Essays on Digital Transformation and Business Decision-Making

Zhijin Zhou

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

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Yong Tan, Chair

Ming Fan

Yingfei Wang

Program Authorized to Offer Degree:
Business Administration

University of Washington

Abstract

Essays on Digital Transformation and Business Decision-Making

Zhijin Zhou

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

The rapid advancements of digital technologies have fostered transformation of various online and offline environments, which in turn created opportunities and challenges on modern management and decision-making strategies. In this dissertation, I examine three areas that are subject to the impact of such transformation, using evidence-based inference and optimization approaches to gain insights and enhance decision-making quality. The first essay examines individual investors' learning in crowdfunded supply chain finance (SCF) markets under the unique presence of loan guarantors in the financing process. My estimation results confirm the existence of investor learning. In addition, I observe that this latent perception has different moderating effects on investor responses to listing attributes, such as interest rate and loan duration. Counterfactual simulations suggest that enabling investor learning from correlated investment experience can help mitigate adverse selection and improve overall market efficiency; for supply chain members, optimizing the structure of loan listings could accelerate investor learning, which in turn can help simulate fundraising performance as a desirable outcome of reputation building. The second essay investigates the effect of scarcity-Induced demand on the crowdfunding market. Using a hierarchical Bayesian framework, this essay empirically validates the positive effect of the scarcity strategy in the crowdfunding market; Interestingly, my mechanism analysis reveals that individual backers are more susceptible to demand-induced scarcity: given the relative scarcity level, individuals are more

attracted to invest in rewards that achieve this scarcity due to excess demand rather than limited supply. In my third essay, I use data-driven analytics to facilitate medical decision-making based on electronic medical records (EMRs). Specifically, I consider the problem of designing personalized treatment recommendations for patients with multiple myeloma, which is the second most common blood cancer in the United States. Using clinical data with patients' cytogenetics information, this essay formulates the treatment recommendation problem as a multilevel Bayesian contextual bandit, which sequentially selects treatments based on contextual information about patients and therapies, with the goal of maximizing overall survival outcomes. I also propose a causal offline evaluation framework which integrates the structural econometric model into bandit optimization and generates counterfactuals for policy evaluation. This novel framework provides reliable performance measures when field experiment or long log data are not available. Compared with clinical practices and benchmark strategies, my method suggests a rise in overall survival outcomes, with higher improvement for aging or high-risk patients with more complications.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to take this opportunity to express my deepest gratitude to my advisor, Professor Yong Tan, for educating, guiding, and supporting me over the past five years. I have been blessed to have a supervisor who cared so much about my work and growth, valued me as a researcher, and provided me with unconditional support and encouragement in critical times throughout my journey. I would forever feel lucky to have the chance to encounter him, become his student, and I hope to emulate his open-mindedness, creativity, scientific rigor, carried over the intellectual independence that he fostered to my own career forward.

I would like to extend my appreciation to my committee members, Professor Jing Tao, Professor Ming Fan, Professor Yingfei Wang, and Professor Shi Chen, for their guidance, insightful comments, constructive feedback through my research study.

My special thanks also go to Shawna Reimers, Jaime Banaag, Beau Kirkeby, and other staff at the Foster School of Business, who provide consistent support and caring during my doctoral studies. I also want to thank my peers in the Mackenzie Hall, and also my academic families: Zixuan Meng, Vipul Aggarwal, Jiaying Deng, Kyungmin Park, Tongxin Zhou, Yang Jiang, Yi-Chun Ho. I enjoy every discussion with you, and I am grateful to have your support, advice, help and accompany in the journey.

Finally, I would like to thank my family, my friends, for their unconditional support, love, encouragement in every step along the way. I could not have made it this far without their support.

Chapter 1

INTRODUCTION

The rapid advancements of digital technologies have fostered transformation of various online and offline environments. Instead of simply enhancing and supporting traditional methods, such transformations may also empower new types of creativity and business model innovation, facilitating new markets, advocate novel decision-making strategies. These transformations are fundamentally changing the environment ecosystems, which in turn created opportunities and challenges on modern management and decision-making strategies. In this dissertation, I investigate three areas that are subject to the impact of such transformations.

In Chapter 3, I investigate the business model innovation enabled by digitization. In Chapter 3, I study a novel technology-driven Fintech initiative, online Supply Chain Finance (SCF) [1], trying to understand how it transforms individual investors' decision-making compared to traditional microloan markets. In this study, I focus on a unique platform policy wherein a fundraiser is a supply chain that involves three firms: an upstream supplier, a downstream buyer, and a large-scale enterprise referred to as a core firm. Three supply chain members work as a single business entity to request a loan on behalf of the supplier facing a liquidity problem. Under this contract, the core firm acts as the guarantor of a loan; it may step in and make repayments to investors in an event of fundraiser defaults. Since the core firm of a supply chain tends to have well-established reputation and reliability, it usually serves as the guarantor of multiple loans in the crowdfunded SCF markets.

We posit that sequential investor-fundraiser interactions can give rise to investor learning of the guarantor reliability – a dynamic individual behavior that has not been examined in the crowdfunding literature. Understanding how such learning behavior governs investment

[1] https://www.investopedia.com/terms/s/supply-chain-finance.asp

decisions generate useful implications for platform designers who desire to encourage investor participation and stimulate fundraising performance. More formally, I aim to answer the following questions: 1) Do investors exhibit learning toward the reliability of loan guarantors as they have more interactions with each other over time? 2) If so, how would such dynamic learning govern individuals' investment behavior? 3) How can supply-chain fundraisers better design their loan listings to boost the fundraising performance?

Drawing on the Bayesian learning framework, we consider that investors learn about the guarantor reliability via a dynamic process. When a new investor joins the platform, she forms an initial reliability perception toward various guarantors and makes investment decisions based on such perceptions and other listing and individual characteristics. Upon receiving scheduled loan repayments in which a guarantor is involved (or lack thereof), the investor updates her reliability perception of that guarantor and then uses the updated perception for subsequent decision-making opportunities. To better understand how this learning mechanism governs investment behavior, we model an investor's decisions of whether to invest (referred to as incidence decision) and how much to invest (referred to as amount decision) as two separate yet correlated processes. Taken together, these modeling efforts allow us to empirically investigate how perceptions of guarantor reliability evolve over time and, more importantly, how this time-variant component dynamically affects investment behavior at different decision-making stages.

Our analysis yields several interesting findings. First, we find supportive evidence that individuals' investment decisions are, at least partially, driven by their perceptions of guarantor reliability. Guarantors' perceived reliability is found to have dissimilar effects at different decision-making stages. To be more specific, while the reliability perception has a positive impact on both individual decisions of whether to invest and how much to invest, the impact exhibits a marginal diminishing effect on the investment incidence decision. Furthermore, our results show that individual perception of guarantor reliability has a moderating effect on how investors respond to listing attributes. In particular, when investors perceive the guarantor of a given listing to be more reliable, they tend to pay less attention to the inter-

est rate when picking a listing to invest, but tend to invest a larger amount in a listing with a longer duration.

In Chapter 4, I examined the effect of scarcity strategy, which is popular among advertisers and retailers in various offline markets to promote sales, in the context of online reward-based crowdfunding platforms. Advertisers and retailers usually base their promotional strategy on the scarcity principal, using restrictions as an incentive. For example, some marketers promote their products by limiting the offering time, while others use phrases such as "limited edition" or "limited number" to restrict the availability of their products. These restrictions are aimed at increasing consumers' perceived scarcity of the products. Previous studies have documented increased desirability among consumers to purchase in various offline markets by the market's adopting such a marketing strategy and, in particular, by limiting information about the product (Eisend 2008; Jung and Kellaris 2004).

The scarcity strategy has become pervasive in reward-based crowdfunding markets, which, as an alternative source of financing for start-up ventures, recently have experienced explosive growth. To attract more backers and to facilitate the success of their fundraising campaigns, many fundraisers on these platforms have sought to integrate the scarcity strategy into their reward scheme design. These raisers put quantity restrictions on some of their reward options to limit the number of backers who can take advantage of the offering, with an attempt to act on individuals' cognitive process to render their judgment favorable in regard to the focal reward.

Although the implementation of the scarcity strategy on crowdfunding platforms is garnering increased attention, doubts about its effectiveness exist. On the one hand, it is likely that such a tactic would lead to value enhancement: although individuals may be unaware of the scarcity of a product in various offline markets, the supply and demand changes are highly transparent on these online platforms, making the scarcity of a product a more salient feature. On the other hand, the link between individuals' perceived scarcity and their corresponding backing intention, in the presence of information asymmetry and high investment uncertainty, has not been established. Although some preliminary studies have examined the

effect, they did not consider how campaign characteristics affect backers' choice behaviors (Joenssen and Müllerleile 2016) or were conducted under conditions (e.g., lab experiment with a small sample size) that leaves unclear the generalizability and applicability of their findings to real crowdfunding markets.

To understand to what extent does the scarcity strategy influences backers' investment intention in the reward-based crowdfunding markets, and the underlying motivational mechanism that gives rise to such an effect, we collected dynamic panel data from JD.com, one of the leading reward-based crowdfunding platforms in China. On this platform, fundraisers can choose to limit the total availability of each reward option to a certain amount; once the number of backers that a reward has attracted reaches the pre-set upper limit, that reward becomes immediately unavailable for any future backers. For all of the observed 617 ongoing campaigns, we collected information on their static characteristics and tracked the dynamic fundraising statistics for all reward levels, resulting in a series of 2,938 continuous observations, where each series represents the dynamics for a distinct reward option.

We start our analysis with a zero-inflated negative binomial (ZINB) model to examine how the scarcity of a reward affects backers' investment intention, as measured by the total number of attracted backers within a day. The result of this model presents primary evidence of the scarcity effect in the crowdfunding market. Next, we extend our baseline model to a hierarchical Bayesian framework to address the unobserved nested reward heterogeneity, and we detail our identification strategy to empirically uncover the underlying mechanism that governs individuals' evaluation process in the presence of quantity scarcity.

Our results suggest that backers are prevention-motivated on the JD crowdfunding platform. Given the relative scarcity of a reward (i.e., its current stock level relative to its total supply), a larger number of predecessors amplifies the attraction of that reward option. In contrast, a reward that is scarce due to limited supply, characterized by fewer previous backers given the relative scarcity, is less preferred. This result suggests that backers' valuation of a reward choice increases with the number of existing adopters, echoing the evidence of network externality in individuals' adoption decisions in regard to high-technology products

(Tucker 2008). This result also indicates that, facing great investment uncertainty and information asymmetry, avoiding making bad decisions and minimizing risk are the principal considerations when backers are deciding which product to contribute to in the crowdfunding markets. In this case, their desire for uniqueness and exclusiveness becomes subordinate. We prove the appropriateness and fitness of our full model by comparing it with its three nested specifications, and we further validate our findings through additional robustness checks.

In Chapter 5, I study how to use data-driven analytics approach to facilitate personalized medical decision-making. Personalized treatment design and, more specifically, precision medicine, aims to provide the right intervention to the right patient at the right time by taking into consideration patients' heterogeneity in needs and treatment responses. With the wide availability of electronic medical records (EMRs) on both population and individual level information, it becomes possible to transform raw data into optimal policies for health interventions.

In this study, we consider the problem of designing personalized treatment recommendations for patients with multiple myeloma. When a physician makes the treatment decision in the medical practice, an assumption is made about the efficacy or toxicity of the therapy based on extrapolation from similar treatments, or prior experience administering the treatment to similar patients. While the physician would always try to treat a patient as efficiently as possible by selecting the treatment combination that appears the best (exploitation), it could in fact be suboptimal due to imprecise knowledge. Alternatively, randomized clinical trials (RCTs) can be performed to gain knowledge on the efficacy of alternative treatments combinations (exploration). However, given that the number of available treatments approved by FDA has surged over the past decade, conducting RCTs on all available treatment combinations are extremely expensive, and at the cost of sacrificing trial patients' survival outcomes by selecting a sub-optimal combination during the trials (Bastani et al. 2017, Bertsimas and Vayanos 2017). Moreover, it is difficult to synthesize all efficacy evidence for direct comparisons of all available treatment combinations (van Beurden-Tan et al. 2017). Hence, a critical question raised by the healthcare community is how to quickly identify the

treatment that gives the best patient response from a large number of alternatives (exploration/learning), while treating patients as effectively as possible (exploitation/earning).

We propose formulating the physician treatment decision-making process as a Bayesian contextual bandit that combines the multilevel Bayesian linear model with TS based treatment selection rule for generating personalized recommendation. In the multi-armed bandit (MAB) model, the algorithm is faced with a set of possible actions called arms and seeks to balance the exploration-exploitation trade-off by sequentially selecting actions to maximize the cumulative reward. In our research context, the objective is to find the optimal treatment selection rule (i.e., the policy) that sequentially assigns treatments (i.e., the arms) to patients, which maximizes their cumulative survival outcome (i.e., the reward). We build the treatment selection rule based on the Thompson sampling (TS) algorithm, randomly drawing each arm according to its probability of being optimal (Thompson 1933). This algorithm also takes an online view, continuously using accumulating data to update our model as new patients become available.

Using the proposed offline causal evaluation approach, we evaluate the performance of proposed model on clinical data collected from 803 patients treated at Seattle Cancer Care Alliance (SCCA). Our experiment results suggest that by balancing the exploration-exploitation trade-off using our proposed TS policy, we are able to achieve a more efficient treatment recommendation at the individual level by considering patient disease dynamics, cytogenetic variants, and unobserved heterogeneity in treatment responses. This advantage is significantly more remarkable for high-risk sub-populations, whose disease usually involves more complications that cannot be appropriately treated with conventional therapies. Moreover, through the counterfactual analysis from the structural econometric model, we validate the value of the TS algorithm in fast learning by comparing it to other widely-used MAB algorithms. This is critical in healthcare applications where the clinical data is usually sparse and limited and where ethical issues prevent arbitrary clinical trials without careful empirical evidence from observational data.

