	(0.0060)	(0.0054)
α_{21}	0.0277^{***}	0.0278^{***}	0.0291^{***}
	(0.0023)	(0.0023)	(0.0024)
α_{02}	0.2106^{***}	0.1801^{***}	0.2169^{***}
	(0.0175)	(0.0162)	(0.0173)
α_{12}	0.0000	0.0000	0.0000

	(0.0001)	(0.0001)	(0.0001)
α_{22}	0.0912^{***}	0.0938^{***}	0.0877^{***}
	(0.0122)	(0.0125)	(0.0118)
	Temporal Distribution		
k_0	0.2624^{***}	0.2659^{***}	0.2582^{***}
	(0.0012)	(0.0012)	(0.0012)
k_1	0.1564^{***}	0.1510^{***}	0.1599^{***}
	(0.0034)	(0.0033)	(0.0034)
k_2	0.1213^{***}	0.1196^{***}	0.1213^{***}
	(0.0068)	(0.0069)	(0.0068)
θ_0	0.0745^{***}	0.0641^{***}	0.0880^{***}
	(0.0007)	(0.0006)	(0.0008)
θ_1	2.3016^{***}	2.6966^{***}	2.0891^{***}
	(0.1919)	(0.2348)	(0.1688)
θ_2	4.3432^{***}	4.3235^{***}	5.9528^{**}
	(1.2205)	(1.2693)	(1.8967)
	Spatial Distribution		
k_1'	5.1683^{***}	5.1248^{***}	5.1361^{***}
	(0.2337)	(0.2444)	(0.2249)
k_2'	1.5089^{*}	1.3953^{*}	1.4955^{*}
	(0.7273)	(0.7020)	(0.7077)
θ_1'	0.4005^{***}	0.3981^{***}	0.4015^{***}
	(0.0046)	(0.0050)	(0.0045)
θ_2'	0.1609^{*}	0.1389^{*}	0.1668^{*}
	(0.0771)	(0.0675)	(0.0802)

Note: Standard Errors are reported under the parameter estimates in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.3.2 Idea Submitting Users Only

In this subsection, we run the analysis using only those users who have submitted at least one idea for any challenge. Our results remain robust to this subset of users.

Table 3.10: Estimates based on only Idea Submitting Users

Parameters	Model: All Users	Model: Submitting Idea Users

	Baseline Intensities	
μ_0^{avg}	1.1352^{***}	0.6911^{***}
	(0.0077)	(0.0093)
μ_1^{avg}	0.2335^{***}	0.6473^{***}
	(0.0037)	(0.0103)
μ_2^{avg}	0.0246^{***}	0.0567^{***}
	(0.0012)	(0.0028)
	Exciting Effects	
α_{00}	0.7146^{***}	0.8326^{***}
	(0.0030)	(0.0042)
α_{10}	0.0575^{***}	0.0918^{***}
	(0.0032)	(0.0066)
α_{20}	0.0768	0.0246
	(0.0780)	(0.0384)
α_{01}	0.2463^{***}	0.3173^{***}
	(0.0061)	(0.0070)
α_{11}	0.1791^{***}	0.1593^{***}
	(0.0056)	(0.0052)
α_{21}	0.0277^{***}	0.0232^{***}
	(0.0023)	(0.0020)
α_{02}	0.2106^{***}	0.2490^{***}
	(0.0175)	(0.0190)
α_{12}	0.0000	0.0000
	(0.0001)	(0.0001)
α_{22}	0.0912^{***}	0.0817^{***}
	(0.0122)	(0.0111)
	Temporal Distribution	
k_0	0.2624^{***}	0.2528^{***}
	(0.0012)	(0.0014)
k_1	0.1564^{***}	0.1491^{***}
	(0.0034)	(0.0033)
k_2	0.1213^{***}	0.1135^{***}
	(0.0068)	(0.0069)
θ_0	0.0745^{***}	0.1383^{***}
	(0.0007)	(0.0016)
θ_1	2.3016^{***}	1.7924^{***}
	(0.1919)	(0.1433)
θ_2	4.3432^{***}	1.9135^{**}
	(1.2205)	(0.5350)
	Spatial Distribution	
k_1'	5.1683^{***}	4.6200^{***}
	(0.2337)	(0.2397)
k_2'	1.5089^{*}	1.6698^{*}

	(0.7273)	(0.7110)
θ_1'	0.4005^{***}	0.3842^{***}
	(0.0046)	(0.0061)
θ_2'	0.1609^{*}	0.1550^{*}
	(0.0771)	(0.0715)

Note: Standard Errors are reported under the parameter estimates in parenthesis. * p < 0.05, ** p < 0.01, *** p < 0.001

3.3.3 Stationarity of exogenous risk parameters

In the context of the Hawkes process, the counting process $N(t)$ has stationary increments if and only if the expected number of events generated through excitation is less than one (Bacry et al., 2015). Our estimated excitation/branching coefficients are less than one implying the stationarity of the Hawkes process. The stationarity of the exogenous risk, μ, implies that the expected number of events caused by the exogenous risk (occurring in any time interval) is a product of the duration of the interval and a constant rate. Therefore, if this were the case, the expected rate of events occurring in arbitrarily defined time intervals would be approximately constant. This also forms the basis for the diagnostic test for stationarity of the homogeneous Poisson process (Daley and Vere-Jones, 2003).

We follow the instructions in Daley and Vere-Jones (2003) for testing the stationarity of the exogenous risk μ. Our current dataset consists of users' trajectory data from time 0 to T (T is the timestamp of the last activity). We create new datasets where we only include user activities between time 0 to $\frac{T}{2}$ and 0 to $\frac{2*T}{3}$. We estimate the mutually exciting spatio-temporal point process model for these datasets.

If the expected baseline risks (for these datasets) are approximately equal to our main dataset with users' activities from 0 to T, then it implies that the baseline risk is stationary (Daley and Vere-Jones, 2003). We tabulate the estimated ratios in Table 3.11. We find that the exogenous risks are approximately equal and consistent, suggesting that the baseline risk is stationary in our setting.

Table 3.11: Baseline Rates for models with different time ranges

Time Range	μ_0^{avg}	μ_1^{avg}	μ_2^{avg}
0 to T	1.1352	0.2335	0.0246
0 to $\frac{T}{2}$	1.1317	0.2376	0.0241
0 to $\frac{2*T}{3}$	1.1360	0.2336	0.0244

3.3.4 Time Delay in Results

The ideas submitted by the platform participants are adjudged to be successful or not after a certain duration (as is common in tournament settings). For example, an ideator submits an idea at t_1. The results are declared at T. Therefore, it is hard to say whether events generated in $(t_1, T]$ are due to successful or unsuccessful ideation as the results are not known until T.

Nonetheless, with overlapping challenges, i.e., multiple challenges occurring around the same time, it does not seem possible to incorporate this delay in our mutually exciting model. Consequently, we set out to test whether our estimates would change drastically if we could model this time delay. Our strategy is as follows:

1. Assume that, instead of successful and unsuccessful ideation processes, there is only one process i.e., ideation process. In other words, our test model has two processes: knowledge acquisition and ideation processes.

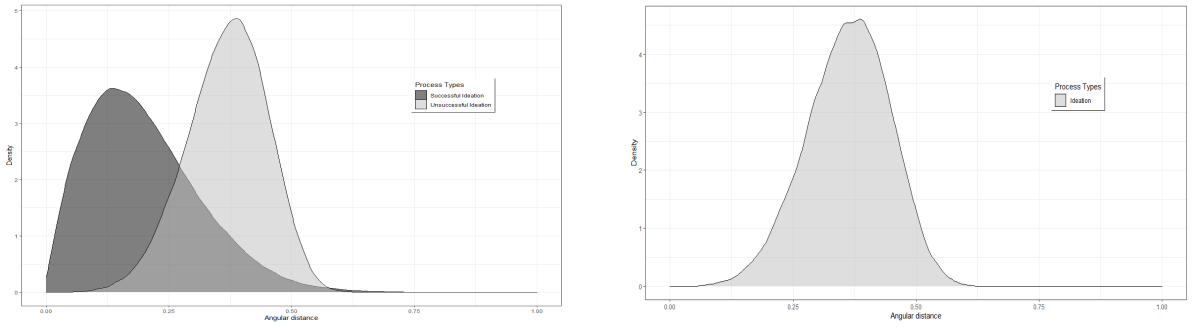
2. Estimate a mutually exciting spatio-temporal point process model on the same data for this simplified specification.

The results are presented in Table 3.12. From this analysis, we find that knowledge acquisition excites ideation activities, and the distribution of the spatial effect is some weighted average of the spatial effects of our main model with the knowledge acquisition process, unsuccessful ideation process, and successful ideation process. As 88% of all idea submissions are unsuccessful, we see that the estimated spatial distribution for the ideation process leans more towards the unsuccessful ideation process (Figure 3.5).

Table 3.12: Estimate for the model with two processes

Parameter Name	Estimate	Standard Error
α_{00}	0.7711	0.0031
α_{10}	0.0675	0.0038
α_{01}	0.3072	0.0065
α_{11}	0.2011	0.0057
k_0	0.2485	0.0011
k_1	0.1549	0.0031
θ_0	0.1396	0.0013
θ_1	2.2681	0.1847
k'	4.8760	0.2164
θ'	0.3944	0.0048
μ_0	0.6488	0.0046
μ_1	0.1570	0.0023

What this simplified model tells us is that knowledge acquisition has an effect on idea generation and the spatial effect has a mean value of 0.3616. Because the time delay is immaterial in this model (as there is no winning or losing status, only idea generation status), it demonstrates that irrespective of the time delay, the current results would hold. When we split the ideation process into successful and unsuccessful processes, we see more insight into the effect of conceptual distance on creativity outcomes. Though this is not a perfect test, it approximately shows that time delay would not drastically change our findings.



(a) (b)

Figure 3.5: We present the density plots for the effect of spatial distance on ideation efforts. (a) Model with three processes. (b) Model with two processes.

3.3.5 Negative Binomial Baseline Rates

In our main model, the baseline rates are restricted to follow the Poisson distribution. We relax this assumption to allow for over-dispersion. Specifically, we model the baseline rate of event occurrence as following the Negative Binomial distribution. This necessitates the need to derive the new conditional intensity function for the baseline rate. We present our derivation and the results below. We do not find any evidence of over-dispersion as it is captured by the Poisson cluster process which we model explicitly through our Hawkes process framework (Daley and Vere-Jones, 2003).

Derivation

For a general Point Process, we know

$$\mathbb{P}\left(N(t+h)-N(t)=m|\mathcal{H}_t\right)=\begin{cases}\lambda(t|\mathcal{H}_t)h+o(h), & m=1,\\ o(h), & m>1,\\ 1-\lambda(t|\mathcal{H}_t)h+o(h), & m=0,\end{cases} \tag{3.14}$$

Let $\lambda(t|\mathcal{H}_t,\nu)=\lambda(t|\mathcal{H}_t)\nu=\lambda\nu$ where ν is gamma distributed with mean 1 and variance $\frac{1}{\phi}$ i.e. shape = ϕ and scale = $\frac{1}{\phi}$ for the baseline process. Thus,

$$\begin{aligned}\mathbb{P}\left(N(t+h)-N(t)=1|\mathcal{H}_t\right)&=\frac{\mathbb{P}\left(N(t+h)-N(t)=1,\mathcal{H}_t\right)}{\mathbb{P}\left(\mathcal{H}_t\right)}\\&=\frac{\int\mathbb{P}\left(N(t+h)-N(t)=1|\mathcal{H}_t,\nu\right)\mathbb{P}\left(\mathcal{H}_t|\nu\right)g(\nu)d\nu}{\int\mathbb{P}\left(\mathcal{H}_t|\nu\right)g(\nu)d\nu}\\&=\frac{\int\lambda(t|\mathcal{H}_t,\nu)h\mathbb{P}\left(\mathcal{H}_t|\nu\right)g(\nu)d\nu}{\int\mathbb{P}\left(\mathcal{H}_t|\nu\right)g(\nu)d\nu}\\&=\frac{\int\lambda\nu h\mathbb{P}\left(\mathcal{H}_t|\nu\right)g(\nu)d\nu}{\int\mathbb{P}\left(\mathcal{H}_t|\nu\right)g(\nu)d\nu}\end{aligned}$$

Now, conditional on ν, $\mathbb{P}(\mathcal{H}_t)$ follows Poissonian distribution. Hence,

$$
\begin{aligned}
\mathbb{P}\left(N(t+h)-N(t)=1|\mathcal{H}_t\right) &= \frac{\int \lambda\nu h\mathbb{P}\left(\mathcal{H}_t|\nu\right)g(\nu)d\nu}{\int \mathbb{P}\left(\mathcal{H}_t|\nu\right)g(\nu)d\nu} \\
&= \frac{\int \lambda\nu h\frac{e^{-\lambda\nu}\lambda\nu^{N(t^-)}}{N(t^-!)}\frac{\nu^{\phi-1}e^{-\nu\phi}\phi^{\phi}}{\Gamma(\phi)}d\nu}{\int \frac{e^{-\lambda\nu}\lambda\nu^{N(t^-)}}{N(t^-!)}\frac{\nu^{\phi-1}e^{-\nu\phi}\phi^{\phi}}{\Gamma(\phi)}d\nu} \\
&= \frac{\int_0^\infty \lambda h\nu^{N(t^-)+\phi}e^{-\nu(\lambda+\phi)}d\nu}{\int_0^\infty \nu^{N(t^-)+\phi-1}e^{-\nu(\lambda+\phi)}d\nu} \\
&= \lambda h\left(\frac{\Gamma(N(t^-)+\phi+1)(\lambda+\phi)^{N(t^-)+\phi-2}}{(\lambda+\phi)^{N(t^-)+\phi-1}\Gamma(N(t^-)+\phi)}\right) \\
&= \lambda h\left(\frac{N(t^-)+\phi}{\lambda+\phi}\right) \qquad (3.15)
\end{aligned}
$$

Thus,

$$
\lambda(t|\mathcal{H}_t) = \lambda\left(\frac{N(t^-)+\phi}{\lambda+\phi}\right) \qquad (3.16)
$$

Results

We estimate a model using these conditional intensity function with baseline rates following the Negative Binomial distribution. The results are presented in Table 3.13. The results indicate no evidence for the presence of over-dispersion in the baseline intensities. The over-dispersion in the data is captured by the Poisson cluster process which we model explicitly through our Hawkes process framework (Daley and Vere-Jones, 2003).