Chapter 2

LITERATURE REVIEW

I study the impact of digital transformations on three different contexts. In this section, I review literature in related fields, and illustrate how my work extends the understanding of the previous studies. Section 2.1 shows the related literature in microloan lending, and investor learning under uncertainty. Section 2.2 reviews the studies on online reward-based crowdfunding, market mechanisms, reward design strategies and scarcity effect. Section 2.3 presents the literature in data-driven healthcare analytics, precision medicine, and multi-armed bandit for adaptive decision-making.

2.1 Supply Chain Finance and Individual Learning Under Uncertainty

Supply Chain Finance (SCF) is a technology-driven financing service,[1] wherein supply chain members collaborate to optimize the working capital within the chain (Wuttke et al. 2016). In a typical supply chain setting, suppliers issue an invoice to buyers upon order delivery, collecting accounts receivable instead of immediate payments. Such payment delays may give rise to liquidity problems for upstream suppliers who are usually small businesses with relatively limited cash flow and creditworthiness (Tunca and Zhu 2017). To deal with this issue, supply chain members can reach an SCF agreement under which a thirtd-party funder (e.g., a commercial bank) is involved to provide liquidity to both buyers and sellers. SCF has been proved an effective contracting practice, with the ability to stabilize a supply chain and strengthen the relationship between buyers and suppliers (Lekkakos et al. 2016, Zhang et al. 2019a).

With the surging popularity of online crowdfunding, an innovative "Fintech" initia-

[1] https://www.investopedia.com/terms/s/supply-chain-finance.asp

tive—termed online crowdfunded SCF – has emerged. Compared to traditional SCF services, the crowdfunded SCF transforms the financial flows, allowing a crowd of individual investors to collectively serve as a funder. By eliminating financial intermediaries that often demand a high service fee, crowdfunded SCF platforms not only lower the financing costs (via lower interest rates) for supply chain members but provide alternative opportunities for crowd investors.

Despite several merits, the crowdfunded SCF presents a tough challenge to individual investors as well. Unlike large-scale lending institutions, individuals may not be sophisticated in evaluating the default risk of a fundraiser (Zhang and Liu 2012). Such information asymmetry (Allon and Babich 2020) leads to high uncertainty on investors' financial returns, dampening their willingness to participate in the market. To combat information asymmetry, crowdfunding platforms implement various platform designs and policies for information disclosure and value creation (Allon and Babich 2020, Chen et al. 2020).

In crowdfunded SCF markets, a fundraiser is a supply chain that involves three firms: an upstream supplier, a downstream buyer, and a large-scale enterprise referred to as a core firm. In ordinary crowdfunding contexts (e.g., peer-to-peer lending), an investor rarely has a chance to transact with the same fundraiser more than once as an overwhelming majority of fundraisers are one-time capital seekers. As a consequence, there are fewer opportunities for fundraisers to build reputation (Agrawal et al. 2014) and investor trust. In the crowdfunded SCF market, however, an investor may fund multiple loans over time that are actually guaranteed by the same core firm. Given that crowd investors may not be sophisticated in assessing the risk of a loan listing (Allon and Babich 2020), the shared attribute across listings – i.e., the same guarantor – provides useful information based on which investors can draw inferences of a guarantor's reliability. In other words, guarantors can bear reputation, and their presence may lead to sequential correlation among a series of decisions by the same investor. This unique feature distinguishes the crowdfunded SCF market from ordinary crowdfunding settings where one's decisions are considered relatively independent.

Our research presents several novelties. First, we examine a novel operational FinTech

solution that aims to solve the liquidity problem confronting businesses situated in a supply chain setting. To the best of our knowledge, such technology-enhanced SCF services has not yet been explored in the OM research or, more broadly, the vast crowdfunding literature. In contrast to much of the prior work focusing on ordinary crowdfunding contexts where fundraisers are often one-time money requesters (Lin and Viswanathan 2016a, Zhang and Liu 2012), our work explores a unique setting in which investors may interact with the same fundraiser (i.e., the same guarantor to be most precise) for multiple times. This platform policy feature allows us to empirically investigate whether investors undergo a learning process as they interact with the same guarantors over time.

The second novelty of this research is that we conceptualize and explicitly model individual perception of the guarantor reliability as a subjective attitude underlying one's perceived risk of a loan listing. Instead of treating this latent perception as a homogeneous and static construct, we consider that individuals hold heterogeneous perceptions and update their beliefs based on the private information signals they receive. Such modeling efforts allows us to gain more insights into how this time-varying construct governs individuals' investment decisions in a dynamic fashion.

Lastly, we consider individual investment behavior as a joint process consisting of two separate yet interdependent decisions: the incidence and amount decision. We model and estimate the factors that affect these two decisions simultaneously, while accounting for the potential interdependence between them. This distinguishes us from previous work which only look into one specific investment outcome. Moreover, we add a hierarchical structure to the Bayesian learning model to obtain individual-specific parameters. With this rich specification, we are able to uncover the learning dynamics and the role of guarantor reliability at a more granular, investor level.

Our paper also contributes to the growing body of research on crowdfunding in operations management literature. Allon and Babich (2020) and Chen et al. (2020) review the existing literature and illuminate several directions for advancing our knowledge in crowdfunding markets, among other surging online platforms. Recently, operations management

scholars have shown interests in better understanding 1) the impact of market designs of a crowdfunding platform (e.g., Belavina et al. (2020), Fatehi and Wagner (2019), Strausz (2017)) and 2) how fundraiser should design their listings to maximize success probability and revenue (e.g., Chakraborty and Swinney (2021)). Nonetheless, there is lesser empirical work that studies how platform policies impact user interactions and subsequent investor behavior (Burtch et al. 2020). In this stream of literature, much of extant work examines campaign success based on cross-sectional variations, which abstracts away from the dynamic nature of the crowdfunding process. We differentiate ourselves from prior research by leveraging time-series data to uncover investor learning over time. As a result, we are able to offer new insights into the role of loan guarantors in shaping investor behavior in a dynamic process. where there exists shared information across individual loans, leading to individual investors' learning process under uncertainty and state dependence among their behavior dynamics over time.

2.2 Online Crowdfunding Markets and Scarcity Strategy

In Chapter 4, I examine the effect of scarcity strategy, in the context of online reward-based crowdfunding platforms. This essay is closely related to a growing literature on online crowdfunding. Crowdfunding, according to Benna and Benna (2018), is "a collective effort of many individuals who networked and pooled their resources to support efforts initiated by other people or organizations" (21). Based on the types of returns that are given after individuals' contributions, these crowdfunding platforms can be further classified into four types: donation-based, reward-based, lending-based, and equity-based (Hemer 2011).

Among the four types of crowdfunding platforms, particularly relevant to our study is the reward-based crowdfunding platform. These platforms provide a forum for any venture with an innovative idea to pitch to vast individual backers. Any backer who is interested in supporting the idea can provide capital support to the fundraiser in exchange for various rewards. This new business mechanism creates a win-win situation for both parties. Campaigners have more flexibility and less cost when receiving funding from a large num-

ber of people with small monetary contributions, while individual backers can access the products of interest as early consumers who enjoy the most favorable price as well as other exclusive benefits. Although this sounds appealing, individual backers may suffer from considerable uncertainty on the crowdfunding platform when deciding whether to contribute to a proposed new product with unknown quality (Agrawal et al. 2014, Mollick 2014). Such uncertainty usually leads to perceptions of high risk, lowering their willingness to transact, which impedes the market growth and efficiency (Jarvenpaa et al. 1999).

A substantial body of information science literature focuses on ways to maintain the sustainability of these markets. One stream of research concerns the factors that motivate backers' engagement. In this vein, Zhang and Liu (2012) and Burtch et al. (2018) find that individuals' decisions are subject to peer influence. With imperfect information, they may assume that other people are more knowledgeable and may, thus, choose to follow their predecessors' actions. Bikhchandani and Sharma (2000) an overview of the recent theoretical and empirical studies on the herd behavior in the financial markets, summarizing the causes as well as the effect of such behavior on various financial markets. Jiang et al. (2018) support the notion of peer influence, finding the existence of herding at the platform level. In addition to peer influence, backers' decisions are subject to payoff externality. Under an "all or nothing" fundraising mechanism, whereby fundraisers can collect the raised funds only if the pre-specified fundraising target has been met, consumers may hesitate to make an investment in campaigns that have yet to receive enough funds, as they may fear the opportunity cost of their contributions. This argument has been empirically supported by previous studies (Zhang and Liu 2012, Burtch et al. 2018). Further, previous research also has shown that individuals may draw clues from the fundraisers' network structure (Liu et al. 2015) and that geographical proximity is an important consideration when individuals make decisions (Lin and Viswanathan 2016b).

Another stream of research focus on examining the influence of different market mechanisms and various reward design strategies on the fundraising outcomes, from the perspective of fundraisers and platform operators. For example, Wei and Lin (2016) compare two

different prevailing market mechanisms, auction and posted price, both theoretically and empirically. Their results show that, although the posted price mechanism causes fundraisers to be charged a higher interest rate, their fund requests are more likely to succeed. At the same time, it also is more likely that the fundraisers will default. Burtch et al. (2016) study the influence of information control policy in the crowdfunding platform, whereby users can conceal or disclose their private information to subsequent visitors and find that the concealing information will negatively influence successors to contribute. We bridge the above two streams of research, focusing on how the use of the scarcity strategy in the reward scheme design influences backers' investment intention. Scarcity, or limited availability, is commonly used by advertisers and retailers to promote sales. According to commodity theory, anything that is useful and can be conveyed to other people "will be valued to the extent that it is unavailable" (Brock 1968). Thus, by intentionally limiting consumers' ability to purchase, these marketers can manipulate perceived scarcity of the product, which, in turn, can alter consumers' judgment regarding the overall attractiveness of the offering (Inman et al. 1997).

It is important to note that quantity scarcity may be the consequences of two types of events: limited supply or excess demand. Although either could lead to the observed increased subjective desirability, they act on individuals' valuation processes in distinct ways (Verhallen 1982). Thus, depending on individuals' underlying motivational goal and the way they process information cues, their favorableness toward the product with these two types of scarcity may differ. One type of consumer cares more about responsibilities, duties, and security as well as the presence or absence of negative outcomes (Higgins 1998). Thus, their decisions are motivated by avoiding making mistakes or being left in an undesirable situation to the greatest extent possible. These consumers are prevention-motivated. Their behavior may be subject to the influence of conformity, which is defined as a desire to "fit in a group" to achieve a sense of security (Cialdini and Trost 1998). In addition, they may engage in "bandwagon reasoning," relying on peers' actions as the clue to infer the quality of the object (Worchel et al. 1975). The other type of consumer is motivated to obtain a sense of personal advancement through consumption, to whom we refer as promotion-motivated

consumers. Acquisition of products characterized by supply-induced scarcity would increase their uniqueness, distinguishing themselves from others. Thus, these individuals may show less caution or risk-aversion as compared to prevention-motivated consumers (Higgins 1998). A number of prior studies have documented favorable outcomes based on the implementation of such a tactic in various markets. DeGraba (1995) suggests that this strategy will induce a "buying frenzy" among consumers: When consumers perceive the product to be scarce, they may rush to purchase it before being fully informed or without deliberation. Balachander et al. (2009) examine the relationship between introductory levels and consumers' preference in the U.S. automobile industry and find that the relative scarcity of a car is positively associated with higher consumer preference for the car.

The above findings, however, cannot be directly generalized to the crowdfunding market wherein individuals usually rely on multiple information cues when making decisions. This extra information acts on individuals' valuation process in dissimilar ways, making the effects of product scarcity unclear. Among the very few studies that provide us with some initial insight, Joenssen and Müllerleile (2016) analyze the performance of over 40,000 projects in Indiegogo.com. Their results suggest that having at least one limited reward option reduces the probability of a campaign's success. They did not, however, control for any other variables, such as campaign attributes, that may influence backers' decisions, which makes it difficult for researchers to draw causal inferences from their results. Weinmann et al. (2017) conducted a controlled experiment with 166 online recruited participants. Contrary to the results of Joenssen and Müllerleile (2016), they find that including rewards with limited availability can be favorable under certain conditions. Their sample size is relatively small, however, and the design of their experiments does not mirror any real setting in the crowdfunding market to represent the empirical population, which limits the generalizability and applicability of their findings. Perhaps the most closely related prior work is that of Yang et al. (2017), who study how the reward option design, which in particular the decision whether to include a limited reward option, and the number of reward of options to limit, influence the fundraising outcome. Using a reduced form approach, they find that the

number of limited reward options has an inverted U-shaped relationship with fundraising performance.

While this essay is motivated by the scarcity effect in the offline setting and shares some similarities in terms of research context with the aforementioned studies in the crowdfunding community, we add valuable insights to the existing literature in the following notable ways. First, and most directly, we offer the first effort to explore the behavioral mechanisms underlies the scarcity effect in the crowdfunding context, thus it differs from the previous studies without an explanation for why such outcome occurs.

Second, when it comes to consider the influence of scarcity strategy, our dynamic panel dataset allow us to differentiate the degree of scarcity across different reward options, as well as within the same reward option at different time points throughout the fundraising cycle. This provides critical insight into how to design an effective reward scheme to best enhance backers' valuation of the product and motivate them to contribute, especially in the online setting where the supply and demand are highly transparent, and the degree of scarcity is dynamically changing over time. By contrast, all the previous studies use cross-sectional data, focusing on on identifying the existence an effect (Joenssen and Müllerleile 2016; Weinmann et al. 2017), or quantifying the impact of implementing such strategy on the overall fundraising performance (Yang et al. 2017) at the aggregate level.

Third, unlike all the aforementioned literature focuses on the campaign level fundraising outcome, this essay looks into individuals' choice dynamics at the reward level, which allows to control for the reward-level unobservables and their heterogeneous effects on backers' investment intention. Indeed, as we can see from the estimation results, we have identify reward-level heterogeneity even after controlling for the observed and unobserved campaign-level attributes. None of the existing studies have properly addressed such level of heterogeneity, which leads to potential endogeneity issue and limits the causal interpretation of their results. It also allows us to explore several factors that characterizing the fundraising environment at a more granular level – namely the within campaign and across campaign spillover effects, providing a more nuanced understanding of how these interactions influence

the outcome of interest.