3.3.6 Usage of Fixed Effects instead of Random Effects to model user heterogeneity

In this section, we develop a model of idea generation where unobserved user heterogeneity is included via fixed effects. We define the multivariate Hawkes model, with user-level fixed effects ν^i, as follows:

$$
\lambda_k^i(t|\mathcal{H}_t^i,\nu^i) = \mu_k^i + \sum_{j=1}^{D}\sum_{m:t_m^j<t}\phi_{kj}(t-t_m^j,\mathbf{c}_t-\mathbf{x}_m^j) \text{ where } \mu_k^i = \mu_k\nu^i
$$

Table 3.13: Parameter Estimates of a Mutually Exciting Spatio-Temporal Point Process Model with Negative Binomial Baseline Rates

Process Types	Interaction	Unsuccessful Ideation	Successful Ideation
	Baseline Intensities		
	μ_0	μ_1	μ_2
	0.6489^{***}	0.1423^{***}	0.0145^{***}
	(0.0046)	(0.0022)	(0.0007)
	ϕ_0	ϕ_1	ϕ_2
	626647.2234	185992.0396	4159.1421
	(8187392.4911)	(987900.6560)	(28449.3428)
	Exciting Effects		
Interaction	α_{00}	α_{01}	α_{02}
	0.7711^{***}	0.3036^{***}	0.3445^{***}
	(0.0032)	(0.0068)	(0.0224)
Unsuccessful Ideation	α_{10}	α_{11}	α_{12}
	0.0576^{***}	0.1839^{***}	0.0000
	(0.0032)	(0.0057)	(0.0001)
Successful Ideation	α_{20}	α_{21}	α_{22}
	0.0361	0.0309^{***}	0.0970^{***}
	(0.0356)	(0.0025)	(0.0129)
	Temporal Distribution		
Shape	k_0	k_1	k_2
	0.2485^{***}	0.1600^{***}	0.1281^{***}
	(0.0011)	(0.0034)	(0.0069)
Scale	θ_0	θ_1	θ_2
	0.1396^{***}	2.0533^{***}	6.9819^{**}
	(0.0013)	(0.1739)	(2.2051)
	Spatial Distribution		
Shape	k_0'	k_1'	k_2'
	n.a.	5.1696^{***}	1.7681^{*}
	n.a.	(0.2305)	(0.6895)
Scale	θ_0'	θ_1'	θ_2'
	n.a.	0.4003^{***}	0.2202^{*}
	n.a.	(0.0046)	(0.0931)

Note: Standard Errors are reported under the parameter estimates in parenthesis. Since we did not estimate the effect of spatial distribution over generation of new comments, we have placed n.a. for the same.

$$\phi_{kj}(t-t_m^j, \mathbf{c}_t - \mathbf{x}_m^j) = \begin{cases} \alpha_{kj} f_k^{temporal}\left(t-t_m^j\right) f_{kj}^{spatial}\left(|\mathbf{c}_t - \mathbf{x}_m^j|_\theta\right), & \text{if } t \geq t_m^j \text{ and } (k,j) \in \{(1,0), \\ & (2,0)\} \\ \alpha_{kj} f_k^{temporal}\left(t-t_m^j\right), & \text{if } t \geq t_m^j \text{ and } (k,j) \in \{(0,0), \\ & (0,1),(0,2),(1,1),(1,2), \\ & (2,1),(2,2)\} \\ 0, & \text{otherwise} \end{cases} \tag{3.17}$$

ν^i are the N dummy variables for fixed effects. N is the number of users. We set $\nu^1 = 1$ for identification purposes.

Estimation

We use the EM algorithm to estimate the model parameters. In the Expectation step, we compute the posterior branching probabilities for the r^{th} iteration (i.e. probability that the event of process k at time t_n is generated by any of the prior events or is the result of exogenous risk)

$$\text{Prob}(Z_{kn} = z \mid \mathcal{H}_t, \theta^{(r)}) = \frac{\lambda_z^{(r)}(t_{kn})}{\sum_{z' \in z_k} \lambda_{z'}^{(r)}(t_{kn})} \tag{3.18}$$

In the maximization step, we take the posterior probabilities as given and maximize the expected complete data log-likelihood.

$$Q(\theta|\theta^{(r)}, \mathcal{H}_t) = \sum_{users}\left(\sum_{k=1}^{D}\left(\sum_{z\in z_k}\left(\sum_{n=1}^{N_k}\log(\lambda_z(t_{kn})) \times \text{Prob}(Z_{kn} = z \mid \mathcal{H}_t, \theta^{(r)}) - \int_0^T \lambda_z dt\right)\right)\right) \tag{3.19}$$

Equation (3.19) has closed form solutions for α_{kj}. For ν^i, we derive its optimal value in terms of μ_k as follows:

$$\nu^i = \frac{\sum_{k=1}^{D}\sum_{n=1}^{N_k^i} \text{Prob}(Z_{kn} = z \mid \mathcal{H}_t^i, \theta^{(r)})}{\sum_{k=1}^{D} \mu_k T} \tag{3.20}$$

i denotes the user. D is the number of processes under study. Therefore, we maximize (3.19) subject to the restriction in (3.20).

Results

Our results remain robust to usage of fixed effects for unobserved user heterogeneity (see Table 3.14). While non-parametric random effects allow us to model multi-dimensional heterogeneity (i.e. heterogeneities in α_{kj} along with μ_k), fixed effects only allow for heterogeneity in the exogenous risk parameters μ_k.

Table 3.14: Estimates for Fixed Effects

Parameters	Model: Random Effects	Model: Fixed Effects
	Baseline Intensities	
μ_0^{avg}	1.1352^{***}	0.8494^{***}
	(0.0077)	(0.0090)
μ_1^{avg}	0.2335^{***}	0.1432^{***}
	(0.0037)	(0.0018)
μ_2^{avg}	0.0246^{***}	0.0148^{***}
	(0.0012)	(0.0007)
	Exciting Effects	
α_{00}	0.7146^{***}	0.6728^{***}
	(0.0030)	(0.0029)
α_{10}	0.0575^{***}	0.0426^{***}
	(0.0032)	(0.0028)
α_{20}	0.0768	0.0303
	(0.0780)	(0.0295)
α_{01}	0.2463^{***}	0.1559^{***}
	(0.0061)	(0.0049)
α_{11}	0.1791^{***}	0.1682^{***}
	(0.0056)	(0.0053)
α_{21}	0.0277^{***}	0.0273^{***}
	(0.0023)	(0.0023)
α_{02}	0.2106^{***}	0.1461^{***}
	(0.0175)	(0.0146)
α_{12}	0.0000	0.0000
	(0.0001)	(0.0001)
α_{22}	0.0912^{***}	0.0914^{***}
	(0.0122)	(0.0122)
	Temporal Distribution	
k_0	0.2624^{***}	0.2752^{***}

	(0.0012)	(0.0013)
k_1	0.1564^{***}	0.1420^{***}
	(0.0034)	(0.0030)
k_2	0.1213^{***}	0.1160^{***}
	(0.0068)	(0.0030)
θ_0	0.0745^{***}	0.0445^{***}
	(0.0007)	(0.0004)
θ_1	2.3016^{***}	1.5360^{***}
	(0.1919)	(0.1476)
θ_2	4.3432^{***}	4.2428^{**}
	(1.2205)	(1.3639)
	Spatial Distribution	
k_1'	5.1683^{***}	5.1958^{***}
	(0.2337)	(0.2691)
k_2'	1.5089^{*}	2.0167^{**}
	(0.7273)	(0.7043)
θ_1'	0.4005^{***}	0.3961^{***}
	(0.0046)	(0.0053)
θ_2'	0.1609^{*}	0.2225^{**}
	(0.0771)	(0.0784)

Note: Standard Errors are reported under the parameter estimates in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 4

STRUCTURAL ANALYSIS OF COMPETITION IN BITCOIN AND BITCOIN-CASH MINING

4.1 Model

We consider a theoretical model of I forward-looking mining pools, subscripted as i. Time is discrete and a single time period is an interval of 6 hours. At the beginning of every time period t, miners make decisions regarding their willingness to participate in mining BTC or BCH or neither or both. As such, every miner draws a private shock, at the beginning of t, regarding mining feasibility of BTC and BCH. We model their choices as a partial equilibrium, dynamic decision problem. In our setup, miners decide on the number of blocks of each chain to mine[1]: that is, $n_{i,t,btc} \in \{0, ..., K_1\}$ and $n_{i,t,bch} \in \{0, ..., K_2\}$. We use the terms miner and mining pool[2] interchangeably.

While the actual decision-making process within the mining pool could be complicated and opaque, we make simplifying assumptions regarding the timing of the process. First, we assume that decision about BTC is made prior to BCH by every mining pool[3]. Second, we assume that coins earned in time t are sold in the same time period at the prevailing market price[4]. Third and most importantly, we assume that BTC and BCH mining represent two different dynamic games. That is, we assume that there are two separate decision-makers for BTC and BCH within every mining pool. And, BCH decision-maker takes the decisions of the BTC decision-maker as given. Our first assumption allows us to separate these two mining operations

[1]Miners decide on the allocation of hash power or computational power to each blockchain. We do not observe these hash power shares. Since, number of mined blocks are strongly correlated with quantum of hash power used, we posit the decision-making problem from the perspective of number of blocks. The structural errors allow us to rationalize the inherent mining stochasticity.

[2]Mining pools are comprised of individual miners who have different preferences over BTC and BCH. We abstract away from their individuality and treat the outcomes of their collective preferences as the overall decisions of the mining pool.

[3]Given the dominance and popularity of BTC, it is very probable.

[4]Ideally, miners are required to wait until 100 new blocks have been added to the chain after their block before they can monetize their coinbase rewards and transaction fees by selling them on the secondary market (coin exchanges). Since, we do not know the actual time of their sale (it cannot be traced on the blockchain either), we assume that they get the current time period's market price.

and treat them as separate mining games. For miner i, we refer to the BTC decision-maker as BTC miner i and BCH decision-maker as BCH miner i.

In our setup, miners are forward-looking i.e. their goal is to maximize expected discounted profits from mining BTC and BCH at time t. BTC miner i decides on the number of BTC, $n_{i,t,btc}$, to mine to maximize its expected discounted profits from mining BTC. In a similar fashion, but with separate function for BCH profits, BCH miners makes a decision about BCH blocks, $n_{i,t,bch}$ while taking its BTC counterpart decision, $n_{i,t,btc}$, as *exogenous*. Miner incurs computational costs while mining these blocks. Inclusion of coinbase reward and transaction fees brings revenue to the miner conditional on the market prices. Let $\chi_{it,btc}$ be the set of payoff-relevant BTC's state variables for miner i at time t, i.e., $\chi_{it,btc} = \left(s^R_{it,btc}, s^C_{it,btc}\right)$. We define $\chi_{it,bch}$ similarly. $s^R_{it,btc}$ is the set of BTC state variables which affect the revenue accrual to the miner such as BTC's market prices and prevailing transaction fees levels. Likewise, $s^R_{it,bch}$ is the set of BCH state variables. Let p denote price and f denote transaction fees. $s^C_{it,btc}$ is the set of variables affecting the cost such as difficult levels, denoted by d_{btc}. $s^C_{it,bch}$ includes BCH's difficulty levels, d_{bch} and the BTC miner's decision, $n_{i,t,btc}$. $\epsilon_{i,t,btc}$ is a privately known structural error, independent and identically distributed, and drawn from a known distribution, $F_{btc}(\epsilon_i|\chi_{it,btc})$ with support $\epsilon_i \subset \mathbb{R}$. It represents miner i's private information about BTC mining, such as hash power contribution of pool members toward BTC or information regarding profit distribution among pool members or the inherent randomness in mining. $\epsilon_{i,t,bch}$ is the miner i's private information about the share of hash power allocated to BCH mining and the inherent randomness in mining. $\epsilon_{i,t,bch}$ is drawn from $F_{bch}(\epsilon_i|\chi_{it,bch})$. The reduced-form one-shot payoff functions for miner i at time t is given by:

$$\begin{aligned}
\pi^{btc}\left(n_{i,t,btc}, n_{-i,t,btc}, \chi_{it,btc}, \epsilon_{i,t,btc}\right) &= R(n_{i,t,btc}, n_{-i,t,btc}, s^R_{it,btc}) - C(n_{i,t,btc}, n_{-i,t,btc}, s^C_{it,btc}) \\
&\quad + \epsilon_{i,t,btc} \\
\pi^{bch}\left(n_{i,t,bch}, n_{-i,t,bch}, \chi_{it,bch}, \epsilon_{i,t,bch}\right) &= R(n_{i,t,bch}, n_{-i,t,bch}, s^R_{it,bch}) - C(n_{i,t,bch}, n_{-i,t,bch}, s^C_{it,bch}) \\
&\quad + \epsilon_{i,t,bch}
\end{aligned} \tag{4.1}$$

Miners' per-period profit (4.1) comprises of revenues and costs associated with mining BTC and BCH.