2.3 Precision Medicine and Adaptive Medical Decision-Making

My third essay is closely related to the literature in the healthcare analytics, in particular to the growing number of studies using data-driven analytics to facilitate medical decision-making based on electronic medical records (EMRs). EMRs are real-time, patient-centered records that usually contain a patient's medical history, diagnoses, medications, laboratory test results, and all other information involved in a patient's care. With the embedded role of Information Systems (IS) in medical equipment, the EMRs have become much more widely accessible, giving researchers opportunities to gain insights that facilitate medical decision-making that was not possible previously.

Most of the early analytics studies shed light on medical decision-making at the population level. For example, using patient admission records and clinical data from 67 hospitals in the North Texas region, Bardhan et al. (2014) develop an analytics model to predict the re-admissions pattern of patients diagnosed with congestive heart failure, which allows hospitals to develop better predictive capabilities to profile patients who pose greater readmission risk. Based on the data from 414 clinical trials for advanced gastric cancer, Bertsimas et al. (2016) build statistical models to suggest regimens to be tested in different phases of clinical trials. While these models are useful for making decisions at the population scale, they are not suitable for individualized decision-making.

Personalized medicine is a promising direction for improving patients' survival outcomes because it tailors medical decisions to best benefit a patient based on their individual characteristics (van de Klundert 2016). This revolutionary paradigm has been punctuated by rapid development in the availability of new types and sources of information on individual characteristics, such as health conditions, treatment histories, and more particularly, cytogenetic profiles. Some of the promising examples of personalized medicine come from genetic-driven personalization, such as the use of tumor genomic profiling to improve clinical outcomes for patients with cancer (Garraway 2013). Another example is that genetic sequencing can

enable early diagnosis of inherited diseases for which targeted therapies are available (Soden et al. 2012).

In the medical literature, various approaches have been proposed to identify personalized treatment strategies. One stream of studies applies various statistical approaches such as sequential multiple assignment randomized trials (SMART) (Almirall et al. 2012, Murphy 2005), models in control theory (Rivera et al. 2007), doubly robust estimation (Brinkley 2014, Zhang and Liu 2012), and reinforcement learning (Alousi et al. 2017, Pineau et al. 2009) to identify the optimal treatment regime from historical data. However, one drawback of these studies is that they are offline, which means they do not punish for errors during the training phase and emphasize only maximizing post-training outcomes. In the context of medical decision-making, ignoring the cost that occurs during the training phase might be inappropriate, as testing a suboptimal decision on a trial patient may inflict irreversible harm to that patient.

Closest to our work is the second stream of studies that focus on adaptive treatment regime. These studies formulate the physicians' decision dilemma using a multi-armed bandit (MAB) framework, which explicitly models the training (exploration) and implementing (exploitation) phases as an integrated optimization problem, with the goal of maximizing the total patient survival outcome. The methodology in these studies spans from the commonly used bandit algorithm in the machine learning community, such as the family of Upper Confidence Bound (UCB) algorithms (Lai and Robbins 1985, Auer et al. 2002), TS (Chapelle and Li 2011), to bandit algorithms that tailored specifically to healthcare applications (Wang et al. 2017, Bastani et al. 2017). In particular, Wang et al. (2017) propose a knowledge-gradient (KG) policy that offers fast learning when the learning budget is limited and the cost per experiment is expensive. Bastani et al. (2017) develop an exploration-free greedy algorithm that is preferable in settings where exploration is extremely costly or even unethical, such as personalized medicine in the present study. In contemporaneous work, Tomkins et al. (2020) present a contextual bandit framework that delivers personalized behavioral treatment to users in the mobile health setting. This paper generalizes the TS algorithm,

incorporating the individual and time-specific effects in modeling patient's heterogeneous response. Although these papers provide methodological tools and theoretical prototypes for healthcare decision-making, their results are either theoretical, or from simulation studies which are not based on empirical data. Thus, scientific empirical evidence that to what extend the analytic methods can improve health in real-world scenarios is lacking. Our theoretical framework firmly grounds on the empirical context, and we validate the performance of our proposed method directly on the clinical dataset that consists of observations from real cancer patients. Thus, our study contributes to the literature by proposing a promising solution prototype for precision medicine when randomized controlled trials (RCTs) are not feasible or inefficient, but also providing more solid evidence to promote the potential usage of MAB in clinical trials.

Indeed, despite the activity in developing analytic methods and bandit algorithms for healthcare decision-making, the empirical application and evaluation of bandit models in the field of personalized medicine lags behind. Among the very few advances, Wang and Powell (2016) study the personalized treatment recommendation for knee replacement patients. They formulate the problem as a Bayesian contextual bandit that considers both patient and treatment features, and they guide the treatment assignment using KG policy which takes advantage of domain knowledge to product rapid learning. While our paper shares a similar problem formulation (e.g., Bayesian contextual bandit), our paper differs from Wang and Powell (2016) in several distinct ways. First, they did not consider patient heterogeneous response in the reward predictive model. In contrast, our developed multilevel Bayesian linear model, which, in turn, a generalization of Tomkins et al. (2020), allows the treatment efficacy to vary both at the patient level, and the LOT level, with the individual-level fixed effects being a function of their demographic features. As we will show in later sections, we observe considerable variations in both individual and LOT level coefficients. This suggests the treatment responses are heterogeneous across patients and at different disease stages; moreover, our ablation analysis suggests that the multilevel model component improves our model performance by 9.7%, endorsing the necessity of a richer, more granular-level model

that takes into account both levels of heterogeneity like ours.

Second, Wang and Powell (2016) uses KG as the treatment allocation rule, which involves computing an expected value over the predictive reward distribution. When the likelihood function is complex, and computing the posterior distribution is intractable without closed-form solutions, Markov Chain Monte Carlo (MCMC) or other approximation method is needed for performing approximate inference, which could be extremely computationally expensive. Given that we model patients' reward distribution using a multilevel Bayesian linear model, and we have considered the parameter heterogeneity across multiple dimensions, we thus use the TS as the backbone for treatment allocation. Proposed by Thompson (1933), TS balances the exploration-exploitation trade-off by randomly selecting an arm according to its posterior probability of being optimal. Recently, TS has attracted considerable attention, due to its ease in implementation, flexibility in accommodating a wide range of structured decision problems in various practical settings, yet competitive performance compared to other bandit algorithms (Chapelle and Li 2011). For example, Ferreira et al. (2018) study a price-based network revenue management problem where a retailer aims to maximize revenue from multiple products with limited inventory over a finite selling season. This study utilizes TS to balance the exploration-exploitation trade-off under the presence of inventory constraints. They empirically demonstrate the promising numerical performance of their TS-contextual algorithm. Schwartz et al. (2017) conceptualize the advertising allocation problem as a hierarchical, MAB problem using a TS allocation rule. Through extensive experiments, their results also suggest a better empirical performance of TS based algorithms compared to other popular strategies such as UCB, Gittins, epsilon-greedy, or test-rollout.

Finally, we would like to highlight one uniqueness of our study, that distinguishes us from the previous work, is that we are able to directly observe from our clinical data set the genetic profiling related to the disease for each patient, which is not commonly available. This allows us to take into account the influence of patients' genetic differences when developing the personalized treatment regimen, leading to a more precise recommendation. As for as we know, none of the aforementioned studies have incorporated cytogenetic information in their

adaptive personalized treatment designs. Our empirical examination reveals that incorporating the cytogenetic information significantly improves the model performance by 19.75%. Thus, our study closes the gap by providing personalized treatment recommendations from the lens of the precision medicine paradigm (Collins and Varmus 2015).

Chapter 3

INVESTORS' LEARNING ON CROWDFUNDED SUPPLY CHAIN FINANCE PLATFORMS

3.1 Context and Data

3.1.1 Context

Our research context is a leading online crowdfunded SCF platform in China. As of November 2020, the platform had raised more than $2.08 billion and provided funding to over 1,325 fundraisers. As we previously discussed, a fundraiser on this platform is a supply chain entity consisting of an upstream supplier, a downstream buyer, and a core firm that serves as the loan guarantor. Most suppliers and buyers on this platform are small-sized companies that are least known to the public and tend to have limited cash flow. Core firms, on the other hand, are usually large-scale enterprises with a well-established reputation and more working capital.

Figure 3.1 illustrates the flow of the supply chain financing process on the platform. When a supplier encounters a liquidity problem and requires external financing, three supply chain members reach a SCF agreement and together raise funds as a single business entity for the upstream supplier. To request for a loan, the fundraiser first needs to submit a listing that specifies the target amount to raise, the interest rate they are willing to offer, and the duration of loan. Once the listing is approved and posted, interested investors can browse the listing and then decide whether to invest and, if so, how much to invest. After a listing is fully funded and materializes into a loan, the downstream buyer is supposed to make repayments to investors as scheduled. If the buyer defaults, the guarantor may step in and take the repayments responsibilities.

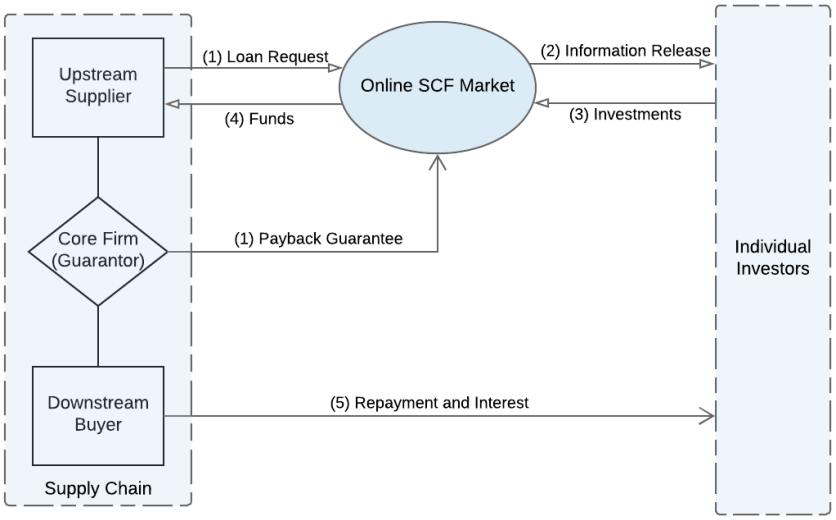


Figure 3.1: Lending Procedure in the Online SCF Market

3.1.2 Data

Our data contain all listings posted on the platform from August 14, 2017, to April 10, 2019. For each listing, we observe the posted time, fundraising target, interest rate, loan duration, and identity of the fundraiser and investors. For each investor, we know the loans in which she invests, along with the time and amount of each investment transaction.

Our empirical objective is to model the behavior of investor learning – a dynamic process in which individuals update their beliefs of guarantor reliability over time. With the original data set composed of 3,868 individual investors and 88 guarantors, a computationally challenging task arises. That is, we would need to dynamically compute a 3,868 $\times$ 88 matrix for belief updating.[1] To circumvent this computational challenge, we draw a systematic sample following the prior work that deals with a similar issue. Specifically, following Moe and

[1]We exclude individuals who have invested only once in our observation period for identification purposes to avoid purely impulse-driven investors and ensure an eloquent empirical analysis. The one-time lenders only account for 0.16% of the total funded amount and 0.13% of the total number of investments on this platform during the observation period.

Table 3.1: Results of Bootstrap Two-sample T-tests

Variable	P-value
Listing-level attributes	
Fundraising Target (in $1,000USD)	0.5050
Annualized interest rate (%)	0.4971
Loan duration (in days)	0.5934
Maximum amount allowed per investment	0.4656
Number of investors	0.4209
Investor-level attributes	
Invest amount per loan	0.7191
Total number of loans invested	0.8188

Schweidel (2012), we sample eight guarantors[2] that are involved in a total of 99 distinct supply chains (fundraisers). This sampling results in a data set of 11,587 investment decisions made by 2,411 investors and funded to 1,967 loans. To show our sample is representative of the population, we perform a series of bootstrap two-sample t-tests to evaluate whether or not the variables of our interest are systematically different between our sample and the full data set. From Table 3.1, we see that the p-values of all variables are well above 0.05, suggesting no evidence of significant difference.[3]

Table 3.2 reports the descriptive statistics of the key variables used in the empirical analysis. The fundraising target ranges from $530 to $140,000, with a median of $28,000 (in USD). The annualized interest rate across all loans ranges from 6% to 13%, with a median of 8.65%. The loan duration ranges from 17 to 360 days, with a median of 90 days.

[2]Following Moe and Schweidel (2012), we first sample the top four guarantors that guarantee the most loans to ensure sufficient investment observations. To ensure the variation across guarantors and the overall investment environment of the platform, we also include the other 4 guarantors that were chosen at random.

[3]As a robustness check, we also construct alternative samples by taking random draws of the remaining four guarantors for multiple times. We obtain consistent results with alternative samples.

Table 3.2: Descriptive Statistics of Main Variables

Variable	Mean	Std. Dev.	Min	Max	Median
Listing attributes					
Fundraising Target (in $1,000USD)	50.05	46.41	0.53	140	28
Annualized interest rate (%)	8.65	1.12	6	13	9
Loan duration (in days)	125.99	69.58	17	360	90
Maximum amount allowed per investment	5,414.55	10,654.88	140	140,000	1,400
Listing-level fundraising dynamics					
Number of investors	44.02	32.73	1	377	34
Average amount per investment	1,217.42	1,453.96	14	126,000	984
Investor-level fundraising dynamics					
Investment amount per listing	1,391.64	1,543.86	14	33,600	1,062.66
Number of listing invested	45.82	56.67	1	321	23
Number of fundraiser encountered	38.99	46.12	1	259	21
Number of guarantor encountered	32.78	36.05	1	226	19

3.2 Model Development

In this section, we develop a dynamic learning model to understand how the presence of loan guarantor shapes investment behavior in the crowdfunded SCF market. The general modeling context is that an individual investor (or individual for brevity) undergoes the following steps in the deciding-making process:

1. An individual holds a perception of guarantor reliability at a given period based on her private knowledge.

2. The individual evaluates a listing's attributes, contingent on her having browsed the listing.

3. The individual decides whether to invest in a listing or not; if she decides to invest, she also decides how much to invest.

4. Upon receiving scheduled repayments (or lack thereof), the individual obtains private signals about the guarantor's reliability and uses this information to update her perception.

It is worth highlighting two important features that motivate our model development. First, an observed individual investment outcome is constituted by two separate yet correlated decisions: 1) the incidence decision of whether to invest and 2) the amount decision of how much to invest. Since the mechanisms underlying these simultaneous decisions may not be identical, we must explicitly model them in a separate yet joint manner. Second, individuals make decisions based on their own beliefs, which are subject to evolve upon the receipts of repayment signals over time. This dynamic nature leads to state dependence among each individual's decisions, suggesting the necessity of a richer, dynamic model. In what follows, we first introduce the latent construct central to our work – i.e., individual perception of guarantor reliability – and discuss how it governs two investment decisions. We then discuss how such a perception is updated over time under the Bayesian learning framework.

3.2.1 Modeling Investment Decisions

Given that a guarantor may be involved in multiple listings, we use the doublet (k, t) to represent a listing guaranteed by guarantor k at fundraising occasion t, with t being k-specific. Let $R_{ik(t-1)}$ denote the reliability perception individual i holds toward guarantor k (Step 1). Since one can only update her perception upon the realization of the scheduled repayments (Step 4), we use index $t-1$ to indicate the most current perception before individual i having interactions with the listing (k, t). In our model, we consider that such a reliability perception will affect one's incidence and amount decision in a separate yet joint process.

Incidence Decision

We model the probability that individual i makes an investment in listing (k, t), $P(d_{ikt} = 1)$, is governed by two latent constructs:

$$P(d_{ikt} = 1) = P(z_{ikt} = 1) \cdot P(d_{ikt} = 1 \mid z_{ikt} = 1). \tag{3.1}$$

The first term, $P(z_{ikt} = 1)$, represents the possibility that an investment decision-making session exists. The inclusion of this latent session construct plays a critical modeling role in our context. From our data, we may observe that an individual does not invest in a given listing (i.e., $d_{ikt} = 0$), which may occur for two reasons. First, it could be a deliberate decision conditionacl on an individual having browsed and evaluated the listing's attributes (i.e., $d_{ikt} = 0 \mid z_{ikt} = 1$). Second, it could arise due to the fact that the decision-making session doesn't exist in the first place (i.e., $z_{ikt} = 0$). For example, an individual may not have a chance to browse the listing and thus overlook it due to some personal reasons, or the listing has been fully funded and become unavailable when she browses it. Our latent session construct takes into account the second possibility. We specify this construct as a function of investor and listing characteristics. More specifically, we have $z_{ikt} = 1$ if

$$\lambda_0 + \lambda_{1:2}\Upsilon_{kt} + \varepsilon_{z,ikt} = z^*_{ikt} + \varepsilon_{z,ikt} > 0, \tag{3.2}$$

where $\varepsilon_{z,ikt} \sim N(0,1)$. γ_0 captures an average investor's propensity to browse and consider a listing posted on the platform. γ_{kt} represents a vector of control variables that may impact the outcome of the latent session construct. We include the total fundraising target and the maximum amount per investment transaction, both of which are fixed during the fundraising window. Since these two variables together determine the number of total possible investment slots approximately, we expect them to be correlated with the outcome of the latent session construct.

The second term, $P(d_{ikt} = 1 \mid z_{ikt} = 1)$, represents the probability that individual i makes an incidence decision conditional on the existence of a decision-making session. We have $d_{ikt} = 1 \mid z_{ikt} = 1$ if

$$\beta_{i0}+\beta_{i1}R_{ik(t-1)}+\beta_{i2}R^2_{ik(t-1)}+\beta_{3:6}X_{d,ikt}+\beta_{7:10}(X_{d,ikt}\times R_{ik(t-1)})+\varepsilon_{d,ikt} = d^*_{ikt}+\varepsilon_{d,ikt} > 0, \quad (3.3)$$

where $\varepsilon_{d,ikt} \sim N(0,1)$. The individual-specific coefficient β_{i0} allows for variation in baseline investment propensity across investors. This specification helps control for individual heterogeneity in investment preference (e.g., frequency). To capture the effect of perceived guarantor reliability on the incidence decision, we include both the linear and the quadratic term of $R_{ik(t-1)}$. Coefficients β_{i1} and β_{i2} 2 together will capture the heterogeneous effect of our interest in a non-linear fashion. $X_{d,ikt}$ represents a vector of control variables. We include the listing's interest rate and loan duration. To control the potential temporal effect that arises at the platform level and beyond, we also include the time difference between the listing's posted date and the first day of our sample period (in days). Individual i's cumulative number of investments is also added to control for the individual's investment experience. Finally, we allow $X_{d,ikt}$ to interact with $R_{ik(t-1)}$ to further explore how the effect of perceived guarantor reliability is moderated by those control variables.

Given the above specification, the unconditional probability that i decides to invest in listing (k,t) can be written as:

$$P(d_{ikt} = 1) = P(z^*_{ikt} > 1) \cdot P(d^*_{ikt} > 1 \mid z^*_{ikt} > 1) = \Phi(d^*_{ikt}) \cdot \Phi(z^*_{ikt}), \quad (3.4)$$

where $\Phi(\cdot)$ denotes the cumulative density function of the normal distribution.

Amount Decision

Suppose now that individual i has decided to invest in listing (k,t). We model that the log-transformed[4] investment amount is a function of i's perception of guarantor reliability at the current time as well as other control variables:

$$\log(y_{ikt}) = \gamma_{i0} + \gamma_{i1} R_{ik(t-1)} + \gamma_{i2} R^2_{ik(t-1)} + \gamma_{3:6} X_{y,ikt} + \gamma_{7:10}(X_{y,ikt} \times R_{ik(t-1)}) + \varepsilon_{y,ikt} = y^*_{ikt} + \varepsilon_{y,ikt}, \tag{3.5}$$

where $\varepsilon_{y,ikt} \sim N(0,1)$. Similar to the incidence model, we include individual-level coefficients to account for investor heterogeneity. γ_{i0} allows for variation in baseline amount that arises due to unobserved characteristics that vary across individuals, such as disposable incomes. γ_{i1} and γ_{i2} together will measure how responsive each individual is to the perceived guarantor reliability while making an amount decision. $X_{y,ikt}$ includes the listing's fundraising target, interest rate, loan duration, and the individual's cumulative investments amount. [5] Consistent with (3.3), $X_{y,ikt}$ is allowed to interact with the reliability construct.

Given the above specifications, the conditional probability that individual i invests an amount of $\exp(a)$ in listing (k,t) is given by $P(\log y_{ikt} = a \mid d_{ikt} = 1) = \varphi(a - y^*_{ikt})$, where $\varphi(\cdot)$ denotes the standard normal probability density function.

Interdependence Between Incidence and Amount Decisions

While we have developed two separate equations governing individual incidence and amount decisions, the covariance matrix of the equation systems has not yet been clearly specified. Since the reliability construct simultaneously enters the two equations, parameter estimations could be biased if the interdependence between two decisions is not properly modeled. To address this issue, we assume the error terms in the equation systems to follow a bivariate

[4]We take log-transformation since the amount variable exhibits a hyper-dispersion property.

[5]Note that the control variables included in the two equations are non-identical. This satisfies the inclusion restriction of our equation system.

distribution:

$$\begin{bmatrix} \varepsilon_d \\ \varepsilon_y \end{bmatrix} \sim BVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_d^2 & \rho\sigma_d\sigma_y \\ \rho\sigma_d\sigma_y & \sigma_y^2 \end{bmatrix}\right). \tag{3.6}$$

Based on Equations (3.1), (3.4) and (3.5), the joint likelihood of observing individual i who makes m investments among a total of N listings during the observation period is given by:

$$L_i(y_{i\cdot}, d_{i\cdot}) = \underbrace{\prod_{(k,t)\in(d_{ikt}=0)} (1 - \Pr(d_{ikt}=1))}_{\text{N - m terms}} \underbrace{\prod_{(k,t)\in(d_{ikt}=1)} \Pr(d_{ikt}=1)\Pr(\log y_{ikt} = a | d_{ikt}=1)}_{\text{m terms}}. \tag{3.7}$$

The likelihood of observing the entire investment decision sets made by all individuals in our sample is $\prod_{i=1}^{N} L_i(y_{i\cdot}, d_{i\cdot})$.

3.2.2 Updating of Individual Perception of Guarantor Reliability

In the spirit of the Bayesian learning framework, we consider that investors are Bayesian learners who will update their perception of guarantor reliability as they experience more scheduled repayment events over time (Step 4). We assume that the initial belief of reliability individual i holds toward guarantor k follows a normal distribution: $\tilde{R}_{ik0} \sim N\left(R_0, \sigma_{R_0}^2\right)$. Parameters R_0 and $\sigma_{R_0}^2$ respectively denote the initial mean and variance of guarantor k's reliability.

There are two different types of repayment events that may trigger belief updating: 1) the realization of interest repayments, and 2) realization of the principal repayment, which of which are scheduled to be delivered on pre-announced dates. The realization of these two events could be either a successful repayment or no repayment. Given that the true reliability of a guarantor is unobservable, the realizations of repayment events, which directly signal the guarantors' commitment to protect investors, provide useful information for investors to learn about the true reliability of the associated guarantor. In line with the Bayesian learning models in the prior literature, we assume those signals to be normally distributed around the guarantor-specific true reliability. Accordingly, the m-th interest repayment signal and

the $m\prime$-th principal repayment signal, received between occasion $t-1$ and t, are given by $\tilde{F}_{iktm} \sim N(R_k, \sigma_F^2)$ and $\tilde{P}_{iktm'} \sim N(R_k, \sigma_P^2)$, respectively. All the signals received during this time period can be represented by their sample mean, which is normally distributed as well. Letting nf_{ikt} and np_{ikt} denote the total number of two different types of signals, we can express these sample means as:

$$\tilde{F}_{ikt} = \frac{\sum_m \tilde{F}_{iktm}}{nf_{ikt}} \sim N(R_k, \frac{\sigma_F^2}{nf_{ikt}}), \tag{3.8}$$

and

$$\tilde{P}_{ikt} = \frac{\sum_{m'} \tilde{P}_{iktm'}}{np_{ikt}} \sim N(R_k, \frac{\sigma_P^2}{np_{ikt}}). \tag{3.9}$$

Starting with her initial belief, the individual combines her prior belief of reliability toward guarantor k (i.e., $\widetilde{R}_{ik(t-1)}$) with all repayment signals received during the time period $(t-1, t)$, and applies Bayes' rule to formulate her posterior belief ($\widetilde{R}_{ikt}$). With the normality assumption for the initial beliefs and two repayment signals, the posterior belief at each fundraising occasion is also normally distributed, as given by $\tilde{R}_{ikt} \sim N\left(R_{ikt}, \sigma^2_{R_{ikt}}\right)$, where:

$$R_{ikt} = \frac{\sigma^2_{R_{ikt}}}{\sigma^2_{R_{ik(t-1)}}} R_{ik(t-1)} + nf_{ikt} \frac{\sigma^2_{R_{ikt}}}{\sigma_F^2} \tilde{F}_{ikt} + np_{ikt} \frac{\sigma^2_{R_{ikt}}}{\sigma_p^2} \tilde{P}_{ikt}, \tag{3.10}$$

and

$$\sigma^2_{R_{ikt}} = \frac{1}{1 \Big/ \sigma^2_{R_{ik(t-1)}} + nf_{ikt}/\sigma_F^2 + np_{ikt}/\sigma_P^2}. \tag{3.11}$$

The prior in period $t = 0$ is $R_{ik0} = R_0$, $\sigma^2_{R_{ik0}} = \sigma^2_{R_0}$.

3.3 Estimation

While the simulated maximum likelihood estimation methods are commonly used to estimate learning models in the prior literature, the approach is inconvenient in the presence of a number of individual-specific parameters. To estimate the proposed model, we use a hierarchical Bayes approach instead. In this section, we briefly explain the model hierarchy and identification. Technical details of the estimation procedure can be found in the appendices.

3.3.1 Parameter Estimation and Model Hierarchy

We apply the following estimation strategy for parameters that are subject to certain constraints. First, for the correlation coefficient ρ, we estimate the inverse arctangent transformation of it. This transformation maps the support $[-1, 1]$ to a real line. For the variance parameters $\left\{\sigma_d^2, \sigma_y^2, \sigma_F^2, \sigma_P^2\right\}$, we estimate the logarithm transformation of them since they take positive values only. As for the prior variance $\sigma_{R_0}^2$, we fix it at 1 for the identification purpose, as we will discuss in Section 3.3.2.

Let ψ denote the set of population-level parameters shared among individuals. Accordingly, we have $\psi \equiv \left[\lambda_{0:3}, \beta_{3:10}, \gamma_{3:10}, R_0, R_{1:8}, \log\sigma_d^2, \log\sigma_y^2, \log\sigma_F^2, \log\sigma_P^2\right]'$. We use vector $\theta_i \equiv [\beta_{i,3:10}, \gamma_{i,3:10}]'$ to denote the individual-specific parameters that vary across investors. We assume θ_i to follow the multivariate normal distribution:

$$\theta_i \sim MVN(\bar{\theta}, \Sigma), \tag{3.12}$$

where $\bar{\theta}$ denotes denotes the mean effects that persist across investors, and Σ represents the covariance matrix of θ_i.