$R(\cdot)$ is the revenue function and $C(\cdot)$ is the cost function. The per-period payoff, from BTC and BCH, is parametrically defined by equations (4.2) and (4.3) respectively. These payoffs have very similar specifications. The profit from mining on each chain increases linearly in the number of mined blocks as each mined block carries significant coinbase rewards. $\theta_3^{btc} n_{i,t,btc}$ and $\theta_3^{bch} n_{i,t,bch}$ represents the revenue accrued by BTC and BCH miners respectively through coinbase rewards. Transaction fees incentivize miners further, in addition to coinbase rewards, to allocate precious mining power and enter additively into their respective profit functions as $\theta_2^{btc} f_{t,btc}$ and $\theta_2^{bch} f_{t,bch}$. Miner gains $\theta_1^{btc} f_{t,btc}$ and $\theta_1^{bch} f_{t,bch}$ due to exogenous movements in market prices of BTC and BCH respectively. The erosion of profits due to competition is quantified by $\theta_5^{btc} n_{-i,t,btc}$ and $\theta_5^{bch} n_{-i,t,bch}$ where $n_{-i,t,btc}$ and $n_{-i,t,bch}$ are the number of BTC and BCH blocks mined by competitors of miner i at time t. The expenses related to mining are quantified by $\theta_4^{btc} d_{t,btc}$ and $\theta_4^{bch} d_{t,bch}$. Moreover, the BCH miner further faces competition from its BTC counterpart for the limited hash power owned by their mining pool. Its profits are further eroded by $\theta_6^{bch} n_{i,t,btc} + \theta_7^{bch} n_{i,t,btc}^2$.

$$\pi_{i,t}^{btc}\left(n_{i,t,btc}, n_{-i,t,btc}, \chi_{it,btc}; \theta^{btc}\right) = [\theta_1^{btc} p_{t,btc} + \theta_2^{btc} f_{t,btc} + \theta_3^{btc} n_{i,t,btc} - \theta_4^{btc} d_{t,btc} - \theta_5^{btc} n_{-i,t,btc} + \epsilon_{i,t,btc}].1\left\{n_{i,t,btc} > 0\right\} \quad (4.2)$$

$$\pi_{i,t}^{bch}\left(n_{i,t,bch}, n_{-i,t,bch}, n_{i,t,btc}, \chi_{it,bch}; \theta^{bch}\right) = [\theta_1^{bch} p_{t,bch} + \theta_2^{bch} f_{t,bch} + \theta_3^{bch} n_{i,t,bch} - \theta_4^{bch} d_{t,bch} - \theta_5^{bch} n_{-i,t,bch} - \theta_6^{bch} n_{i,t,btc} - \theta_7^{bch} n_{i,t,bch}^2 + \epsilon_{i,t,bch}].1\left\{n_{i,t,bch} > 0\right\} \quad (4.3)$$

Thus, the structural parameters defined in our model for the BTC miner are $\theta^{btc} = \left(\theta_1^{btc}, \theta_2^{btc}, \theta_3^{btc}, \theta_4^{btc}, \theta_5^{btc}\right)$. BCH miner's structural parameters are defined by the vector $\theta^{bch} = \left(\theta_1^{bch}, \theta_2^{bch}, \theta_3^{bch}, \theta_4^{bch}, \theta_5^{bch}, \theta_6^{bch}, \theta_7^{bch}\right)$.

We study the mining game being played by BTC miners from April 2017 to February 2018. We split this period into three sub-periods: (a) pre-fork period: BTC wasn't forked in this period and, therefore, BCH did not exist; (b) EDA-active period: BTC was forked on 1st August 2017, and immediately after the fork, BCH developers instituted the emergency difficulty adjustment algorithm (EDA) which was active

until 13th November 2017 ; (c) Non-EDA period: covers the time from end of EDA to February 2018. These sub-periods were characterized by very different equilibria due to (a) Miners focused their attention only on BTC in the pre-fork period as BTC fork hadn't occurred; (b) presence of EDA algorithm in the BCH chain during EDA-active period, (c) BCH switched to a new difficulty adjustment algorithm[5] during the Non-EDA period, and (d) Non-EDA period witnessed the cryptocurrency boom leading to massive overheating of the cryptocurrency market. The Non-EDA period did not see any wild fluctuation in the BCH's rate of mining and hash power contributions. As such, we compute the equilibria governing these periods separately.

In the BTC and BCH mining process, difficulty changes and transactions fee levels are a function of the collective mining rate in the recent time periods. Higher rate of mining leads to upward adjustment in the difficulty levels and vice versa to maintain a constant rate of mining (one block every 10 minutes). Introduction of EDA led to one-way downward difficulty adjustments on the BCH chain during the time EDA mechanism was active. Fees incentive miners to process transactions during periods of high demand Houy, 2014; Easley et al., 2019. The rate of mining increases to cater to the increased demand for blocks and as a consequent, more transactions are processed. This mechanism allows users to increase the priority of their transactions. Once the excess demand subsides (as increased rate of mining eventually catches up), the fee levels decrease in response. This endogeneity of transaction fees is an important component of our structural model and has not been considered in earlier theoretical studies of Bitcoin mining (Ma et al., 2018; Dimitri, 2017).

We explain our forward-looking model for BTC miner only. The model and its explanation is exactly the same for BCH miner. Given an initial state $\chi_{it,btc}$ at time t, BTC miner i's expected discounted profits, before the realization of private shock, $\epsilon_{i,t,btc}$, are given by

$$\mathbb{E}\left[\sum_{\tau=t}^{\infty}\beta^{\tau-t}\pi_i^{btc}\left(n_{i,\tau,btc},n_{-i,\tau,btc},\chi_{i\tau,btc},\epsilon_{i,\tau,btc}\right)\middle|\chi_{it,btc}\right], \tag{4.4}$$

where $\beta \in [0,1)$ is the discount factor. The goal of BTC miner i is to maximize its expected discounted profits, taking the actions of rival BTC miners as given. The expectation in (4.4) is taken over i's private

[5]New DAA smoothed the future estimates of difficulty levels by using a simple moving average of the hash power contributed to the BCH blockchain for the previous 144 blocks to avoid erratic block generation times.

shock and other BTC miners' actions in time period t as well as over the future values of BTC-relevant state variables, private shocks, and BTC miners' actions.

We focus on pure strategy Markov Perfect Equilibrium (MPE) in specifying the equilibrium behavior. BTC Miners follow anonymous, symmetric, and Markovian strategies i.e. miners' actions depend only on the current state vector and their individual private shocks. Let $S_{i,btc}$ be the state space of all possible values of $\chi_{it,btc}$ for miner i and S_{btc} be the collection of $S_{i,btc}$. Let $n_{i,t,btc} \in N_{i,btc} \subset \mathbb{N}^0$. Also, $\epsilon_t \subset \mathbb{R}$ is the collection of i.i.d. private shocks $\epsilon_{i,t,btc}$ for all BTC miners at time t. The Markov strategy profile for BTC miner i can be characterized as a function $\sigma_i^{btc} : S_{btc} \times \mathbb{R} \to N_{i,btc}$ that maps the BTC payoff-relevant state variables and private information to the set of possible choices. The profile of Markov strategies is denoted by $\sigma^{btc} = \left(\sigma_1^{btc}, ..., \sigma_I^{btc}\right)$.

Given the Markov profiles σ_{btc}, expected profit of BTC miner i in state χ_{btc} (dropping the time subscripts) is defined recursively as

$$\mathbb{V}_i^{btc}\left(\chi_{btc}; \sigma^{btc}\right) = \mathbb{E}_\epsilon\left[\pi_i^{btc}\left(\sigma(\chi_{btc}, \epsilon_{btc}), \chi_{i,btc}, \epsilon_{i,btc}\right) + \beta\mathbb{E}\left[\mathbb{V}_i^{btc}(\chi'_{btc}; \sigma^{btc})|\chi_{btc}, \sigma(\chi_{btc}, \epsilon_{btc})\right]|\chi_{btc}\right] \tag{4.5}$$

$\mathbb{V}_i^{btc}$ is the BTC miner i's ex-ante value function. It reflects expected profits from mining BTC before the private shocks are drawn. $\sigma^{btc}(\chi_{btc}, \epsilon_{btc})$ is the action profile $\left(\sigma_1^{btc}(\chi_{btc}, \epsilon_{1,btc}), ..., \sigma_I^{btc}(\chi_{btc}, \epsilon_{I,btc})\right)$. The inner expectation is with respect to future value of state variable, χ'_{btc}, conditional on the current state χ_{btc} and the action of all the BTC miners in current time period. The outer expectation is over the current values of the private shocks, ϵ_{btc}.

An action profile σ^{btc} is a MPE if BTC miner i has no incentive to unilaterally deviate from its strategy σ_i^{btc} while the opponent miners are playing by σ_{-i}^{btc}. This implies that there is no alternative Markov strategy $\widetilde{\sigma}_i^{btc}$ that will yield a higher ex-ante value function than σ_i^{btc}. Therefore, σ^{btc} is a MPE if for all BTC miners

i, states χ_{btc}, and alternative Markovian strategies $\widetilde{\sigma_i}^{btc}$,

$$\begin{aligned}\mathbb{V}_i^{btc}(\chi_{btc};\sigma_i^{btc},\sigma_{-i}^{btc}) &\geq \mathbb{V}_i^{btc}(\chi_{btc};\widetilde{\sigma_i}^{btc},\sigma_{-i}^{btc}) \\ &= \mathbb{E}_\epsilon\big[\pi_i^{btc}\left(\widetilde{\sigma_i}^{btc}(\chi_{btc},\epsilon_{i,btc}),\sigma_{-i}^{btc}(\chi_{btc},\epsilon_{-i,btc}),\chi_{i,btc},\epsilon_{i,btc}\right)+ \\ &\quad \beta\mathbb{E}\left[\mathbb{V}_i^{btc}(\chi'_{btc};\widetilde{\sigma_i}^{btc},\sigma_{-i}^{btc})|\chi_{btc},\widetilde{\sigma_i}^{btc}(\chi_{btc},\epsilon_{i,btc}),\sigma_{-i}^{btc}(\chi_{btc},\epsilon_{-i,btc})\right]|\chi_{btc}\big]\end{aligned} \tag{4.6}$$

The forward-looking model for BCH miner is defined in an equivalent fashion. Though studies have proven the existence of MPE in general Markovian games (Dutta and Sundaram, 1998), however, little is known about equilibrium existence in a dynamic game with continuous states as ours. As such, we assume that an MPE exists and the same MPE is expected to be played in all time periods (e.g., Bajari et al., 2008; Blevins et al., 2017).

4.1.1 Theoretical Understanding of Competition in Mining

We define a theoretical equilibrium mining model for any coin, whether BTC or BCH. We consider N identical miners who choose their rate of computations (or hash power allocations), $\lambda_i \in R^+$ in every time period t. The difficulty of mining is set to $D \in \mathbb{N}^0$. The difficulty adjusts to maintain a time interval of δ between two blocks on average. Let $X_i(t)$ denote the Poisson distributed random variable for the number of computations done by miner i at time t (see Altman et al., 2019; Ma et al., 2018). In expectation, $\frac{D \times 2^{32}}{\delta}$ computations[6] are required to find the hash below the target value. Since finding the hash is a stochastic process and independent of past computations done by miners, the inter-arrival time between two successive computations on the network is a random variable following the Exponential distribution with rate parameter, $\sum_i^N \lambda_i$. Therefore,

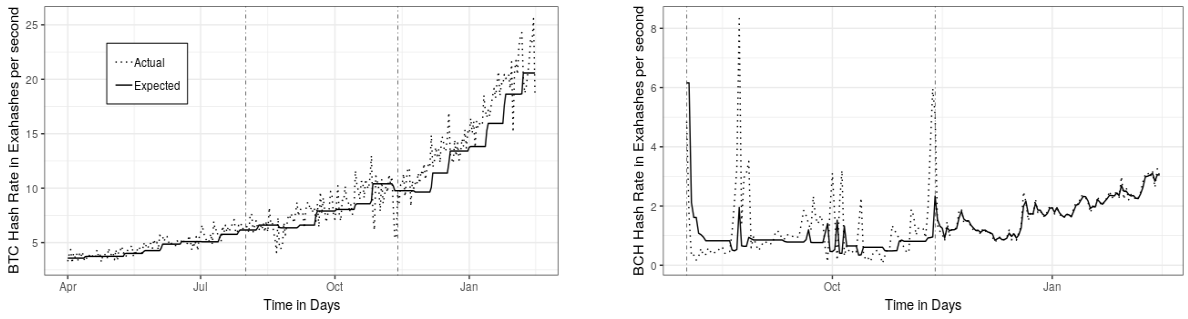
$$\sum_i^N \lambda_i = \frac{D \times 2^{32}}{\delta} \tag{4.7}$$

In equilibrium, every miner chooses the same λ. Let $M = D \times 2^{32}$.

$$\lambda = \frac{M}{N \times \delta} \tag{4.8}$$

[6]https://en.bitcoin.it/wiki/Difficulty

We plot the expected total network hash rates (according to our result, $\frac{M}{\delta}$) with the observed total network hash rate for BTC and BCH. We find that our theoretical predictions for total network hash rate closes matches with the observed data. However, we do find drastic, though expected (because of EDA-triggered difficulty reductions), deviations from our predictions during the EDA-active period for BCH. We also find consistent (yet, expected) deviations during the Non-EDA period for BTC (period characterized by cryptocurrency boom).