As we do not directly observe the repayment signals investors receive at each time period, belief updating would be stochastic. Moreover, the mean of the belief of guarantor reliability, $\bar{R}_{ikt}$ is also a random variable. As can be seen from Equations (3.10) and (3.11), R_{ikt} is a function of repayment signals, which are unobservable to researchers. This implies that the updating process of $\bar{R}_{ikt}$ is also stochastic. Since all signals are assumed to be normally distributed, its linear combination R_{ikt} is also a stochastic variable with normal distribution. As a result, the distribution of R_{ikt} , conditional on $R_{ik(t-1)}$, can be expressed as:

$$R_{ikt}|R_{ik(t-1)} \sim N(\bar{R}_{ikt}, \nu_{ikt}^2), \tag{3.13}$$

where

$$\bar{R}_{ikt} = \frac{\sigma_{R_{ikt}}^2}{\sigma_{R_{ik(t-1)}}^2} R_{ik(t-1)} + nf_{ikt}\frac{\sigma_{R_{ikt}}^2}{\sigma_F^2} R_k + np_{ikt}\frac{\sigma_{R_{ikt}}^2}{\sigma_P^2} R_k, \tag{3.14}$$

and

$$\nu_{ikt}^2 = nf_{ikt}\frac{\sigma_{R_{ikt}}^4}{\sigma_F^2} + np_{ikt}\frac{\sigma_{R_{ikt}}^4}{\sigma_P^2}. \tag{3.15}$$

Therefore, the unobserved reliability belief can be drawn from the following natural hierarchy:

$$R_{ikt} \mid R_{ik(t-1)} \sim N\left(\bar{R}_{ikt}, \nu^2_{ikt}\right),$$
$$R_{ik(t-1)} | R_{ik(t-2)} \sim N\left(\bar{R}_{ik(t-1)}, \nu^2_{ik(t-1)}\right),$$
$$\dots$$
$$R_{ik1} | R_{ik0} \sim N\left(\bar{R}_{ik1}, \nu_{ik1}\right).$$

The full hierarchy of the model can be summarized as:

$$Y_{ikt}, D_{ikt} \mid R_{ikt}, \psi, \theta_i, \sigma^2_z, \sigma^2_d, \sigma^2_y;$$
$$R_{ikt} \mid R_{ik(t-1)}, R_k, \sigma^2_F, \sigma^2_P, nf_{ikt}, np_{ikt};$$
$$\theta_i | \bar{\theta}, \Sigma.$$

We adopt a Markov Chain Monte Carlo (MCMC) procedure to estimate the model parameters. Specifically, we combine Gibbs sampler and Metropolis algorithm to recursively draw from the conditional distributions of model parameters based on the above hierarchy structure. Our MCMC sampling procedure consists of draws from four different Markov chains, each of which has 10,000 draws. The first half of the draws are discarded as the burn-in samples, and we use the remaining 5,000 draws in each chain to construct the posterior samples, resulting in a total of 20,000 draws for inferences. The result of our estimation is a set of posterior distributions, approximated by the sample draws of each unknown parameter.

3.3.2 Identification

In this section, we briefly discuss some intuition as to how the model parameters are identified. In our model, individuals make investment decisions based on their perceived utility. The variance of the two signals, σ^2_F and σ^2_D, are identified from the dynamics of individuals' investment behaviors over time. Every time when an individual receives a signal from a repayment event, she would acquire some new information about the guarantor's reliability and use it to update her belief, which, in turn, will impact her subsequent investment decisions.

Table 3.3: Model Fit and Comparison

Model	Full Model	Model 1	Model 2	Model 3
Heterogeneity	Yes	No	No	No
Latent decision session construct	Yes	Yes	No	No
Interaction Term	Yes	Yes	Yes	No
Log-Likelihood	-9,456.15	-9,687.45	-9,987.36	-11,369.78
WAIC	11,723.80	12,456.36	12,968.41	14,539.14
ΔWAIC w.r.t. full model		732.56	1,244.61	2,815.34

The mean of the prior belief of reliability R_0 is identified from the direction of change in investment behavior as individuals receive more repayment signals. For example, if an individual's incidence probability increases after she receives a repayment signal, then it implies that the mean of the individual's prior belief is lower than that particular signal. Thus, the direction of the change and the extent of the change from the initial investment behaviors helps identify a given individual's prior belief.

The true reliability level, R_k, is identified from the likelihood of a listing, associated with guarantor k, being invested and the amount funded. As an individual receives more signals, her belief R_{ikt} will evolve toward the true reliability R_k. After we fix the variance of individuals' initial belief (i.e.,$\sigma^2_{R_0} = 1$), the estimates of σ^2_F and σ^2_D would represent the relative magnitude of the signal variance compared to the variance of the individual's initial belief.

3.3.3 Model Fit and Comparison

Our model allows investors to have individual-specific parameters, and we also include linear interaction terms to capture a more flexible utility function form. Before we discuss the estimation results, we first demonstrate our proposed model's fitness by comparing our full model with several nested ones in Table 3.3.

To draw the comparison, we use the widely applicable information criterion (WAIC),

which is an extension of the Akaike Information Criterion (AIC) that is more fully Bayesian than the Deviance Information Criterion (DIC). As reported in Table 3.3, we obtain the best fit from our full model. The model performs worse if we do not control for heterogeneity among investors (Model 1) and the latent decision-making session (Model 2). Moreover, the model fit significantly drops if we exclude the interaction term in the model (Model 3), suggesting the necessity of considering the moderating effect of the perceived reliability when formulating the investment utility function.

3.4 Results

Table 3.4 reports the posterior means of parameters from the full mode. Since a Bayesian approach is used for estimation, we assess the significance levels of parameter estimates based on the highest posterior density (HPD). An HPD interval describes the information we have about the location of the true parameter upon the data being observed. Parameter estimates are considered significant if a high percentage of the HPD interval does not contain zero.

The estimates from our proposed models provide supportive evidence that individual investment behavior is, at least partially, driven by their perception of guarantor reliability. Individuals who perceive a guarantor to be more reliable not only are more likely to invest ($\beta_{i1} > 0$) but also tend to invest a higher amount ($\gamma_{i1} > 0$) in the associated listings. However, the estimated coefficients of the quadratic term reveal that this latent construct has dissimilar effects across two decision-making stages. The estimated coefficient from the incidence model (β_{i2}) is negative and significant. This means that the marginal positive effect of reliability perception diminishes as the investors perceive guarantors as more reliable. This non-linear pattern, however, is not found at the amount decision stage (γ_{i2}).

Next, we turn to discuss the estimates of other parameters and dig into the moderating role of the reliability perception on investor responses to other attributes. First, we see that the estimated coefficient of the interest rate is positive and significant from both models (β_3, γ_4), meaning that investors are more likely to invest in and invest a higher amount in the listings that offer higher returns, with everything else being held equal. The estimated

Table 3.4: Main Estimation Results[†]

	Parameter	Description	Mean
	β_{i0}	Mean effect of baseline investment propensity	-0.0353
	β_{i1}	Perceived reliability - linear (R_{ikt})	0.6803^{***}
	β_{i2}	Perceived reliability - quadratic (R_{ikt}^2)	-0.4614^{***}
	β_3	Interest rate	0.6678^{**}
Investment	β_4	Loan duration	-0.1337
Incidence	β_5	Loan relative start date	-0.4404
Model	β_6	Individual cumulative number of investments	-0.1204
	β_7	$R_{ikt}\times$ interest rate	-0.1725^{***}
	β_8	$R_{ikt}\times$ loan duration	-0.1337
	β_9	$R_{ikt}\times$ loan relative start date	-0.5451
	β_{10}	$R_{ikt}\times$ individual cumulative number of investments	0.0552
	γ_{i0}	Mean effect of baseline investment amount	1.0896^{***}
	γ_{i1}	Perceived reliability - linear (R_{ikt})	0.3595^{***}
	γ_{i2}	Perceived reliability - quadratic (R_{ikt}^2)	0.6294
	γ_3	Fundraising target	0.7495^{***}
Investment	γ_4	Interest rate	1.3284^{***}
Amount	γ_5	Loan duration	-0.2459
Model	γ_6	Individual cumulative investment amount	0.2444^{***}
	γ_7	$R_{ikt}\times$ fundraising target	0.3020
	γ_8	$R_{ikt}\times$ interest rate	-1.6973^{*}
	γ_9	$R_{ikt}\times$ loan duration	0.0578^{***}
	γ_{10}	$R_{ikt}\times$ individual cumulative investment amount	0.5018
Interdependence	ρ	Interdependency between two decisions	0.6414^{***}

[†]Notes: ***99%, **95%, *90% of the HPD interval does not contain 0.

coefficient of the fundraising target from the amount model suggests that a higher requested amount is likely to induce a larger size of contributions per transaction ($\gamma_3 > 0$). Moreover, the estimated coefficient of the individual cumulative investment amount is significant ($\gamma_6 > 0$), indicating that investors are interested in contributing a higher amount as they have invested more on this platform increases. An implication of these results is that the platform could take into account individual heterogeneity and design different marketing strategies across different segments of investors to incentivize them to contribute more effectively.

The estimated coefficients of the interaction terms, reported in Table 3.4, reveal how taste coefficients (i.e., coefficients of listing attributes) are moderated by individual perception of guarantor reliability. Specifically, the estimated coefficients of the interaction between reliability perception and interest rate are negative and significant from both models. This indicates that when making investment decisions, investors are more sensitive to a listing's interest rate when the extent of perceived guarantor reliability is low. In addition, we observe that the estimated coefficient of the interaction term between reliability perception and loan duration is positive and significant from the amount model. This implies that while a higher extent of perceived reliability motivates investors to make a larger investment, a longer loan duration would accentuate such a motivation. Based on this finding, supply chain fundraisers can find opportunities to improve the design of their listings to attract more funding. Finally, the estimated coefficient of ρ is positive and significant, confirming an intuition that investors who have a higher propensity to invest in a given listing tend to fund a larger amount to that listing.

Table 3.5 reports the estimation results for other model parameters, including coefficients governing latent decision-making sessions and investor learning. The estimates of γ_1 is positive whereas those of γ_2 is negative. Taken together, these two results confirm our earlier intuition that a higher fundraising target and a lower amount cap can be approximately translated into more investment slots, which are able to draw more eyeballs from the crowd.

The estimates of the learning coefficients confirm the existence of investor learning in the online crowdfunded SCF market. First of all, the estimated true reliability varies across

Table 3.5: Estimation Results of Other Parameters[†]

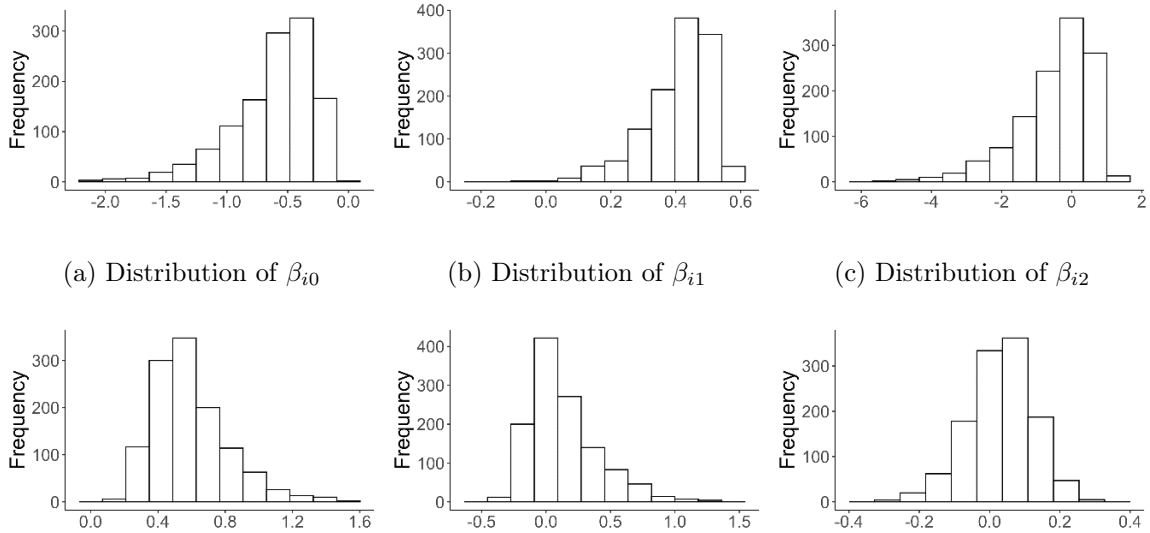
	Parameter	Description	Mean
Latent Decision-making Session Parameters	λ_0	Mean effect of baseline propensity	0.6017^{***}
	λ_1	Fundraising target	0.1912^{**}
	λ_2	Maximum amount per investment	-0.5200^{*}
Learning Parameters	R_1	True reliability - guarantor 1	0.1674^{**}
	R_2	True reliability - guarantor 2	-0.5819^{***}
	R_3	True reliability - guarantor 3	-0.4401
	R_4	True reliability - guarantor 4	0.6075
	R_5	True reliability - guarantor 5	0.5966^{***}
	R_6	True reliability - guarantor 6	0.6933^{***}
	R_7	True reliability - guarantor 7	-0.1424^{***}
	R_8	True reliability - guarantor 8	-0.3019
	R_0	Initial prior on guarantors' reliability	-0.1866
	σ_F^2	Variance of interest repayment signals	0.8113^{***}
	σ_P^2	Variance of principal repayment signals	0.2455^{***}
	σ_d^2	Variance of errors in the incidence model	0.8308^{***}
	σ_y^2	Variance of errors in the amount model	0.6106^{***}

[†]Notes: ***99%, **95%, *90% of the HPD interval does not contain 0.

guarantors, implying that guarantors are not perceived as equally reliable. The estimates of the two repayment signal variances, σ_F^2 and σ_P^2, are 0.8113 and 0.2455, respectively. Recall that the signal variance represents its relative magnitude compared to the variance of prior belief $\sigma_{R_0}^2$ (fixed at 1). These estimates suggest that the two signals provide more precise information about guarantor reliability, relative to investors' own initial beliefs. In other words, receiving repayment signals indeed help investors learn about the true reliability of the associated guarantors. A quick comparison between the two estimated variances reveals that the principal repayment signals provide more precise information about the true reliability, allowing investors to learn at a higher rate. This result is reasonable because, in our context, principal repayments are associated with a much larger amount than interest counterparts. As a result, a successful principal repayment carries more informational value from the investors' perspective.

We plot the distribution of the individual-level parameters in Figure 3.2. As we can see, there is considerable heterogeneity across investors. We find from Figure 3.2(a) that there is significant variation in individuals baseline propensity to invest (β_{i0}). This indicates that investors in the crowdfunded SCF market are quite different in their investment preference. From Figure 3.2(b) and Figure 3.2(c), we also observe variations in their response to guarantor reliability they perceive, but most of the individuals share the same sign of the two taste coefficients ($\beta_{i1} < 0$ and $\beta_{i2} > 0$). This suggests that individuals' taste for the perceived reliability are relatively homogenous, although some are more susceptible to the latent perception than the others. Regarding the individual-level deviations for the coefficients of the amount model, the distribution of the intercept coefficient (γ_{i0}, as illustrated in Figure 3.2(d)) signifies investors' heterogeneous baseline investment amount. Moreover, as Figure 3.2(f) shows, while most individuals invest a higher amount with increased perceived reliability ($\gamma_{i1} > 0$), the sign of the estimated coefficient of the quadratic term is quite inconsistent among individuals. This result suggests that the perception influences investors' decision outcome in distinct non-monotonic ways across individuals.

To further examine the interdependence between the investment incidence and amount



(a) Distribution of β_{i0} (b) Distribution of β_{i1} (c) Distribution of β_{i2}

(d) Distribution of γ_{i0} (e) Distribution of γ_{i1} (f) Distribution of γ_{i2}

Figure 3.2: Distribution of Individual-specific Parameters

decisions, we present the correlation matrix of these individual parameters in Table 3.6. The negative correlation between β_{i0} and β_{i1} suggests that investors with low baseline propensity to invest are more subject to the influence of their perceived reliability values; similarly, the significant and negative correlation between γ_{i0} and γ_{i1}, indicating the investors who are more conservative in their total contribution would be influenced by their perceived reliability to a larger extend.

3.5 Managerial Implications

Our empirical result yields some interesting findings and offers important managerial insights for both supply chains and platform managers who attempt to facilitate fundraising through this novel crowded SCF finance market. Our results suggest that repayment signals are informative in revealing the guarantors' true reliability. Being exposed to this external informational signal, their learning process is accelerated, and more sound decisions can

Table 3.6: Correlation Matrix†

	Incidence Model			Amount Model		
	β_{i0}	β_{i1}	β_{i2}	γ_{i0}	γ_{i1}	γ_{i2}
β_{i0}	1					
β_{i1}	-0.9300*	1				
β_{i2}	0.9272	-0.8498*	1			
γ_{i0}	0.9581	-0.8612*	0.8738	1		
γ_{i1}	-0.9368	0.8691*	-0.8311*	-0.8997*	1	
γ_{i2}	-0.9224	0.8058	-0.8307	-0.9540	0.9045	1

†Notes: $*$95% of the HPD interval does not contain 0.

thus be made. A practical implication for supply chain manufactures, is that they could optimize the structure of loan contracts, breaking down their fixed payments to more frequent repayment events to accelerate crowd investors' correlated learning, which in turn help build the credit profile among individual investors.

Our empirical estimation results also suggest that the existence of guarantors and shared information across individual loans influence crowd funders' learning on loan riskiness, and thus their investment behaviors. Several actionable strategies are immediate from the findings from our estimation results. For example, for lending markets, the platform may consider building branding systems to help build customers' correlated learning, thus promote individuals' investment decision outcomes. Further, we identified that high-interest rate and short loan duration positively correlated with their investment utility, and this utility is moderated by their perceived reliability level. An effective recommendation system can be adapted to better segment the market by targeting individual investors with loans that better match their favorable preferences.

To further illustrate how the platforms could increase their fundraising performance and

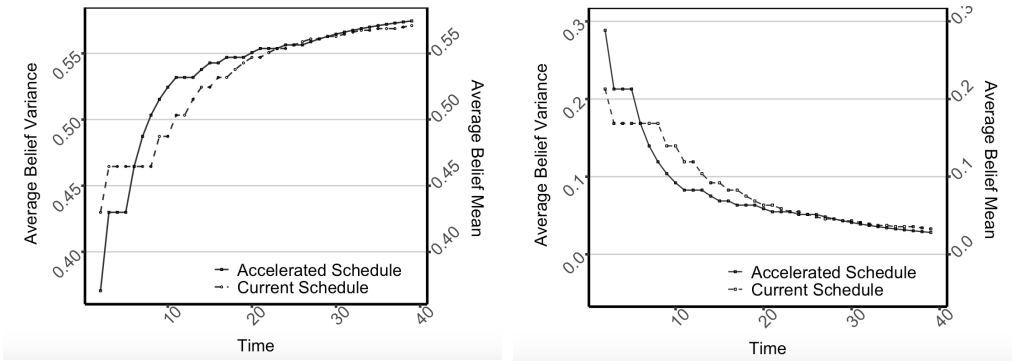
Table 3.7: Simulation Results

			(1) $R_2 = -0.5819$		(2) $R_5 = 0.5966$	
	Scenario	Change	$\bar{d}$	$\bar{y}$	$\bar{d}$	$\bar{y}$
Current	1	With belief updating, actual repayment schedule	0.0125	6546	0.0272	10293
Alter Belief Updating	2	No belief updating, prior belief	0.0128	7896	0.0222	8187
	3	No Belief Updating, True mean reliability	0.0101	5118	0.0314	13569
Alter Payment Schedule	4	With belief updating, bi-weekly repayment schedule	0.0105	5876	0.0289	12985

the overall investor contribution if we have shared information across loans and optimize the repayment schedule, we conduct three counterfactual simulations to see the impact numerically. We chose two guarantors, R_2 and R_5, where R_2 has its true estimated mean reliability lower than the individuals' prior ($R_2 = 0.5819$) and R_5's true mean reliability is higher than the individuals' prior belief ($R_5 = 0.5966$). We randomly sample 100 individuals and their heterogeneous taste parameters from the estimates model. We simulate the investment outcomes separately for loans associates with each of the guarantors in each scenario. We then compare the simulated investment outcomes averaged over 5,000 iterations in three different counterfactual scenarios, with those from our current model settings.

Simulation 1: Altering the belief updating

In this simulation, we compare our current setting with two different scenarios to examine how the correlated learning influences the fundraising outcome. In the first scenario, we consider the investors' behavior without belief updating, and they do not have private information. Thus, the crowd investors hold the same prior belief for all guarantors, and the beliefs are constant over time without learning from the past. This is similar to the conventional lending environment where individuals' investment decisions are independent across loan occasions without correlated learning. In the second scenario, individuals have constant belief on guarantors with the full information; i.e., their beliefs kept being the true mean reliability of the guarantor over time. In this case, the perceived reliability construct



(a) Mean Evolution (b) Variance Evolution

Figure 3.3: Evolution of Belief Comparison

R_{ikt} reduces to the guarantor fixed effect $\bar{R}_i$ without dynamic evolving.

By comparing the investment outcome between our current case and scenario 1 for R_2, we find that by enabling the customers' correlated learning, the investors have lower investment probability and investment amount on loans whose guarantors' true reliability is lower than the prior mean. This result indicates that the shared information across loans allows investors to learn the unobserved attributes of loans from correlated experience, mitigating the degree of adverse selection and improving overall market efficiency. For guarantors whose true reliability was under evaluated (e.g., R_5), the crowdfunded SCF market provided an efficient way to help the guarantors quickly build their credit profile among investors and thus achieve better investment outcomes compared to that through traditional SCF channels. As we can see from the investment outcome for scenario 1 for R_5, the crowdfunded SCF market can help alleviate the price discrimination for those loan raisers with a relatively short history and will endure high costs when raising funds through the traditional SCF channel.

Simulation 2: Altering interest repayment schedule

In this simulation, we consider individuals as Bayesian learners who will update their beliefs over time when they receive new signals. However, we alter the interest repayment schedule from monthly to bi-weekly to see how a more frequent repayment signaling would influence the overall fundraising outcome. From the simulations results presented in Table 5.10, we find that by varying the interest repayment schedule from monthly to bi-weekly, the guarantor with under-evaluated credibility (i.e., R_5) is able to increase the investment propensity by 6.25%, and investment amount by 26.15% compared to the current scenario.

We further plot out individuals' belief evolution Figure 3.3 to visualize the trend of investors' belief evolving under different repayment schedules. As we can see from the graph, investors can update their beliefs at a much faster rate when receiving the extra information regarding guarantors' credibility through the bi-weekly payment schedule.

3.6 Conclusion

SCF is a new financing method aiming to liquidize financial flows within and between companies through cross-company integration within a supply chain. With the rise of online crowdfunding platforms, a novel concept called crowdfunded SCF has gained popularity among supply chain members and individual investors. Comparing to the traditional SCF market, the crowdfunded SCF market simplifies the capital-seeking process by directly connecting supply chain raisers with potential crowd investors without the professional intermediaries. However, for individual investors, they may lack sufficient information to evaluate the riskiness of a loan; as a result, they usually seek and rely on multiple sources of information and make decisions based on their perceptions of the unobserved attributes of the loan. As an increasing number of supply chains attempt to adopt this alternative financing, being able to understand how individual investors' perceptions are formulated, updated and how it operates at individual's different decision stages to influence the investment outcomes is critical for a better fundraising strategy that can best enhance investors valuation and invest

motivation.

In this paper, we explore investors' behavioral mechanisms in this novel crowdfunded SCF market. Whereas past studies have examined individuals' investment behavior in P2P lending platforms that are more independent on different occasions, this paper, what is to our knowledge, presents the first attempt to explore a context where there exists shared information across individual loans, leading to state dependence among their behavior dynamics over time. By modeling individuals as Bayesian learners, we explicitly conceptualize the additional utility associated with the shared attributes across loans. We distinguish among different sources of information individuals may rely on when formulating the perceptions, and we take an in-depth look at the interactions between them. Such an approach allows us to examine investors' learning process under uncertainty and explore the heterogeneous response to the perceptions.

Using a rich panel dataset that consists of individual-level investment decision dynamics and static loan attributes, we empirically validate the existence of individuals' learning, and we find supportive evidence that individuals' decisions of making investments are driven by their subjective perceptions in a non-monotonic manner. We also observe the moderating role of the perceived reliability: crowd investors are more sensitive to the loan interest rate. They have low perceived reliability, and a longer loan duration amplifies the extra incentive from the increased perceived reliability associated with a loan guarantor. A more granular-level examination on the individual-level coefficients reveals a heterogeneous non-linear influence of the perceived reliability on investment decisions. More conservative investors are more sensitive to the guarantors' perceived reliability level.

Our results provide important managerial implications for both supply chain fundraisers and lending platform managers. Our simulation suggests that the individual learning induced by shared information allows investors to learn the unobserved attributes of loan from correlated experience, mitigating the degree of adverse selection and thus helps improve the overall market efficiency; moreover, the supply chain fundraisers can consider better structure the loan contracts by having more frequent repayments events to accelerate crowd investors'

correlated learning, which in turn help build the credit profile among individual investors.

This study, however, has some limitations and can be extended in the following three directions. First, when modeling individuals' investment incidence and amount decisions, we only include past total investment history to account for the variation of taste coefficients due to data limitation. Demographic information, such as gender, age, and income, is shown by many previous studies correlated with an individual's investment pattern. Second, due to data availability, we modeled latent choice set probability as a function of loan observed attributes. Future work can consider obtaining click-stream data from this website to enrich the econometric model. To some extent, limiting our analysis within a platform with one financial product may reduce the applicability of our results to other financial products and different types of crowdfunding platforms. But we believe our results are still insightful for managers in crowdfunding platforms since the basic principles to maximize utility under uncertainty, which drive individuals' decisions are similar.

Chapter 4

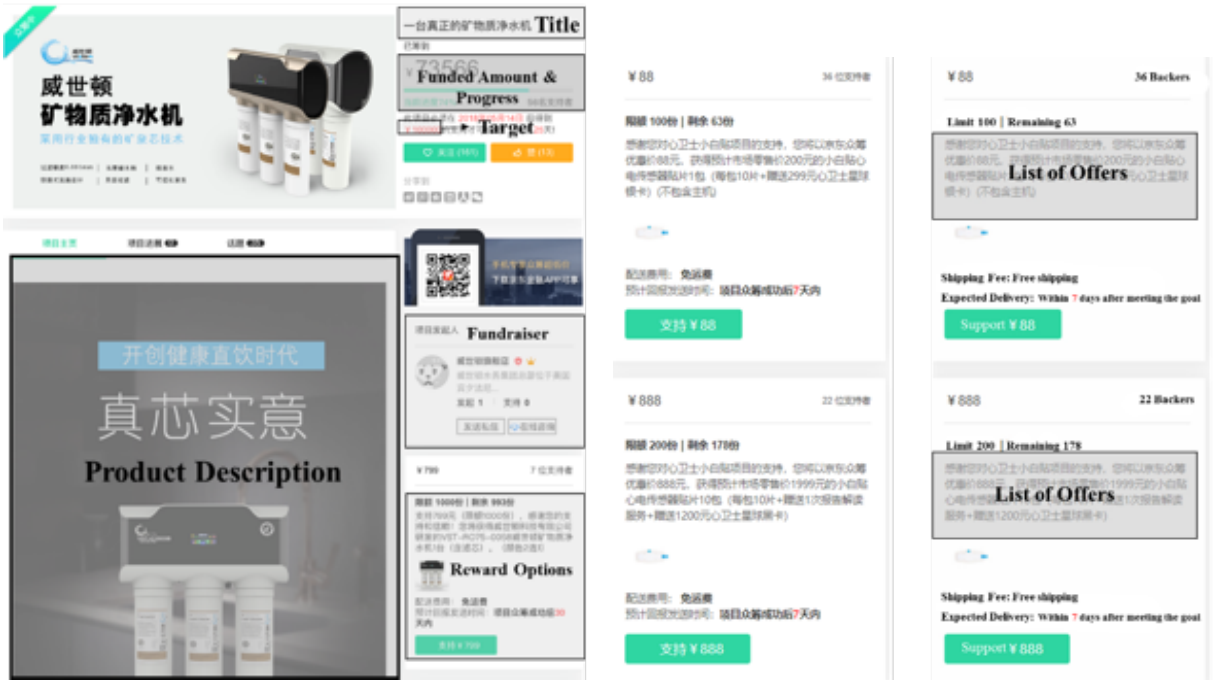
EFFECTS OF SCARCITY-INDUCED DEMAND ON THE CROWDFUNDING MARKET

4.1 Data and Descriptive Analyses

4.1.1 Background and Data

The data used for this study were collected from the JD crowdfunding platform. Opened to the public in 2014, JD crowdfunding has grown rapidly into one of the largest reward-based crowdfunding platforms in China. In 2017, it facilitated the success of over 4,000 fundraising campaigns and collected more than $250 million from crowd backers.