(a) (b)

Figure 4.1: Expected vs Observed Total Hash Power on the Bitcoin and Bitcoin-Cash blockchains. Vertical lines mark the time of BTC fork and EDA deactivation. (a) Bitcoin network hash power. (b) Bitcoin-Cash network hash power.

Therefore, each miner has an equal probability of winning the block generation competition, which is given by $\frac{1}{N}$. Block generation brings a total reward of R (including fees and coinbase) coins. Further, every miner faces the same cost function, $c(\lambda_i)$ which is strictly increasing and convex. Hence, miner i's expected payoff, $U_i(\lambda_i)$ is

$$\begin{aligned} U_i(\lambda_i) &= \frac{R}{N} - c(\lambda_i) \\ &= \frac{R}{N} - c(\lambda) \\ &= \frac{R}{N} - c\left(\frac{M}{N \times \delta}\right) \end{aligned} \tag{4.9}$$

Through derivatives, we find that the expected utility is increasing in N until a threshold value of $N^* =$

$\frac{M}{\delta(c^{-1})'(\frac{R\delta}{M})}$ where $(c^{-1})'(\cdot)$ is the first derivative of the inverse of $c(\cdot)$. Drastic changes in difficulty levels or coin prices, such as those observed for BCH during EDA-active periods or BTC price boom, can impact this competitive threshold. Our empirical analysis will allow us to understand our theoretical model's results with the help of actual mining data.

4.2 Empirical Strategy

Searching for an equilibrium in dynamic games is computationally prohibitive and mathematically intractable. Maximum likelihood approaches to estimate dynamic game parameters are computationally very demanding as the estimation procedure needs to solve for an equilibrium at every guess of parameters when even solving for an equilibrium for a single set of candidate values is computationally prohibitive. Moreover, existence of multiple equilibria complicates the procedure further as the econometrician is required to compute the set of all equilibria and then, match these equilibria to the data (Ryan, 2012).

We follow the two-step estimation approach to alleviate these computational issues. These methods do not require one to solve the theoretical equilibrium model explicitly even once (Hotz et al., 1994). The two-step empirical strategy posited in Bajari et al. (2007), henceforth referred to as BBL, is our preferred method to solve the miners' dynamic decision-making problem. In the first step, we compute the reduced-form estimates of the miners' policy functions. We do not impose any structure here and estimate the reduced-form regressions with flexible specifications. We also recover the probability distribution governing the evolution of market prices and transaction fees using standard time series methods. We explicitly include the dependence of transactions fees on the past rate of mining in the evolution of the same. This avoids computing the equilibrium of the game because the policy functions are estimated from the actual equilibrium that is played in the data. Therefore, by estimating the probability distributions for actions and states, we ultimately recover the miners' equilibrium beliefs because they are expected to have correct opinions about their environment and competitors in an equilibrium model. We simulate the ex-ante value functions using the parameters recovered in first stage of BBL. Using these simulated value functions, we impose optimality conditions (4.6) to recover the structural parameters through a minimum distance estimation routine in the second step. We assume that the data follows a first-order Markov process, are generated through a

single MPE, and all the miners expect the same equilibrium to be played in all the time periods of a single duration[7].

4.2.1 Data Sources

The data for our study has been generously provided by BlockChair (Blockchair, 2018). The data consists of the entire BTC and BCH blockchains right from their genesis blocks (the very first block in the blockchain). This includes information such as timestamp of each block, name of the successful miner, number of transactions in the block, total transaction fees, miner reward, and level of difficulty. We collected price-level and trading data from coinAPI.io (CoinAPI, 2018). Using their services, we aggregated price data from multiple exchanges such as Bitfinex, CoinBase, OKCoin, Cexio, and Kraken. We aggregated data at 6 hour intervals i.e. the number of coins mined by our selected mining pools in every 6 hour interval along with information about market prices, difficulty levels, and transaction fees.

4.2.2 First-Stage Estimation

Let ϕ be the vector of all first stage parameters. ϕ includes the reduced-from policy function coefficients and the coefficients of the transition functions for the price and transaction fees. The estimated coefficients in ϕ are tabulated in the appendix.

Policy Functions

We estimate the policy functions dictating the number of blocks to mine on the BTC and BCH blockchains. We use log-linear models to correlate the observed states to the number of miner blocks. Specifically, we model $\log(n_{i,t,btc}+1)$ and $\log(n_{i,t,bch}+1)$ as a linear combination of ϕ_{btc} and ϕ_{bch} respectively. Let $W_{i,t,btc}$ and $W_{i,t,bch}$ be the vectors of BTC's and BCH's policy-relevant state variables. These vectors consists of their respective prices, difficulty levels, transaction fees, lagged decisions for both chains, EDA-active

[7]We compute equilibria for three different time durations: pre-fork, EDA-active, and Non-EDA.

period, and some interactions of the above. We also include miner fixed effects.

$$\begin{aligned} \log(n_{i,t,btc}+1) &= \phi_{btc} W_{i,t,btc} + \epsilon_{i,t,btc} \\ \log(n_{i,t,bch}+1) &= \phi_{bch} W_{i,t,bch} + \epsilon_{i,t,bch} \end{aligned} \tag{4.10}$$

We try different specifications of the vectors $W_{i,t,btc}$ and $W_{i,t,bch}$ and select the best fit based on R^2 values.

Transition Equations

We estimate the transition equations for the state variables in χ_t, namely, $p_{t,btc}$, $p_{t,bch}$, $f_{t,btc}$, $f_{t,bch}$ and an indicator variable denoting whether EDA was active in that time period or not. We specifically model the evolution of $\log(x_t/x_{t-1})$ where $x \in \{p_{btc}, p_{bch}, f_{btc}, f_{bch}\}$. Let us denote $\log(x_t/x_{t-1})$ as $\Delta(x_t)$. $c_{t,btc}$ and $c_{t,bch}$ is the total number of BTC and BCH blocks mined at time t respectively. Evolution of EDA is deterministic. EDA was activated with the BTC fork on August 1^{st} 2017 and deactivated on November 13^{th} 2017.

$$\Delta(p_{t,btc}) = \zeta_{0,0}\Delta(p_{t-1,btc}) + \zeta_{0,1}\nu_{0,t-1} + \nu_{0,t}$$

$$\nu_{0,t} \sim (i.i.d.)\mathcal{N}\left(0, \kappa_0^2\right) \tag{4.11}$$

$$\Delta(p_{t,bch}) = \zeta_{1,0}\Delta(p_{t-1,bch}) + \zeta_{1,1}\nu_{1,t-1} + \nu_{1,t}$$

$$\nu_{1,t} \sim (i.i.d.)\mathcal{N}\left(0, \kappa_1^2\right) \tag{4.12}$$

$$\Delta(f_{t,btc}) = \zeta_{2,0}\Delta(f_{t-1,btc}) + \zeta_{2,1}\nu_{2,t-1} + \zeta_{2,2}(c_{t-1,btc} - c_{t-2,btc}) + \nu_{2,t}$$

$$\nu_{2,t} \sim (i.i.d.)\mathcal{N}\left(0, \kappa_2^2\right) \tag{4.13}$$

$$\Delta(f_{t,bch}) = \zeta_{3,0}\Delta(f_{t-1,bch}) + \zeta_{3,1}\nu_{3,t-1} + \zeta_{3,2}(c_{t-1,bch} - c_{t-2,bch}) + \nu_{3,t}$$

$$\nu_{3,t} \sim (i.i.d.)\mathcal{N}\left(0, \kappa_3^2\right) \tag{4.14}$$

4.2.3 Second-Stage Estimation

In the first stage, we recover the parameters $\hat{\phi}$ of the reduced-form policy functions governing miners' mining decisions and transition equations for the evolution of state variables assuming first-order Markov

processes. The estimated policy functions inform us about the beliefs of miners in equilibrium and how they will respond in a given market condition. In an equilibrium model, we assume miners have rational beliefs about state transitions and the actions of their rivals. We can use estimated $\hat{\phi}$ to simulate many paths of plays starting from different market conditions. Therefore, we can effectively simulate miners' reactions under different policy regimes for the same market conditions.

We drop the btc/bch supscripts/superscripts for this part. We use these simulations to construct empirical estimates of the ex-ante value function. Then, we use the equilibrium condition (4.6) to form the minimum distance criterion function to recover the underlying structural parameters, θ. We will use the subset of the estimates $\hat{\phi}$ from first-stage to approximate the MPE profile σ. $\hat{\sigma}$ is the actual policy employed by miners to make mining decisions and is an approximation for σ. $\widetilde{\sigma}$ is the behavior under alternative policy (generated via perturbation of $\hat{\sigma}$). Using θ (or θ') and $\hat{\sigma}$, we can simulate ex-ante values for any miner i starting with initial conditions χ_{t_1}.

$$\begin{aligned}\overline{\mathbb{V}}\left(\chi_{i,t_1};\hat{\sigma},\theta\right) &= \mathbb{E}\left[\sum_{\tau=t_1}^{T}\beta^{\tau-1}\Pi_i\left(\hat{\sigma}\left(\chi_\tau,\epsilon_\tau\right),\chi_\tau,\epsilon_\tau;\theta\right)\middle|\chi_{t_1},\hat{\sigma}\right] \\ &= \frac{1}{\overline{S}}\sum_{s=1}^{\overline{S}}\sum_{\tau=t_1}^{T}\beta^{\tau-1}\Pi_i\left(\hat{\sigma}\left(\chi_\tau^s,\epsilon_\tau^s\right),\chi_\tau^s,\epsilon_{i\tau}^s;\theta\right) \end{aligned} \tag{4.15}$$

where we simulate $\overline{S}$ paths of play until T time periods. Superscript s denotes different paths generated starting from χ_{t_1}. We repeat this procedure B times with different χ. To compute ex-ante value functions for the alternative policies, $\widetilde{\sigma}$, we repeat the above procedure using $\widetilde{\sigma}$ in place of $\hat{\sigma}$. We generate 200 alternate polices with multiplicative perturbations following the recommendations in Srisuma (2013). Our linear-in-parameters specification of the one-shot profit function results in massive computational saving as we only need to forward simulate once. The simulated $\overline{\mathbb{V}}\left(\cdot;\hat{\sigma},\theta\right)$ and $\overline{\mathbb{V}}\left(\cdot;\widetilde{\sigma},\theta\right)$ can then be evaluated for different candidate values of θ.

BBL posits that in equilibrium, actual policy is always weakly preferred to other feasible alternatives at every other state. Therefore, the policy $\hat{\sigma}_i$ for miner i will yield a weakly higher value of expected present discounted profits than any other policy $\widetilde{\sigma}_i$ when the rival miners are playing by $\hat{\sigma}_{-i}$. The differences in value functions generated by these two class of policies, actual and alternative, are used to estimate the

structural parameters. Given B inequalities with different initial conditions ($\chi_{t_1}^b$), we minimize the sum of squared negative deviations from the equilibrium play. In effect, our simulated minimum distance estimator minimizes the violations of the optimality condition through the objective function, $Q(\theta)$, defined as

$$Q(\theta) = \frac{1}{B}\sum_{b=1}^{B}\left(\min\left\{\overline{\mathbb{V}}_i\left(\chi_{t_1}^b;\hat{\sigma},\theta\right) - \overline{\mathbb{V}}_i\left(\chi_{t_1}^b;\widetilde{\sigma}_i,\hat{\sigma}_{-i},\theta\right),0\right\}\right)^2 \tag{4.16}$$

A positive value of the minimum distance, $Q(\theta)$, at $\hat{\theta}$ signifies that at least one miner has a profitable deviation. We simulate each inequality 2,500 times for 100 time periods.

Our identification scheme relies on exploiting the temporal variations in miners' decisions throughout the observation period. In the period of our study, spanning 9,368 miner-time period observations, the observed price volatility (Figure 2.1d) created conditions where one coin was more profitable than the other allowing us to observe variations in miners decision-making.

4.3 Results

The goal of this study is to investigate the driving forces behind the observed mining decisions made by BTC and BCH miners from April 2017 to February 2018. This period includes the BTC fork, introduction of an emergency difficulty adjustment algorithm to the forked cryptocurrency, BCH, and the cryptocurrency bubble that led to massive over-valuation of the cryptocurrency market. We compute three different equilibria for the BTC miners (pre-fork, EDA-active, and Non-EDA) and two for BCH miners (EDA-active and Non-EDA). We specifically focus our attention on the endogenous impact of transaction fees and competition on miners' utilities while allowing for dynamic adjustments in the difficulty levels.

First-stage results are presented in the Appendix in Table 3, 4 and 5. In the second stage, we sampled 1,171 inequalities (out of a total sample of 1,284). Each inequality comprises of a randomly sampled miner, starting time period, tuple of state variables (prices, fees, difficulty levels) for that time period, and the set of alternative policies. We simulated 2,500 paths of play for every inequality for 100 time periods. The discount factor β was set to 0.95[8]. The standard errors for second stage were calculated by bootstrapping using 200

[8]Our results remain robust to other commonly preferred discount factors, such as 0.9 and 0.97.

replications with replacement. Table 4.1 presents our estimates for the structural parameters governing BTC and BCH miners dynamic decision-making during the three sub-periods.