The fundraising process on the JD crowdfunding platform proceeds as follows. When a creator has an innovative business idea but lacks sufficient initial capital, the creator can seek financial support by initializing a fundraising campaign on the platform. Fundraisers must include narrative statements and pictures to describe their business ideas and must specify, in advance, the fundraising goal, duration of the campaign, and the different levels of reward options for backers in exchange for their monetary contributions. These reward options can be highly diversified; they could be as simple as a finished product or a combination package that includes multiples of the product and certain exclusive. Different reward options are usually set at different price levels to cater to the needs and preferences of various backers. In addition, fundraisers can choose to limit the availability of any reward to restrict the number of backers who can take advantage of the offering. Once the upper limit has been reached, that reward becomes unavailable to upcoming backers. During the active fundraising cycle, backers can contribute to a campaign by purchasing its available reward options. If that campaign has reached the pre-set fundraising target, backers will get the offering specified in the reward option by the promised delivery date; if it is not fully funded at the end of the



(a) A Sample Active Campaign

(b) Sample Reward Options

Figure 4.1: Example Campaign and Reward

fundraising duration, all of the raised funds will be returned to the backers.

Figures 1 and 2 show a sample campaign and a sample reward, respectively. As seen, detailed information for each reward, including price, shipping fee, expected delivery date, and the list of offers, is displayed in an eye-catching manner. In addition, the total availability of the reward, as well as its current demand, is visible to any potential backers, which makes the scarcity of each reward option a salient feature when individuals are making decisions. It is worth noting that, for most mainstream reward-based crowdfunding sites, such as Kickstarter, the fundraising mechanism, feasibility of imposing quantity limits on reward options, and the presentation of the information are very similar to those of JD crowdfunding, as we described above. Thus, our data are representative of the backing behavior of investors, typically observed on these other sites as well.

Our dataset contains campaign characteristics and fundraising dynamics for all of the 617 active campaigns and their 2,938 different reward options from May 1, 2017, to June 30, 2017. In our two-month observation period, the fundraising statistics are collected at the reward level on a daily basis. Thus, the individual unit in our panel data is a reward, and the temporal unit is a day. For each reward, we take a snapshot of the number of backers and the amount received at the end of each day. Campaign-level dynamics are obtained by aggregating the fundraising statistics across all associated reward levels. For example, if a fundraiser has specified five different reward levels under the campaign, then the campaign-level dynamics (i.e., the total number of backers, the total funded amount) within a day can be obtained by summing up the statistics across these five reward levels. For all 617 campaigns, we also collected static characteristics, such as fundraising goal, fundraising duration, and the total number of included reward levels. For the 2,938 rewards, we collected price information as well as total availability if the reward was limited.

It is worth noting that our focus on reward-level dynamics distinguishes our work from other studies that commonly examine the fundraising performance aggregately at the campaign level (Yang et al. 2017). This sharper focus allows us to control for a more granular level of heterogeneity. In addition, we are able to model the interactions and mutual influences among different reward levels within a campaign, which is crucial to the understanding of an effective reward scheme design.

4.1.2 Descriptive Analyses

Before we introduce our empirical model, we first summarize the variables in our data and then present some basic patterns that we found in our data set that motivate our subsequent model development and support the need to address several econometric concerns. Table 4.1 presents the summary statistics for the 617 campaigns and their 2,938 different reward levels. In this sample, fundraisers request between 5,000 CNY and 2,000,000 CNY[1], with an

[1] 1 CNY = 0.15 US dollars, per the exchange rate on August 27, 2018.

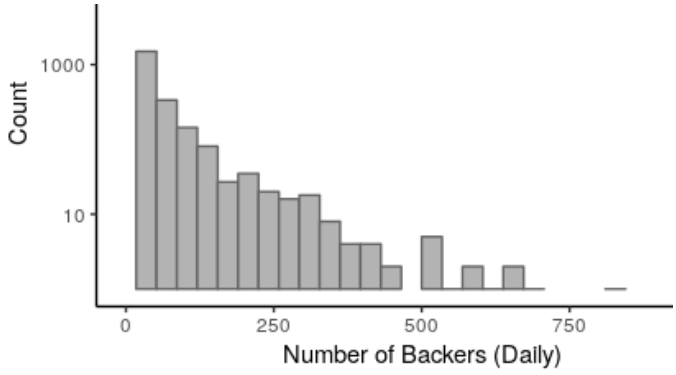


Figure 4.2: Distribution of Daily Backers' Arrival (Log Scale)

average of 99,781.2 CNY. The duration of each campaign is 15 to 60 days, with a mean of 34.30 days. Among all of the observed campaigns, 47% are successfully funded during the fundraising cycle.

To analyze backers' investment behavior patterns, we plot the distribution of the number of backers for each reward within an active fundraising day, as seen in Figure 3. This figure suggests that a large portion of the records is associated with zero backers, which indicates a pattern of zero inflation in our outcome variable. Moreover, through a closer examination of the relevant statistics in Table 4.1, we can see that the number of incoming backers each day appears to be significantly over-dispersed (mean = 10.02, standard deviation = 215.11). These patterns provide evidence to accommodate excess zero and over-dispersion into our model to best fit our data.

To have a better sense of the hierarchical structure of our data, we plot the reward-level fundraising dynamics over time for some sample campaigns in our dataset. Figure 4.3 shows the daily raised amount across different reward levels within Campaigns 127 and 624. As seen, the dynamics for reward levels within the same campaign show similar patterns but differ across campaigns.

To further control for the possibility that these similarities are driven purely by shared characteristics across rewards within the same campaign, we regress the daily raised amount

Table 4.1: Summary Statistics

Variable	Mean	Std Dev	Min	Max	Median
Project Attributes					
Target	99,781.20	117,625.20	5,000	2,000,000	1,00,000
Average Price	5,576.67	15,668.55	5	219,120	1,326
Duration (Days)	34.3	9.58	15	60	30
Success (Yes = 1)	0.47	0.5	0	1	
HasLimitedRewards (Yes = 1)	0.99	0.11	0	1	1
Number of rewards	4.84	2.12	1	26	5
Number of Observations	617				
Reward Attributes					
Price	5,526.10	26,669.27	2	688,800	499
Expected Delivery	22.93	12.17	1	154	30
IsLimited (Yes = 1)	0.96	0.21	0	1	1
Total Availability	515.92	1,314.00	1	50,000	300
Number of Observations	2,938				
Project-Level Funding Dynamics					
Average Daily Backers	51.6	492.39	0	12,055	11.31
Average Daily Raised Funds	14,215.21	68,556.75	0	1,375,385	3852.16
Reward-Level Funding Dynamics					
Number of Backers Each Day	10.02	215.11	0	11,592	0.57
Funded Amount Each Day	2,847.38	30,580.36	0	1,375,000	249.5

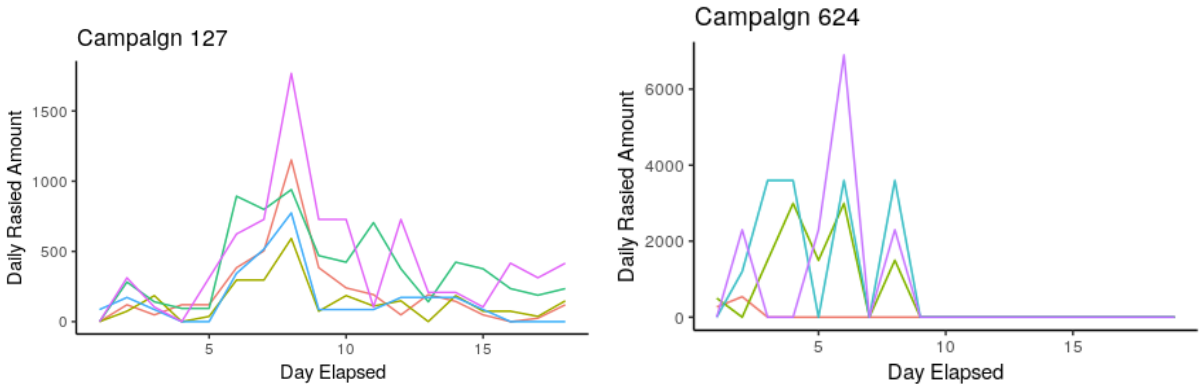


Figure 4.3: The Dynamics of Total Raised Amount for Rewards Within Campaigns

on observed reward characteristics and plot the residuals, as seen in Figure 5. As seen in the figure, after controlling for the observed characteristics, the observed systematic pattern persists. These plots provide suggestive evidence of the existence of structural unobserved heterogeneities across rewards. These heterogeneities could be driven by several factors, such as individuals' hierarchical preference for rewards (i.e., individuals' preference for reward levels are structured by the campaigns to which the rewards belong) or correlated unobservable (i.e., some campaign-level unobserved factors). Either way, it would influence the empirical estimation results if unaddressed, which necessitates the use of a hierarchical model.

Finally, the results presented in Table 4.1 indicate that 96% of the reward options are limited to a certain number of backers and that 99% of fundraisers have reward levels with a quantity limit in their campaign. The distribution of the total availability for each reward, however, shows a high dispersion, with a mean of 507.98 and a standard deviation of 1,294.06. Figure 4.5 presents this pattern in log-scale. Despite the fact that fundraisers on this platform show an eagerness to impose quantity limitations to attract backers, their strategies are heterogeneous and lack patterns. Whether this strategy is indeed beneficial for the fundraising outcome and how to effectively design the reward scheme to best harness the persuasive power of scarcity are still unknown. We thus attempt to address these uses

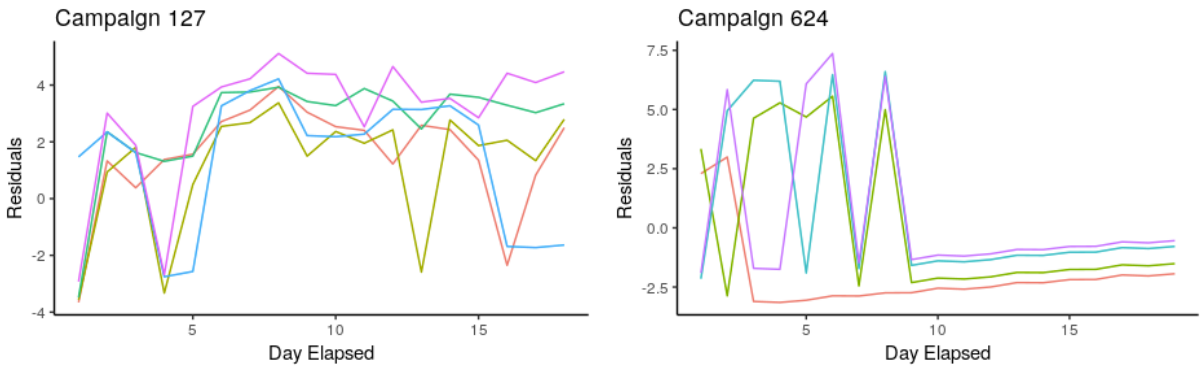


Figure 4.4: Residuals from OLS Regression

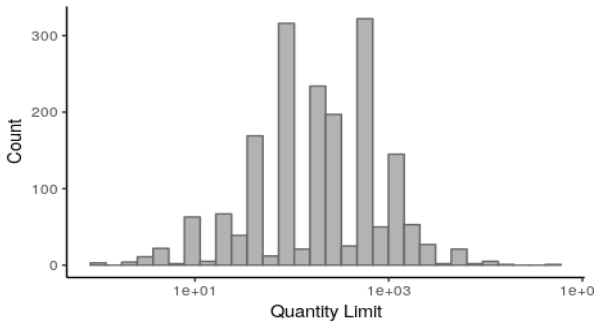


Figure 4.5: Distribution of Quantity Limit in Log-Scale

in the following sections.

4.2 Empirical Approach

4.2.1 Baseline Model and Variable Description

We begin the empirical analysis by using a baseline model to provide the preliminary evidence of the scarcity effect on crowdfunding markets. We use the number of backers that a reward has attracted within a day to measure backers' investment intention. This measurement is equivalent to the daily funded amount once we have controlled for the price of the reward.

Using the counts (i.e., the number of backers) as our dependent variable, however, allows us to better capture the occurrence of receiving zero contribution in the fundraising dynamics from the perspective of modeling. In addition, the number of attracted backers each day serves as a better indication of a campaign's popularity, which provides us stronger insight into how the backers' arrival rate changes in response to the factors of interest.

We use i to index campaign and j to index reward option within that campaign. Thus, the double (i, j) represents campaign i's j-th reward option. Because the fundraising dynamics are summarized on a daily basis, the time frame t in this study is the number of days elapsed since its launch, where . We did not exclude the right-censored campaigns (i.e., the campaigns that are still in active fundraising duration at the end of our observation period) because this would lead to potential bias due to truncation or selection (Heckman 1976). Thus, if all of the fundraising dynamics are fully observed, will be equal to the duration of that campaign; if a campaign is ongoing at the end of the observation, then will be equal to the days elapsed since its launch at the last observation date.