We begin our discussion of the results with BTC miners' structural parameters. As expected, price plays a significantly positive role in BTC miners' utility. During the period of our study, BTC's price increased from \$1,000 to \$19,000 (Figure 2.1d). Our estimates for the θ_1^{btc} for the three periods reflect the increasing importance of price in miners' per-period utility. At the same time, we find BTC miners' preference over the number of mined blocks, θ_3^{btc}, decreases from 0.1053 in pre-fork period to 0.0442 in the Non-EDA period. This indicates that miners, in equilibrium, preferred mining fewer BTC blocks as the prices and difficulties rose. The rising difficulty levels (Figure 2.1c), from pre-fork period to the Non-EDA period, increased the cost of mining significantly and it is visible through the magnitude of θ_4^{btc}. In line with the preferences over BTC blocks, miners' utility was positively impacted by transaction fees (θ_2^{btc}), however, their contribution to the payoff decreases monotonically as we go from pre-fork period to Non-EDA period.

In section 4.1.1, we derive the threshold N^* for number of active miners, below which, expected utility is increasing in competition. Our estimates for the competition among BTC miners is negative and statistically significant i.e. competition has an overall positive impact on miners' utilities. This implies that the number of active miners were indeed below N^* during the observed mining periods. Moreover, we see that the benefit of competition increases from pre-fork period to the Non-EDA period. During the high-demand lucrative periods, BTC mining witnesses contribution of hash power from new miners[9]. This reduces the computational burden on all active miners and consequently, the total cost of mining for the BTC network. Moreover, addition of new hash power has the added benefit of making the BTC network more secure and trustworthy (Houy, 2014).

BCH miners do not show any preference for BCH prices. The price coefficient, θ_1^{bch} is positive but statistically insignificant. BCH difficulty (θ_4^{bch}) has a negative impact on miners' utilities. However, the effect is quite small compared to the cost of mining BTC in the corresponding time periods. As expected, transaction fees (θ_2^{bch}) increase miners' payoff from mining. Their effect increases from 0.0252 during the EDA-Active period to 0.0527 during the Non-EDA period. However, we notice that effect of mining

[9]In our setup, this would be reflected by the growing contributions of Other and Unknown mining pools.

Table 4.1: Estimates for the structural parameters for BTC and BCH miners.

Structural Parameters	Pre-fork *(Apr 2017 - Jul 2017)*	EDA-Active *(Aug 2017 - Nov 2017)*	Non-EDA *(Nov 2017 - Feb 2018)*
BTC Miners			
Price (θ_1^{btc})	0.0960^{***}	0.1273^{***}	0.2025^{***}
	(0.0204)	(0.0273)	(0.0554)
Transaction Fees (θ_2^{btc})	0.0729^{***}	0.0630^{***}	0.0547^{***}
	(0.0096)	(0.0083)	(0.0084)
BTC Blocks (θ_3^{btc})	0.1053^{***}	0.0763^{***}	0.0442^{***}
	(0.0044)	(0.0041)	(0.0044)
Difficulty (θ_4^{btc})	0.0484^{***}	0.0596^{***}	0.0922^{***}
	(0.0055)	(0.0083)	(0.0181)
Competitors' BTC Blocks (θ_5^{btc})	-0.0254^{**}	-0.0464^{***}	-0.0834^{***}
	(0.0084)	(0.0079)	(0.0076)
BCH Miners			
Price (θ_1^{bch})	$n.a.$	0.0185	0.0081
	$n.a.$	(0.0149)	(0.0291)
Transaction Fees (θ_2^{bch})	$n.a.$	0.0252^{***}	0.0527^{***}
	$n.a.$	(0.0028)	(0.0067)
BCH Blocks (θ_3^{bch})	$n.a.$	0.1440^{***}	0.1059^{***}
	$n.a.$	(0.0070)	(0.0066)
Difficulty (θ_4^{bch})	$n.a.$	0.0163^{***}	0.0173^{*}
	$n.a.$	(0.0037)	(0.0084)
Competitors' BCH Blocks (θ_5^{bch})	$n.a.$	0.0528^{***}	-0.0130^{**}
	$n.a.$	(0.0037)	(0.0050)
BTC blocks (θ_6^{bch})	$n.a.$	-0.0443^{***}	-0.0250^{***}
	$n.a.$	(0.0032)	(0.0023)
BTC Blocks^2 (θ_7^{bch})	$n.a.$	0.0182^{***}	0.0085^{***}
	$n.a.$	(0.0011)	(0.0009)

Note: Standard Errors are reported under the parameter estimates in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

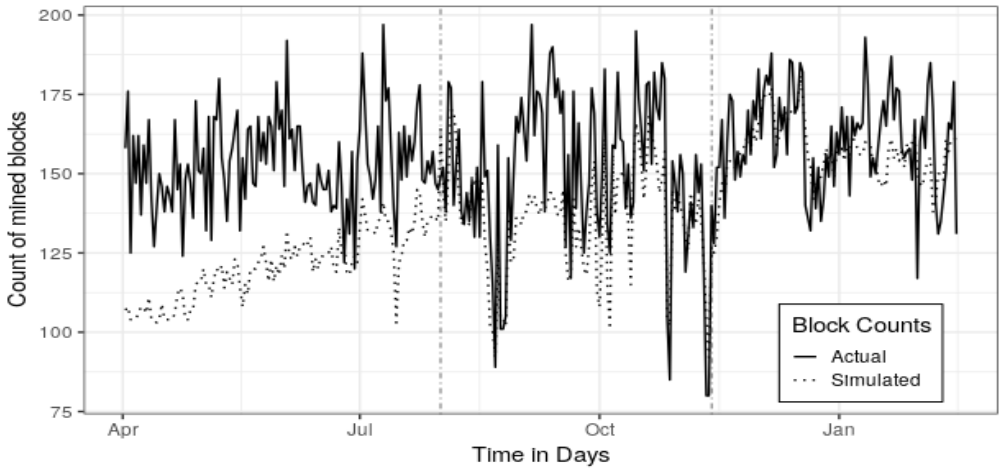
BCH blocks (θ_3^{bch}) on miners' utilities decreases from 0.1440 during the EDA-Active period to 0.1059 during the Non-EDA period. This was a consequent of the EDA algorithm. During EDA-Active periods, EDA-triggered reductions in BCH difficulty made mining BCH much more lucrative. However, after EDA deactivation, fees started playing bigger role in BCH miners' decision-making.

N^* also applies to BCH mining. It's dependence on the difficulty levels will alter the threshold value above which competitions starts eroding BCH miners' expected per-period payoffs. Due to EDA, BCH difficulties were lowered by almost 50-75% (Figure 2.1c). This, in turn, lowered the threshold value for the number of active miners. This is why, we observe that competition led to erosion of miner profits during the EDA-active period ($\theta_5^{bch} = 0.0528$). Difficulty levels stabilized when EDA was deactivated (in the Non-EDA period) and a difficulty adjustment algorithm, based on simple moving average, was introduced. This allowed N^* to stabilize at levels where competition contributed positively to miners' payoff ($\theta_5^{bch} = -0.0130$) by curtailing the expected computational burden on all active BCH miners.

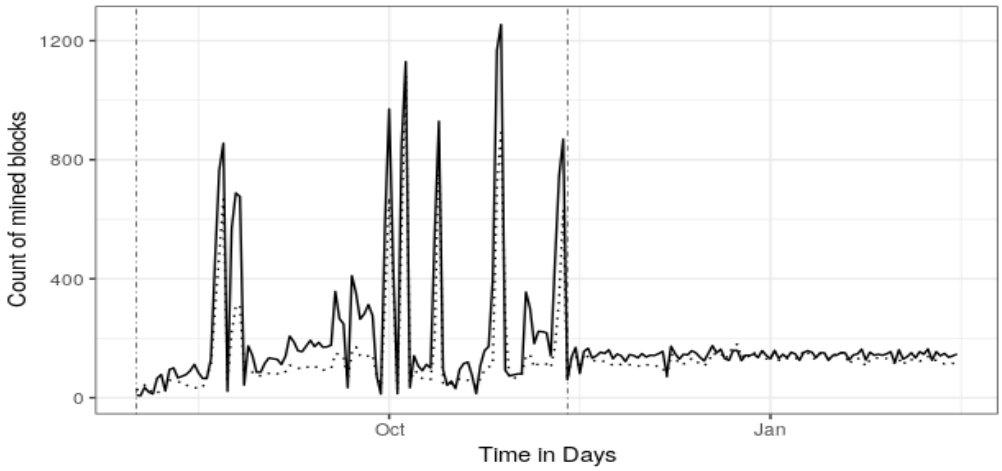
In our forward-looking setup for BCH miners, we assume that the said miner takes the decision of its BTC counterpart as given and then, maximizes its expected discounted profits. Our results indicate, as expected, that the rate of BTC mining (by the same mining pool) has a much stronger effect during the EDA-Active period as compared to Non-EDA period. As the number of BTC miner's blocks increase, the per-period utility of BCH miner is negatively affected where the effect increases quadratically. This is again a consequence of EDA. BCH miner competes with its BTC counterpart for precious hash power. Higher BTC blocks result in lower levels of hash power for the BCH miner. Since EDA made BCH mining very lucrative, lower availability of hash power (due to BTC miner's actions) had a much stronger negative effect during the EDA-Active periods.

4.3.1 Model Fit

We find our estimated model performs well during simulation runs. We contrast the actual rate of block generation with the outcomes using our model. Figure 4.2 presents a succinct overview of its effectiveness in predicting the mining decisions. Model simulations are implemented via the forward simulation approach of Benkard et al. (2010).



(a) BTC blocks.



(b) BCH blocks.

Figure 4.2: Actual vs Simulated number of mined blocks for BTC and BCH. Vertical lines mark the time of BTC fork and the deactivation of Emergency Difficulty Adjustment algorithm on BCH.

4.3.2 Counterfactual Policy Simulations

BCH developers introduced EDA with the intent to motivate miners to contribute their hash power to the newly born minority BCH chain over BTC. However, their reasons to prefer the implemented design for EDA, 20% difficulty reduction after a 12 hours gap, is not very clear. Coinbase rewards are halved after every 210,000 blocks to keep the underlying currency's inflation in check by lowering the supply of new coins (Nakamoto, 2008). Halved rewards would adversely affect individuals' incentives to participate in mining. As such, there is a tension between miners' goals (maximize revenues) and BCH developers' (survival of BCH). We run this set of counterfactual simulations to ascertain the comparative effectiveness of the chosen EDA design with respect to other possible schemes from the perspectives of miners and developers. We look to the forward simulation procedure elaborated in Benkard et al. (2010) and applied by Blevins (2016) to estimate these counterfactuals. This approach does not require one to solve the computationally intractable dynamic model.

Simulated counts of mined BCH blocks and advancement in halving of rewards are presented in Tables 4.2 and 4.3 respectively. Overall level of dispersion (standard deviation) in the BCH block counts is 12,911 which drops to 5,460 after removal of the extreme 3 hours regime.

		Reduction Factor				
		50%	40%	30%	20%	10%
Time Difference (in hours)	3	78359	64932	53176	41088	31453
	6	41538	36315	30549	25930	21257
	9	31442	28508	24862	21741	19848
	12	28524	25600	22378	20417	19255
	15	26475	24055	21217	19423	18370
	18	24582	22638	20485	19208	17784
	21	23274	22054	20032	18696	16985
	24	23082	21638	19313	17945	16371

Table 4.2: Simulated counts of mined BCH blocks during the EDA-active periods for different EDA designs. Highlighted cell represents the count of BCH blocks for the actual EDA design.

Simulations for BCH mining under different EDA designs display very dispersed counts. Extreme design configurations, such as, 20-50% difficulty adjustment after every 3 hours or 40-50% after 6 hours, will hasten

the arrival of the day when mining rewards would be halved. Ideally, a block should be generated every 600 seconds. For extreme designs, we find that almost 2x-4x BCH blocks would have been mined in comparison to the actual design's simulated rate, which would advance halving of rewards by another 4-12 months (Table 4.3).

		Reduction Factor				
		50%	40%	30%	20%	10%
Time Difference (in hours)	3	14.19	11.19	8.55	5.84	3.69
	6	5.95	4.78	3.48	2.45	1.40
	9	3.68	3.03	2.21	1.51	1.09
	12	3.03	2.38	1.65	1.21	0.95
	15	2.57	2.03	1.39	0.99	0.76
	18	2.15	1.71	1.23	0.94	0.62
	21	1.85	1.58	1.13	0.83	0.45
	24	1.81	1.49	0.97	0.66	0.31

Table 4.3: Halving of BCH rewards would be advanced by the above number of months for different EDA designs. Highlighted cell represents the months for the actual EDA design.

Our counterfactual exercise informs us of the feasibility of multiple EDA configurations which would allow BCH developers to preserve similar reward halving advancements. For example, 30% reduction in BCH's difficulty after every 18 hours, would have resulted similar advancement of reward halving.

Chapter 5

CONCLUSION

5.1 Creative Idea Generation in Open Innovation Platforms

The ability to think creatively fuels competitive advantage for companies by enabling them to make novel products and services. According to the Forrester Research (Forrester, 2014), companies that encourage creative thinking are 3.5 times more likely to achieve 10% revenue growth goals and are 50% more likely to enjoy higher market shares over their competitors.

With the heightened interest on creative thinking skills, this study is motivated to advance our understanding of how individuals generate creative ideas. Such an understanding will enable us to offer insights on how to cultivate creative thinking skills. Idea generation process is simultaneously influenced by multiple activities that commonly occur in an open innovation platform: knowledge acquisition from participants' interactions with each others' ideas, deliberate practice through persistent participation, and learning through failures. Due to the dynamic interplay among these activities, which influence creative ideation outcomes and are themselves influenced by these outcomes, it is challenging to identify each activity's influence on creative ideation outcomes using a reduced-form regression analysis. Moreover, individuals' ideation capabilities are expected to vary with their abilities. To overcome these challenges, we employed a novel empirical framework, the mutually exciting spatio-temporal point process model with unobserved heterogeneity, and uncovered their synergetic effects in creative idea generation. We found that knowledge acquired through interactions influences creative ideation outcome by enhancing individual's ability to select relevant problems to ideate for. Further, we found direct impact of deliberate practice in motivating individuals to persist in generating ideas which leads to improvements in ideation skills. Finally, failures were found to drive individuals to learn from others' ideas by increasing their future knowledge acquisition rate and improving their ideation abilities.

Our research makes the following theoretical and methodological contributions. Foremost, our research

contributes to empirical studies on the effectiveness of analogical reasoning in generating creative ideas (Dunbar, 1995; Fu et al., 2013) by introducing a novel empirical framework. Although the creativity theory suggests that knowledge acquisition and its application processes are endogenous, the prior studies share implicit assumptions relating to a) problem selection is exogenous i.e. individuals do not have any hand in selecting the problems they ideate for, and b) all analogies occur independently of past behavior (Dahl and Moreau, 2002; Yu et al., 2014). Such assumptions are inconsistent with innate motivations to participate in creativity-oriented tasks where individuals' intrinsic preferences play a strong role (Amabile, 1988). Our modeling strategy alleviates the need for such assumptions by explicitly modeling knowledge acquisition and ideation processes' ability to affect themselves as well as other processes. Our benchmark analysis reveals the importance of interdependence among these processes which are key to discerning the nuanced impact of acquired knowledge on future ideation outcomes.

Our study also contributes to the literature on community-based open innovation platforms. Members' interactions among themselves and with the sponsoring firms are vital for the overall health of such innovation communities. These interactions can be essential to motivate and guide the participants towards desirable innovation goals (West and Lakhani, 2008). Extant literature focuses its theoretical attention to either one-way communication of firms with participants (e.g. West and O'mahony, 2008) or role of community leaders and their interactions (e.g. Fichter, 2009). This paper, using members' interactions among themselves over their creative ideas, expands our knowledge of the vital role played by knowledge-based communications on creative outcomes as well as their relative importance in the face of success and failure in creativity tasks.

Methodologically, our contribution lies in the extension of two-dimensional spatio-temporal mutually exciting models to high-dimensional semantic data (such as images or text). Earlier applications were limited to 2-dimensional euclidean distances where earthquake occurrences related geospatial data were used to create a spatio-temporal model of seismological activities (Veen and Schoenberg, 2008; Ogata, 1998). Majority of clickstream or trajectory data points are associated with raw unstructured information, such as images or text descriptions, which can be numerically approximated to fixed-length vectors using machine learning methods (Le and Mikolov, 2014). In this paper, we show how the complex high-dimensional in-

tegrals (Equation (3.13), see Appendix .3) can be numerically approximated without needing to analytically solve over many dimensions (see Ogata, 1998 for 2D application). This opens many possibilities to utilize semantic information in clickstream data, such as, including tweet contents in information diffusion on twitter or including advertisement message and image data for calculating click through rates, etc. while accounting for the dynamic interactions between tweets or advertisement types. We also extend the EM algorithm for Hawkes processes to include multi-dimensional unobserved user heterogeneity through non-parametric random effects.

The rise of open innovation platforms and firms' increasing reliance on such platforms for their R&D needs has created a compelling need to find effective methods to increase creative throughput. As such, our research generates recommendations for the platform managers to consider. Since they have active knowledge of their participants and the upcoming problems on the platform, their targeted impetus to participants can increase the platform's creative throughput. Platform managers would find it advantageous to direct ideators, who have not been successful in the past or new ideators, towards challenge-relevant sources of knowledge. While these ideators could be using knowledge from other domains to solve the current challenge, the challenge-relevant knowledge would allow them to tailor their ideas to aspects of problems that are important to solving it. Our results indicate that ideas or solutions, albeit creative, fail to convey their value if they are viewed as being uncorrelated with the challenge. Therefore, recommending challenge-specific knowledge and its subsequent inclusion in the idea submitted by the ideator, can make the idea more relatable to the problem.

We would also like to highlight the importance of encouraging all types of participants to interact with (or acquire knowledge from or provide feedback to) each other's creative outputs. Introduction of incentive mechanisms that can nudge potential ideators and platform participants to share their views with others results in higher knowledge levels for all involved. This act of knowledgeable interaction spawns future knowledgeable interactions with a high probability. While knowledge levels are one part of the creativity process, application of that knowledge is the other part. Since the act of acquiring knowledge leads to future idea submission activity, thereby, encouraging the acquirer to also learn to apply the acquired knowledge. Our findings indicate that learning-by-doing is a strong driver of creativity. Therefore, efforts in that dir-

ection, whether direct such as encouraging past ideators to submit ideas again or indirect such as via the knowledge acquisition - knowledge application route, will lead to higher overall *creative* submissions for a posted challenge.

This study is not without its limitations and as such, provides avenues for future research. In this paper, we are interested in how individuals utilize analogies that they find intriguing enough to acquire knowledge from. Though it is sufficient for our analysis as only intrinsic preferences and their outcomes affect creativity (Amabile, 1988), it would be useful to include passive observation data (i.e. analogies that they did not find interesting enough to acquire knowledge from) into our analysis to model individuals' knowledge acquisition processes more intricately. That is, viewing log data will enable scholars to also analyze what knowledge is actively avoided by community members and whether such evasion is beneficial or not. Secondly, every submission goes through many iterations based on constant feedback from the community. Since we only observe the final output, we are unable to investigate micro changes brought about by others' feedback. Availability of submissions' version histories will provide insights into what kind of comments' characteristics are useful for ideators and what are not. Given the tournament nature of our research setting, participants come to know of their winning status after a certain time delay. It is a limitation of our study that we do not explicitly model this delay, however, we perform robustness check (Appendix **??**) to investigate its effect on our results. Lastly, our findings might only be generalizable to open innovation platforms dealing with socio-economic problems i.e. our results will apply for those sub-populations that are motivated by such problems. Whether these findings will generalize for other platforms, specializing in other problem domains, is left for future research work when other platforms' data becomes available.

Overall, our study provides us with insights into how individuals learn through their interactions and participations on an open innovation platform through a novel application of mutually exciting spatio-temporal point process model. Our work finds its motivations in the growing debate over future of work and importance of open innovation platforms. As such, our recommendations can be used to improve participants' learning experiences as well as the creative throughput of the platform.

5.2 Cryptocurrency Mining and impact of Emergency Difficulty Adjustment Algorithm

Over the last couple of years, cryptocurrency valuations have wildly fluctuated between $100 B and $ 800 B. Their volatility and sky-high valuations have piqued the interest of everyone in industry, academia, and government. Their growing valuations and demand has encouraged individuals and organizations to start mining. Many new currencies have been proposed which aim to provide lightning-fast payments facilities, smart contracts, utilization of idle computing resources, data storage, etc. Quite a few of the proposed currencies are derivatives of older currencies and have been created via code forks. The dynamic ecosystem developing around cryptocurrencies will lead to more code forks and increase the limelight on role played by incentives like EDA and the nature of competition between the parent and children cryptocurrencies.

In this paper, we investigate competition in Bitcoin and Bitcoin-Cash mining. Specifically, we focus our attention around the time period of the very popular cryptocurrency fork that resulted in the creation of Bitcoin Cash from Bitcoin. Further, we analyze the resulting competitive equilibria in BCH, with and without EDA. EDA was introduced to incentivize miners to process BCH's transactions over those of the more popular and dominant BTC's (since BTC and BCH shared the same difficulty algorithms). We aim to estimate the structural parameters that motivate the observed strategic behavior of the forward-looking miners during and after EDA using the two-step estimator of Bajari et al. (2007). We use the publicly-available chain data to find that competition in Bitcoin mining has immense value to Bitcoin mining. Apart from increasing the security and trust in the system, competition also enables miners to share the computational load of mining during high demand or high difficulty regimes. We also find that difficulty adjustment algorithms like EDA, can distort this positive externality of competition which results in lowering of expected payoffs for BCH miners. Through our counterfactual policy simulations, we find that different EDA configurations would have been satisfied developers' halving concerns with similar rates of BCH mining.

Our work, to the best of our knowledge, is the first to empirically investigate the dynamic changes in the competitive mining ecosystem during the tumultuous periods of Bitcoin fork and cryptocurrency boom. Our econometric analysis uncovers the structural determinants of miners' decision-making and their expected contributions to miners' payoffs . We believe cryptocurrency developers would find our research helpful in understanding economic implications of incentives like EDA and their overall impact on participating

miners' profits in comparison to their currency's rewards schedule. Using our analysis, individuals and groups involved in mining can benefit by making better-informed decisions when faced with competing cryptocurrencies.

We make some simplifying assumptions in our modeling which also provide avenues for future research. In our formulation, we assume all profits are realized in the same time period. However, many miners hold on their earned coins and sell them at a later stage. This limitation can be relaxed in future works through investigation of exchange-level trading data of miners. We also assume that miners apply computational power commensurate with the observed mined coins but the relationship is not exactly linear. In future, scholars can analyze the dynamic relationship between hash power contributions and realized mining outcomes using data from mining pools. Scholars will also find it interesting to investigate the recent regulatory interventions in the cryptocurrency industry whose long-term impacts aren't fully understood and are topics for future research.

BIBLIOGRAPHY

Ronald A. Finke, Steven M. Smith, and Thomas B. Ward. *Creative Cognition*. MIT Press, 1992.

Victor Aguirregabiria and Pedro Mira. Sequential estimation of dynamic discrete games. *Econometrica*, 75 (1):1–53, 2007.

Robert Allen. Collective invention. *Journal of Economic Behavior & Organization*, 4(1):1–24, 1983. URL https://EconPapers.repec.org/RePEc:eee:jeborg:v:4:y:1983:i:1:p:1-24.

Eitan Altman, Alexandre Reiffers, Daniel S. Menasche, Mandar Datar, Swapnil Dhamal, and Corinne Touati. Mining competition in a multi-cryptocurrency ecosystem at the network edge: A congestion game approach. *SIGMETRICS Perform. Eval. Rev.*, 46(3):114–117, January 2019. ISSN 0163-5999. doi: 10.1145/3308897.3308950. URL https://doi.org/10.1145/3308897.3308950.

T. M. Amabile. A Model of Creativity and Innovation in Organizations. *Research in Organizational Behavior*, 10:123–167, 1988.

Kenneth J. Arrow. The economic implications of learning by doing. *The Review of Economic Studies*, 29 (3):155–173, 1962. ISSN 00346527, 1467937X. URL http://www.jstor.org/stable/2295952.

Thomas B. Ward, Steven M. Smith, and Ronald A. Finke. *The Creative Cognition Approach*. MIT Press, 1995.

Emmanuel Bacry, Iacopo Mastromatteo, and Jean-François Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01):1550005, 2015.

Patrick Bajari, Benkard C. Lanier, and Levin Jonathan. Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370, 2007.

Patrick Bajari, Phoebe Chan, Dirk Krueger, and Dan Miller. A dynamic structural model of housing demand: Estimation and policy implications. *Department of Economics, University of Minnesota working paper*, 2008.

C Lanier Benkard, Aaron Bodoh-Creed, and John Lazarev. Simulating the dynamic effects of horizontal mergers: Us airlines. *Manuscript, Yale University*, 2010.

BitcoinABC. Difficulty Adjustment Algorithm Update. https://www.bitcoinabc.org/2017-11-01-DAA/, 2017. Accessed: 2019-03-04.

Jason R. Blevins. Sequential Monte Carlo Methods for Estimating Dynamic Microeconomic Models: SEQUENTIAL MONTE CARLO METHODS. *Journal of Applied Econometrics*, 31(5):773–804, August 2016.

Jason R. Blevins, Ahmed Khwaja, and Nathan Yang. Firm Expansion, Size Spillovers, and Market Dominance in Retail Chain Dynamics. *Management Science*, November 2017.

Blockchair. Blockchair. https://blockchair.com, 2018. Accessed: 2019-01-14.

Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, et al. Universal sentence encoder. *arXiv preprint arXiv:1803.11175*, 2018.

Joel Chan, Steven P. Dow, and Christian D. Schunn. Do the best design ideas (really) come from conceptually distant sources of inspiration? *Design Studies*, 36:31 – 58, 2015.

WG Chase and HA Simon. Skill in chess. *American Scientist*, 61(4):394–403, 1973.

Bo T. Christensen and Christian D. Schunn. The relationship of analogical distance to analogical function and preinventive structure: the case of engineering design. *Memory & Cognition*, 35(1):29–38, January 2007.

CoinAPI. Coinapi. https://www.coinapi.io/, 2018. Accessed: 2019-01-14.

John P. Conley. Blockchain and the Economics of Crypto-tokens and Initial Coin Offerings. Technical report, Vanderbilt University Department of Economics, 2017.

Darren W Dahl and Page Moreau. The influence and value of analogical thinking during new product ideation. *Journal of Marketing Research*, 39(1):47–60, 2002.

Andrew M Dai, Christopher Olah, and Quoc V Le. Document embedding with paragraph vectors. *arXiv preprint arXiv:1507.07998*, 2015.

Daryl J. Daley and D. Vere-Jones. *An introduction to the theory of point processes*. Springer, New York, 2nd edition, 2003.

Nicola Dimitri. Bitcoin mining as a contest. *Ledger*, 2:31–37, Sep. 2017. doi: 10.5195/ledger.2017.96. URL https://www.ledgerjournal.org/ojs/ledger/article/view/96.

Kevin Dunbar. How scientists really reason: Scientific reasoning in real-world laboratories. In *The nature of insight.*, pages 365–395. The MIT Press, Cambridge, MA, US, 1995.

Prajit K Dutta and Rangarajan K Sundaram. The equilibrium existence problem in general markovian games. *Organizations with incomplete information*, 5:159–207, 1998.

David Easley, Maureen O'Hara, and Soumya Basu. From mining to markets: The evolution of bitcoin transaction fees. *Journal of Financial Economics*, 134(1):91 – 109, 2019. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2019.03.004. URL http://www.sciencedirect.com/science/article/pii/S0304405X19300583.

Ellen Enkel and Oliver Gassmann. Creative imitation: exploring the case of cross-industry innovation. *R&d Management*, 40(3):256–270, 2010.

Klaus Fichter. Innovation communities: the role of networks of promotors in open innovation. *R&D Management*, 39(4):357–371, 2009.

Lee Fleming. Recombinant uncertainty in technological search. *Management Science*, 47(1):117–132, 2001. doi: 10.1287/mnsc.47.1.117.10671. URL https://doi.org/10.1287/mnsc.47.1.117.10671.

Forrester. Does creativity matter? adobe research shows dramatic impact on business results. https://news.adobe.com/press-release/adobe-creative-cloud-dps/media-alert-does-creativity-matter-adobe-research-shows, 2014. Accessed April 20, 2019.

Nikolaus Franke and Sonali Shah. How communities support innovative activities: an exploration of assistance and sharing among end-users. *Research Policy*, 32(1):157 – 178, 2003. ISSN 0048-7333. doi: https://doi.org/10.1016/S0048-7333(02)00006-9. URL http://www.sciencedirect.com/science/article/pii/S0048733302000069.

Katherine Fu, Joel Chan, Jonathan Cagan, Kenneth Kotovsky, Christian Schunn, and Kristin Wood. The meaning of "near" and "far": the impact of structuring design databases and the effect of distance of analogy on design output. *Journal of Mechanical Design*, 135(2):021007, 2013.

Dedre Gentner, Sarah Brem, Ron Ferguson, and Philip Wolff. Analogy and creativity in the works of Johannes Kepler. In Thomas B. Ward, Steven M. Smith, and Jyotsna Vaid, editors, *Creative thought: An investigation of conceptual structures and processes.*, pages 403–459. American Psychological Association, Washington, 1997.

Karan Girotra, Christian Terwiesch, and Karl T. Ulrich. Idea Generation and the Quality of the Best Idea. *Management Science*, 56(4):591–605, April 2010.

Stuart Haber and W Scott Stornetta. How to time-stamp a digital document. In *Conference on the Theory and Application of Cryptography*, pages 437–455. Springer, 1990.

Peter F. Halpin. A scalable em algorithm for hawkes processes. In Roger E. Millsap, L. Andries van der Ark, Daniel M. Bolt, and Carol M. Woods, editors, *New Developments in Quantitative Psychology*, pages 403–414, New York, NY, 2013. Springer New York.

Peter F. Halpin and Paul De Boeck. Modelling dyadic interaction with hawkes processes. *Psychometrika*, 78(4):793–814, October 2013.

Alan G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1): 83–90, 1971a.

Alan G. Hawkes. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33(3):438–443, 1971b.

John R. Hayes. Cognitive Processes in Creativity. In John A. Glover, Royce R. Ronning, and Cecil R. Reynolds, editors, *Handbook of Creativity*, pages 135–145. Springer US, Boston, MA, 1989.

John R Hayes, Linda Flower, Karen A Schriver, James Stratman, Linda Carey, et al. Cognitive processes in revision. *Advances in applied psycholinguistics*, 2:176–240, 1987.

Matthew Haywood. Miners gaming the BCash emergency difficulty adjustment. https://bravenewcoin.com/insights/miners-gaming-the-bcash-emergency-difficulty-adjustment, 2017. Accessed: 2019-03-04.

E Tory Higgins. Knowledge activation: Accessibility, applicability, and salience. *Social psychology: Handbook of basic principles*, pages 133–168, 1996.

V Joseph Hotz, Robert A Miller, Seth Sanders, and Jeffrey Smith. A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies*, 61(2):265–289, 1994.

Nicolas Houy. The economics of Bitcoin transaction fees. Working Papers 1407, Groupe d'Analyse et de Théorie Economique Lyon St-Étienne (GATE Lyon St-Étienne), Université de Lyon, 2014. URL https://ideas.repec.org/p/gat/wpaper/1407.html.

Jeff Howe. The rise of crowdsourcing. *Wired magazine*, 14(6):1–4, 2006.

Yan Huang, Param Vir Singh, and Kannan Srinivasan. Crowdsourcing New Product Ideas Under Consumer Learning. *Management Science*, 60(9):2138–2159, September 2014.

Elina H. Hwang, Param Vir Singh, and Linda Argote. Jack of all, master of some: Information network and innovation in crowdsourcing communities. *Information Systems Research, Forthcoming*, 2019.

M. Jamshidian and R. I. Jennrich. Standard errors for em estimation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(2):257–270, 2000. doi: 10.1111/1467-9868.00230.

Dmytro Karabash. On stability of hawkes process. *arXiv preprint arXiv:1201.1573*, 2012.

Thomas S Kuhn. *The structure of scientific revolutions*. University of Chicago press, 2012.

Patrick J Laub, Thomas Taimre, and Philip K Pollett. Hawkes processes. *arXiv preprint arXiv:1507.02822*, 2015.

Quoc Le and Tomas Mikolov. Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, ICML'14, pages II–1188–II–1196. JMLR.org, 2014.

Christian Lüthje, Cornelius Herstatt, and Eric von Hippel. User-innovators and "local" information: The case of mountain biking. *Research Policy*, 34(6):951 – 965, 2005. ISSN 0048-7333. doi: https://doi.org/10.1016/j.respol.2005.05.005. URL http://www.sciencedirect.com/science/article/pii/S0048733305000971.

June Ma, Joshua S Gans, and Rabee Tourky. Market structure in bitcoin mining. Working Paper 24242, National Bureau of Economic Research, January 2018. URL http://www.nber.org/papers/w24242.

Eric Maskin and Jean Tirole. Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory*, 100(2):191–219, 2001.

Oren Melamud, David McClosky, Siddharth Patwardhan, and Mohit Bansal. The role of context types and dimensionality in learning word embeddings. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1030–1040, San Diego, California, June 2016. Association for Computational Linguistics. doi: 10.18653/v1/N16-1118. URL https://www.aclweb.org/anthology/N16-1118.

Menegerian. Bringing Stability to Bitcoin Cash Difficulty Adjustments. https://medium.com/@Mengerian/bringing-stability-to-bitcoin-cash-difficulty-adjustments-eae8def0efa4, 2017. Accessed: 2019-03-04.

Swapnil Mishra, Marian-Andrei Rizoiu, and Lexing Xie. Feature driven and point process approaches for popularity prediction. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, CIKM '16, pages 1069–1078. ACM, 2016.

Pamela D. Morrison, John H. Roberts, and Eric von Hippel. Determinants of user innovation and innovation sharing in a local market. *Management Science*, 46(12):1513–1527, 2000. doi: 10.1287/mnsc.46.12.1513.12076. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.46.12.1513.12076.

Satoshi Nakamoto. Bitcoin: a peer-to-peer electronic cash system. 2008.

Bernard A Nijstad and Wolfgang Stroebe. How the group affects the mind: A cognitive model of idea generation in groups. *Personality and social psychology review*, 10(3):186–213, 2006.

Yoshiko Ogata. The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 30(1):243–261, 1978.

Yosihiko Ogata. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association*, 83(401):9–27, 1988.

Yosihiko Ogata. Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2):379–402, 1998.

Tohru Ozaki. Maximum likelihood estimation of hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145–155, 1979.

Morgen E. Peck. Blockchains: How they work and why they'll change the world. *IEEE Spectrum*, 54(10): 26–35, 2017.

Jeffrey Pennington, Richard Socher, and Christopher Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar, October 2014. Association for Computational Linguistics. doi: 10.3115/v1/D14-1162. URL https://www.aclweb.org/anthology/D14-1162.

Martin Pesendorfer and Philipp Schmidt-Dengler. Asymptotic least squares estimators for dynamic games. *The Review of Economic Studies*, 75(3):901–928, 2008.

Henning Piezunka and Linus Dahlander. Distant search, narrow attention: How crowding alters organizations' filtering of suggestions in crowdsourcing. *Academy of Management Journal*, 58(3):856–880, 2015.

Marc Pilkington. 11 Blockchain technology: principles and applications. *Research handbook on digital transformations*, page 225, 2016.

Sam Ransbotham, Gerald C. Kane, and Nicholas H. Lurie. Network characteristics and the value of collaborative user-generated content. *Marketing Science*, 31(3):387–405, 2012. doi: 10.1287/mksc.1110.0684. URL https://doi.org/10.1287/mksc.1110.0684.

Jakob Gulddahl Rasmussen. Bayesian inference for hawkes processes. *Methodology and Computing in Applied Probability*, 15(3):623–642, September 2013.

Raul. Transactions speeds: How do cryptocurrencies stack up to visa or paypal?, 2018. https://howmuch.net/articles/crypto-transaction-speeds-compared. Accessed 30 June 2018".

P. Reynaud-Bouret, V. Rivoirard, and C. Tuleau-Malot. Inference of functional connectivity in neurosciences via hawkes processes. In *2013 IEEE Global Conference on Signal and Information Processing*, pages 317–320, 2013.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 55(5):999–1033, 1987.

Stephen P. Ryan. The costs of environmental regulation in a concentrated industry. *Econometrica*, 80(3): 1019–1061, 2012.

Joseph A Schumpeter. The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle. 1934.

Sonali K. Shah and Mary Tripsas. The accidental entrepreneur: the emergent and collective process of user entrepreneurship. *Strategic Entrepreneurship Journal*, 1(1-2):123–140, 2007. doi: 10.1002/sej.15. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sej.15.

Peter Sims. The No. 1 Enemy of Creativity: Fear of Failure. https://hbr.org/2012/10/the-no-1-enemy-of-creativity-f, 2012. Accessed April 20, 2019.

Param Vir Singh, Yong Tan, and Nara Youn. A hidden markov model of developer learning dynamics in open source software projects. *Information Systems Research*, 22(4):790–807, 2011. doi: 10.1287/isre.1100.0308. URL https://pubsonline.informs.org/doi/abs/10.1287/isre.1100.0308.

Jimmy Song. Bitcoin Cash Difficulty Adjustments. https://medium.com/@jimmysong/bitcoin-cash-difficulty-adjustments-2ec589099a8e, 2017a. Accessed: 2019-03-04.

Jimmy Song. Why Miners Are Mining Bitcoin Cash – and Losing Money Doing It. https://www.coindesk.com/miners-mining-bitcoin-cash-losing-money, 2017b. Accessed: 2019-03-04.

Jimmy Song. Mining BTC/BCH: Past, present and future. https://bitcointechtalk.com/mining-btc-bch-past-present-and-future-e9432315aa47, 2017c. Accessed: 2019-03-04.

P. K. Srijith, Michal Lukasik, Kalina Bontcheva, and Trevor Cohn. Longitudinal modeling of social media with hawkes process based on users and networks. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, ASONAM '17, pages 195–202. ACM, 2017.

Sorawoot Srisuma. Minimum distance estimators for dynamic games: Minimum distance estimators for dynamic games. *Quantitative Economics*, 4(3):549–583, November 2013.

Mark Thorley. The role of failure in developing creativity in professional music recording and production. *Thinking Skills and Creativity*, 30:160–170, 2018.

Olivier Toubia and Oded Netzer. Idea generation, creativity, and prototypicality. *Marketing Science*, 36(1):1–20, 2017.

Alejandro Veen and Frederic P Schoenberg. Estimation of space–time branching process models in seismology using an em–type algorithm. *Journal of the American Statistical Association*, 103(482):614–624, 2008.

Eric von Hippel. Lead users: A source of novel product concepts. *Management Science*, 32(7):791–805, 1986. doi: 10.1287/mnsc.32.7.791. URL https://doi.org/10.1287/mnsc.32.7.791.

Eric Von Hippel. The sources of innovation. Oxford Univ. Press, 1988.

Robert W. Weisberg. *Creativity and Knowledge: A Challenge to Theories*, page 226–250. Cambridge University Press, 1998.

Robert W. Weisberg. On "Out-of-the-Box" Thinking in Creativity. In *Tools for Innovation*. Oxford University Press, Oxford, 2009.

Joel West and Karim R. Lakhani. Getting clear about communities in open innovation. *Industry and Innovation*, 15(2):223–231, 2008.

Joel West and Siobhán O'mahony. The role of participation architecture in growing sponsored open source communities. *Industry and Innovation*, 15(2):145–168, 2008.

N Wishbow. Creativity in poets. *Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, Pennsylvania*, 1988.

Lizhen Xu, Jason A. Duan, and Andrew Whinston. Path to Purchase: A Mutually Exciting Point Process Model for Online Advertising and Conversion. *Management Science*, 60(6):1392–1412, June 2014.

Shuang-Hong Yang and Hongyuan Zha. Mixture of mutually exciting processes for viral diffusion. In Sanjoy Dasgupta and David McAllester, editors, *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 1–9, Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR.

Lixiu Yu, Aniket Kittur, and Robert E. Kraut. Distributed analogical idea generation: inventing with crowds. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pages 1245–1254, Toronto, Ontario, Canada, 2014. ACM Press.

Appendices

.1 Distributional Assumptions for Spatial Effect

While gamma/exponential density function has been used to model temporal decay in earlier studies, we experiment with different distributional forms for the spatial effect. We model the impact of conceptual/spatial distance between the topics of the focal challenge and the prior knowledge acquisition activities as a realization of a random variable. Since distances are positive-valued, we experiment with different distributions which have support on $\mathbf{R}^{+}$. We select the most appropriate distribution for this random variable based on the AIC values of the estimated model. We find weibull distribution leads to a minimum AIC value and the second best model has a relative likelihood of 1.85×10^{-06} with respect to our best model. We present the exact form of all the density functions below. For $f^{weibull}$ and f^{gamma} density functions, k and θ are the shape and scale parameters respectively. For f^{beta}, k and θ are the shape parameters. For $f^{lognormal}$, k and θ are the mean and standard deviation parameters.

$$f^{weibull}(x;k,\theta) = \begin{cases} \frac{k}{\theta}\left(\frac{x}{\theta}\right)^{k-1} e^{-\left(x/\theta\right)^{k}} & x \geq 0, \\ 0 & x < 0 \end{cases} \qquad f^{beta}(x;k,\theta) = \begin{cases} \frac{\Gamma(k+\theta)}{\Gamma(k)\Gamma(\theta)} x^{k-1}(1-x)^{\theta-1} & x \geq 0, \\ 0 & x < 0 \end{cases}$$

$$f^{lognormal}(x;k,\theta) = \begin{cases} \frac{1}{x\theta\sqrt{2\pi}} e^{-\frac{(\log x - k)^2}{2\theta^2}} & x \geq 0, \\ 0 & x < 0 \end{cases} \qquad f^{gamma}(x;k,\theta) = \begin{cases} \frac{x^{k-1}e^{-\left(x/\theta\right)}}{\theta^{k}\Gamma(k)} & x \geq 0, \\ 0 & x < 0 \end{cases}$$

Table 1: AIC values for models with different spatial distributions

Model	AIC
Gamma-Weibull	-214328.2
Gamma-Beta	-214301.8
Gamma-Lognormal	-214276
Gamma-Gamma	-213959

Each model name represents the temporal density function - spatial density function. We use gamma density function for all models for temporal decay.

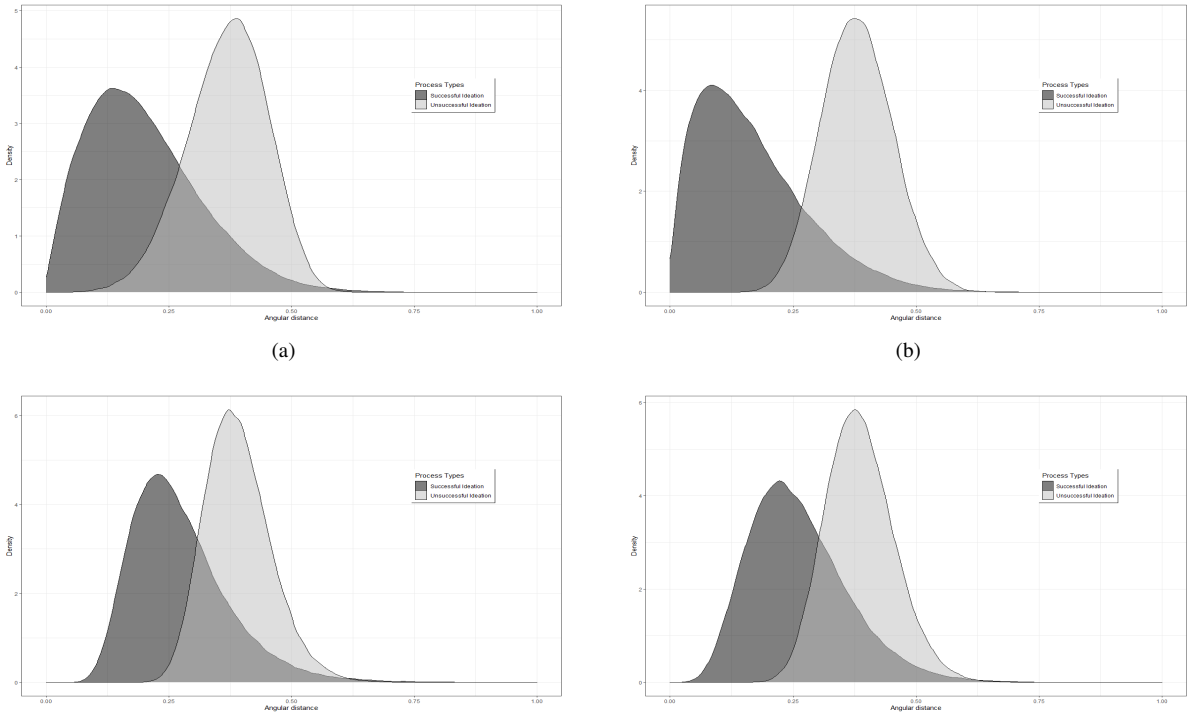


Figure 1: We present the density plots for the effect of spatial distance on ideation efforts. (a) Gamma-Weibull Model. (b) Gamma-Beta Model. (c) Gamma-Lognormal Model. (d) Gamma-Gamma Model.

.2 Neural embeddings for text vectors

In this section, we discuss the natural language processing (NLP) method that was employed to generate the numerical representation of the ideas and challenges posted on the open innovation platform. Extant NLP research for downstream usage of neural embeddings makes a distinction between semantic (e.g. word similarity, sentiment analysis) and syntactic (e.g. part-of-speech tagging) understanding. While shorter vectors are found to contain more syntactic knowledge i.e. about the grammatical structures, high dimensional vectors are found to be much more suitable for semantic knowledge i.e. about the underlying meaning of the words/documents. Majority of research papers use vectors of length 300 due to diminishing returns on vectors with higher dimensionality (Pennington et al., 2014). Moreover, research has found semantic understanding plateaus for vectors larger than 300 (Melamud et al., 2016).

We follow the machine learning literature to decide on the optimal length of the numerical vectors via a grid search. We craft our creativity problem as a classification task i.e. we classify ideas as success or failure in creativity tasks using the numerical representations as input features. We split our data into training and testing set using a stratified (at a challenge level) 80:20 split. For every possible configuration of the hyper-parameters (namely, length of the vector and window size), we create a new paragraph vector model using the training set data, train a logistic regression model using this training set data, infer vectors for the testing set, and then test the performance of our classifier on the test set (Le and Mikolov, 2014) using the F1 score as a measure of accuracy. We use F1 score as our evaluation metric as it takes both precision and recall/sensitivity into consideration.

Through this empirical process, we find that the paragraph vectors with length 300 and window size 7 lead to the best performance with a F1 score of 0.41. The F1 scores for all configurations are given in Table 2.

Table 2: F1 scores for different vector length- contextual window size pairs.

Vector Sizes	Contextual Window Sizes						
	6	7	8	9	10	11	12
200	0.343	0.347	0.347	0.348	0.340	0.345	0.376
250	0.345	0.331	0.333	0.398	0.334	0.339	0.329
300	0.343	0.410	0.380	0.346	0.372	0.359	0.335
350	0.351	0.340	0.374	0.372	0.359	0.307	0.332
400	0.397	0.378	0.321	0.344	0.363	0.335	0.367

.3 Numerical Approximation for Equation (3.13)

The integration in the likelihood function has a specific interpretation which forms the basis of our numerical approximation. The integral represents the likelihood of not observing the events at any other points (in time and space) except their actual realizations. This requires us to know the bounds of their temporal and spatial horizons. In a single dimension, the integral over time is easy to calculate and has a closed form solution (1). However, with the introduction of the spatial component where the dimensionality of the space is high, we run into intractable computational issues due to lack of closed-form solutions as well as the time required to

numerically approximate through simulations.

$$\begin{aligned}\int_0^T \lambda_z^h dt &= \int_0^T \phi_{kj}^h (t - t_m^j) dt \\ &= \int_0^T \alpha_{kj} f_k^{temporal} (t - t_m^j) dt \\ &= \alpha_{kj} F_k^{temporal} (T - t_m^j; \Phi) \end{aligned} \tag{1}$$

We overcome these by using the conceptual understanding of this integral. Given the activity stream of knowledge acquisitions and idea submissions, we can observe the range of conceptual distances that the ideator could have possibly engaged in or actually engaged in as we observe all the challenges that were accepting ideas during the tenure of the ideator on the platform. Therefore, we can observe, for every acquired knowledge, the lowest (d_{min}) and highest (d_{max}) conceptual distance from the set of all possible challenges that the ideator participated in or could have participated in (but chose not to). Now, the multidimensional integral problem is reduced to a single dimension over conceptual distances, instead of the numerical vector coordinates, where the integral represent the likelihood of not observing any other idea submissions in challenges whose conceptual distance from the aforementioned acquired knowledge $\in [d_{min}, d_{max}]$.

.4 First Stage Results

.4.1 Policy Estimates

We estimate log-linear models for the policy functions governing mining BTC and BCH blocks in the first stage. Specifically, we model $\log(n_{i,t,btc}+1)$ and $\log(n_{i,t,bchc}+1)$ as a flexibly-specified function of $W_{i,t,btc}$ and $W_{i,t,bch}$ respectively. Results for BTC and BCH miners, with different specification, are presented in Table 3 and 4 respectively.

.4.2 Estimates for transition equations

The transition equation parameters (Equations 4.11, 4.12, 4.13, 4.14) are given in Table 5. We estimate different transition equations for BTC price before and after the fork. Prior to the fork, BTC price innov-

Table 3: First Stage Policy Estimates for BTC miners

	I		II	
Parameter	Point Estimate	Standard Error	Point Estimate	Standard Error
Intercept	2.5024	1.4536	3.0132	0.8626
Price$_{btc}$	0.2253	0.0508	0.2236	0.0301
Difficulty$_{btc}$	−0.1097	0.0611	−0.1068	0.0363
Fees$_{btc}$	−0.0740	0.0219	−0.0742	0.0130
Price Ratio	0.0002	0.0217	−0.0017	0.0128
Difficulty Ratio	−0.0038	0.0202	−0.0017	0.0120
EDA	−0.0290	0.0993	−0.0247	0.0589
Fork	−0.0461	0.0967	−0.0503	0.0574
EDA × Price Ratio	0.0187	0.0218	0.0206	0.0129
EDA × Difficulty Ratio	−0.0163	0.0206	−0.0184	0.0122
Lagged BTC Total	0.0038	0.0013	0.0038	0.0008
Bitcoin.com	$n.a$	$n.a$	−1.6210	0.0202
BTC.com	$n.a$	$n.a$	−0.4225	0.0202
BTC.TOP	$n.a$	$n.a$	−0.3630	0.0202
F2Pool	$n.a$	$n.a$	−0.6880	0.0202
Other	$n.a$	$n.a$	0.6592	0.0202
Unknown	$n.a$	$n.a$	−1.4723	0.0202
ViaBTC	$n.a$	$n.a$	−0.7107	0.0202
AdjustedR^2	0.01604		0.6536	

Table 4: First Stage Policy Estimates for BCH miners

	I		II	
Parameter	Point Estimate	Standard Error	Point Estimate	Standard Error
Intercept	-3.0784	2.1213	-4.3147	1.4977
Price$_{bch}$	-0.0576	0.0739	-0.0773	0.0521
Difficulty$_{bch}$	0.1797	0.0762	0.2192	0.0538
Fees$_{bch}$	-0.0204	0.0094	-0.0172	0.0066
Price Ratio	-0.1596	0.0284	-0.1762	0.0201
Difficulty Ratio	0.1628	0.0277	0.1772	0.0195
EDA	5.6938	2.7972	6.0828	1.9736
EDA $\times$ Price Ratio	0.0681	0.0304	0.0828	0.0215
EDA $\times$ Difficulty Ratio	-0.0733	0.0299	-0.0858	0.0211
Lagged BCH Total	0.0048	0.0003	0.0052	0.0002
BTC Count	-0.0815	0.0030	-0.0131	0.0034
EDA $\times$ Price$_{bch}$	0.0238	0.1125	0.0618	0.0794
EDA $\times$ Difficulty$_{bch}$	-0.2300	0.1052	-0.2538	0.0742
Bitcoin.com	$n.a$	$n.a$	-0.0984	0.0416
BTC.com	$n.a$	$n.a$	-0.2174	0.0356
BTC.TOP	$n.a$	$n.a$	0.1241	0.0360
F2Pool	$n.a$	$n.a$	-1.0057	0.0379
Other	$n.a$	$n.a$	-0.8876	0.0392
Unknown	$n.a$	$n.a$	1.6199	0.0399
ViaBTC	$n.a$	$n.a$	0.4714	0.0366
AdjustedR^2	0.2918		0.6474	

ations, according to (4.11), are distributed as $\mathcal{N}(0, 0.0005)$ with AR term estimated as 0. Post-fork, BTC price innovations are distributed as $\mathcal{N}(0, 0.0011)$. Innovations for BTC fees, BCH price, and BCH fees are distributed as $\mathcal{N}(0, 0.0438)$, $\mathcal{N}(0, 0.0040)$, and $\mathcal{N}(0, 0.2801)$ respectively.

Table 5: First stage transition equation parameters.

Parameters	Point Estimates	Standard Errors
	BTC Price	
$\zeta_{0,0}$	-0.0880	0.0354
$\zeta_{0,1}$	0.0000	0.0000
	BCH Price	
$\zeta_{1,0}$	-0.3950	0.1360
$\zeta_{1,1}$	0.2658	0.1378
	BTC Fees	
$\zeta_{2,0}$	0.4785	0.0550
$\zeta_{2,1}$	-0.7966	0.0393
$\zeta_{2,2}$	-0.0143	0.0007
	BCH Fees	
$\zeta_{3,0}$	0.3653	0.0726
$\zeta_{3,1}$	-0.8029	0.0528
$\zeta_{3,2}$	-0.0128	0.0006