Our outcome variable of interest is the number of backers' reward (i, j) that has been received at its tth fundraising day, which is denoted as y_{ijt}. As noted in Section 3.2, this outcome variable has two salient features: the relatively large share of zeros (i.e., no incoming backers within a day) and the presence of over-dispersion in the number of daily arrivals. To explicitly take these two features into account, we employ a ZINB. The ZINB is a mixed model that contains two distinct processes: a negative binomial model to account for the over-dispersion feature of the count outcomes and a logit model for predicting the excess zeros. More formally, y_{ijt}. is modeled as:

where the θ is the zero-inflation rate and is the negative binomial distribution given by

$$p(y_{ijt}|\pi, \mu_{ijt}, \varphi) = \begin{cases} \pi + (1-\pi)g(y_{ijt} = 0) \text{if} y_{ijt} = 0 \\ (1-\pi)g(y_{ijt}) \text{if} y_{ijt} > 0 \end{cases} \tag{4.1}$$

where the θ is the zero-inflation rate and $g(y_{ijt})$ is the negative binomial distribution

given by

$$g(y_i) = \Pr(Y = y_{ijt}|\mu_{ijt}, \varphi) = \frac{\Gamma(y_{ijt} + \varphi^{-1})}{\Gamma(\varphi^{-1})\Gamma(y_{ijt} + 1)} \left(\frac{1}{1 + \varphi\mu_{ijt}}\right)^{\varphi^{-1}} \left(\frac{\varphi\mu_{ijt}}{1 + \varphi\mu_{ijt}}\right)^{y_{ijt}}. \quad (4.2)$$

In the above ZINB model, each part of the mixture (i.e., the negative binomial regression model and the logit model) accommodates zeros, and the zero occurrences in our data could be classified into two types: "chance zeros," those that occur due to the stochasticity in the negative binomial model characterized by the heterogeneous mean arrival rate μ_{ijt} and the over-dispersion parameter ϕ; or "structural zeros," those that occur due to some systematic factors across campaigns, such as individuals' degree of attention, which is captured by a zero-inflation rate π.

We further model the log transformation of μ_{ijt}, the expected backers' arrival rate for reward (i, j) at its t-th fundraising day, as a linear function of the explanatory variables:

$$\log(\mu_{ijt}) = \theta_0 + \theta_1 S_{ijt} + \theta_2 Q_{ij} + \theta_3 Y_{ij,t-1} + X_j \cdot \Phi + \Upsilon_{ij} \cdot \alpha + H_{ijt} \cdot \eta + e_{ijt}. \quad (4.3)$$

Our focal independent variable is S_{ijt}, which represents the degree of scarcity of reward (i, j) at time t. We use the lagged stock level relative to the total supply to measure the degree of scarcity. Specifically, if a reward has 100 in stock with a total availability of 1,000, then ; when a reward option is not limited, S_{ijt} is set to be 1 throughout the fundraising cycle. Thus, a smaller S_{ijt} indicates a higher degree of scarcity. We base our analysis on the relative scarcity (i.e., the stock relative to its supply) rather than the absolute scarcity (i.e., the exact stock level), as it makes reward options at different availability levels more comparable. In addition, this measure is consistent with the literature in psychology that suggests that the relative scarcity—how much the object remains in relation to the total availability—is an important factor when speaking of an individual's perceptions of the availability of an object (Gurr 2015). We include the total availability Q_{ij} for each reward to control for any systematic differences in rewards' intrinsic attractiveness at different supply levels. For example, some rewards that are restricted to a small number of backers may be attractive due not to the effect of scarcity but, rather, to the fact that they are inherently

more valuable. Because this quantity is missing for those rewards without any quantity restriction, we discretize this continuous variable along its quantile into four groups, and we specify an additional group to cluster all of the unlimited rewards, which will absorb any systematic differences between the limited and unlimited rewards.

It is important to note that the degree of scarcity S_{ijt} also can be regarded as an alternative measure of the popularity of a reward. When individuals are uncertain about the value of an option, they may infer value from observing others' investment decisions. Therefore, the favorableness of rewards with lower relative stock levels may be attributed not to the effect of scarcity but, rather, to individuals' increased evaluation through observational learning (Zhang and Liu 2012). We disentangle the effect of scarcity from peer influence by including $Y_{ij,t-1}$, the lagged number of attracted backers for each reward. This variable reflects the previous backers' collective investment decisions of that reward option and, thus, may serve as a valid quality signal for subsequent individuals when they are making decisions. By including this lagged number of predecessors into the model, we explicitly control for the effect of previous peer action on subsequent backers' evaluation of the reward. Therefore, any remaining effect caused by the extent of scarcity can no longer be explained by peer influence. In addition to the above main independent variables, we include an extensive set of controls in our analysis. The comprehensive list of covariates we included in our model is presented in Table 4.2.

X is a set of covariates that represent campaign-level characteristics, including the fundraising target and campaign duration; Υ_{ij} is a set of covariates that represent the reward-level characteristics consisting of the expected delivery date, price, and price deviation, which is the difference between the reward price and average price across all reward levels within a campaign, i.e., $\text{PriceDeviation}_{ij} = \text{AvgPrice}_i - \text{Price}_{ij}$. This variable controls for arrival rate differences due to within-campaign price variation.

H is a set of environmental time-varying factors that influence backers' arrival dynamics. In this variable set, we include the days elapsed since the campaign's launch, which we denote as t, and day-of-week dummies to control for the time effects. In accordance with

Table 4.2: List of Covariates

Notation	Coeff.	Control Variable	Description
X_i	ϕ_1	Fundraising Target	Campaign' fundraising goal, in CNY.
	ϕ_2	Duration	Campaign's fundraising duration, in days.
Q_{ij}	θ_2	Available Quantity	A series of dummy variables for the total availability of the rewards, discretized based on the 25th, 50th, and 75th quantiles; rewards without limitation are represented by a separate dummy.
Υ_{ij}	α_1	Price	Reward's price, in CNY.
	α_2	Price Deviation	The difference between the price of a reward and the average price across all reward levels within the same campaign.
	α_3	Estimated Delivery	Reward's estimated date of delivery since the goal is reached, in days.
S_{ijt}	θ_1	Scarcity Measure	Reward's current stock level relative to its total availability, ranging from 0 to 1.
$Y_{ij,t-1}$	θ_3	Lagged Total Number of Backers	The cumulative number of attracted backers at time $t-1$.
H_{ijt}	η_1	Day Elapsed	The days elapsed since campaign i's launch at the observation day.
	η_2	Success	1 if a campaign has reached the fundraising goal at time $t-1$.
	η_3	Stock-out Spillovers	1 if any other reward option within campaign i is unavailable due to excess demand at time t.
	η_4	Within Campaign Spillovers	The average number of backers for all rewards other than the reward j within campaign i at time t.
	η_5	Across Campaign Spillovers	The average funded amount for all campaigns other than campaign i in the platform at time t.
		Day-of-Week	A series of dummies indicating the day of the week.

prior research on the impact of payoff externality on backers' behavior (Burtch et al. 2018, Zhang and Liu 2012), we also include a success indicator, which takes the value of 1 if the campaign has reached the fundraising target to account for any structural change in backers' behavior after the campaign has been fully funded. Further, we consider three types of spillover effects that can arise from the interactions among rewards within the same campaign as well as across campaigns at the platform level:

Within campaign spillovers are the average number of attracted backers of rewards at fundraising day t for rewards $(i, -j)$, which are those from the same campaign as the focal reward but are not the focal reward. This construct controls for the contemporary influence of the popularity of other reward levels within the campaign on backers' investment intention toward the focal reward (i, j).

Across campaign spillovers are the average number of daily attracted backers of other active campaigns $(-i)$ in this platform but not the campaign to which the focal reward belongs. This construct controls for the potential spillover effects across campaigns at the platform level.

Stock-out spillovers are an indicator variable that takes the value of 1 if there is any reward within the campaign that is no longer available due to the quantity limit of the active fundraising cycle. The coefficient of this variable measures how the unavailability of a reward influences backers' investment intention toward other rewards within the campaign. This influence could be either harmful or beneficial. On the one hand, it may signal positive information to individuals about the quality or attractiveness of the campaign, incentivizing new backers who would otherwise not make a contribution. On the other hand, having a reward that is no longer unavailable reduces potential backers' choice set, and they may perceive themselves a not exclusive or distinctive due to not being able to choose a certain option, which may reduce their willingness to contribute.

Column (1) in Table 4.3 presents the estimates for the baseline model. The coefficient for the scarcity measure is negative and significant ($\theta_1 < 0$), suggesting that individuals' investment intention increases when a reward becomes scarcer (i.e.,S_{ijt} decreases). This

positive effect, however, could be driven by different motivational mechanisms that underlie individuals' inference processes, or it may be due to the reward heterogeneity that is not observed by the researchers. Below, we detail our identification strategy to validate the existence of such an effect and to unravel the underlying mechanisms in a rigorous manner.

Table 4.3: Main results

		(1) Homogeneous Model		(2) Two-level Hierarchical Model		(3) Three-level Hierarchical Model		(4) Three-level Hierarchical Model with Interaction	
Notation	Variable	*Mean*	*Est.Error*	*Mean*	*Est.Error*	*Mean*	*Est.Error*	*Mean*	*Est.Error*
	Population Level Coefficients								
$\bar{\theta}_0$		1.07∗	0.65	9.81∗	1.64	5.30∗	1.58	6.98∗	1.61
$\bar{\theta}_1$	Scarcity	−0.92∗	0.12	−0.57∗	0.11	−0.62∗	0.13	−1.25∗	0.26
$\bar{\theta}_2$	Total Availability (100, 300]	1.27∗	0.12	1.04∗	0.14	1.02∗	0.27	1.08∗	0.27
	Total Availability (300, 500]	1.43∗	0.12	1.16∗	0.14	1.14∗	0.28	1.21∗	0.27
	Total Availability (500,)	1.67∗	0.12	1.35∗	0.15	1.04∗	0.29	1.13∗	0.29
	Not Limited (Yes = 1)	0.78	0.14	−0.07	0.69	−0.26	0.74	−0.22	0.74
$\bar{\theta}_3$	Lagged Total Number of Backers	0.36∗	0.01	0.29∗	0.01	0.29∗	0.02	1.25∗	0.11
φ_1	Fundraising Target	0.37∗	0.02	0.54∗	0.09	0.35∗	0.09	0.36∗	0.09
φ_2	Duration	−0.31∗	0.05	−0.80∗	0.22	−0.61∗	0.21	−0.60∗	0.22
α_1	Price	−0.47∗	0.01	−0.70∗	0.06	−0.69∗	0.06	−0.70∗	0.05
α_2	Price Deviation	−0.14∗	0.02	0.14∗	0.06	−0.43∗	0.06	−0.43∗	0.06
α_3	Estimated Delivery	−0.13∗	0.02	−0.1	0.09	−0.15	0.1	−0.15	0.1
η_1	Days Elapsed	−0.14∗	0	−0.10∗	0	−0.09∗	0	−0.09∗	0
η_2	Success (Yes = 1)	0.62∗	0.03	0.03∗	0.05	0.22∗	0.04	0.21∗	0.04
η_3	Stock-out Spillovers	1.34∗	0.28	1.22∗	0.28	0.72∗	0.24	0.95∗	0.23
η_4	Within Campaign Spillovers	1.07	0.65	−0.30∗	0.02	0.56∗	0.02	0.56∗	0.02
η_5	Across Campaign Spillovers	−0.09∗	0.05	−0.68∗	0.07	−0.44∗	0.07	−0.43∗	0.06
$\bar{\theta}_4$	Scarcity Measure × Lagged Total Number of Backers							0.92∗	0.11
	Day-of-Week Fixed Effects	Yes		Yes		Yes		Yes	
	Distribution Parameters								
φ	Over-Dispersion Rate	0.19∗	0	0.26∗	0	0.41∗	0.01	0.41∗	0.01
π	Zero-Inflation Rate	0	0	0	0	0	0	0	0

† Note. *95% of the posterior interval does not contain 0.

4.2.2 Identification Strategy

Unobserved Hierarchical Heterogeneity

Although we have incorporated a rich set of covariates, it is unlikely that these variables are comprehensive in capturing users' preference and utilities across different rewards. As shown in Figure 4.3, the number of daily attracted backers for different rewards within the same campaign are highly correlated even after controlling for both shared and distinct reward characteristics. Thus, an alternative causal explanation is that the popularity of a reward is due to an inherent attractiveness that we cannot observe.

To account for this nested unobserved heterogeneity, we extend our baseline model to a hierarchical Bayesian reward random effects framework. We take into account the unobserved reward characteristics by specifying a reward-specific intercept, and we incorporate the heterogeneous effects of scarcity, quantity limit, and peer influence by allowing the corresponding coefficients to vary across different rewards. The model specification is given as follows:

$$\log(\mu_{ijt}) = \theta_{0ij} + \theta_{1ij} S_{ijt} + \theta_{2ij} Q_{ij} + \theta_{3ij} Y_{ij,t-1} + X_j \cdot \Phi + \Upsilon_{ij} \cdot \alpha + H_{ijt} \cdot \eta + \varepsilon_{ijt}, \qquad (4.4)$$

where $\theta_{ij} = [\theta_{0ij}; \theta_{1ij}; \theta_{2ij}; \theta_{3ij}]'$ is the vector of random coefficients for reward (i, j). Our model accommodates the observed dependency structure of our data by specifying the prior probability distribution of θ_{ij}, following a nested structure. Specifically, we assume $\theta_{ij} \sim N(\Theta_i, \Sigma_R)$, where the term $\Theta_i = [\theta_{0i}; \theta_{1i}; \theta_{2i}; \theta_{3i}]'$ is a vector of campaign-specific mean, and Σ_R is the covariance matrix of the reward-level random coefficients. The campaign-specific mean $\Theta_i \sim N(\theta, \Sigma_C)$ is further distributed around a population mean Θ with a covariance matrix Σ_C. Thus or each reward-specific coefficient in θ_{ij}, there are two variance components: among rewards within campaign Σ_R, and the variation between campaigns Σ_C. For simplicity, we assume that both Σ_R and Σ_C have a diagonal structure and are homogeneous across different units. Figure 4.6 provides a graphical summary of the structure of our random effects framework.

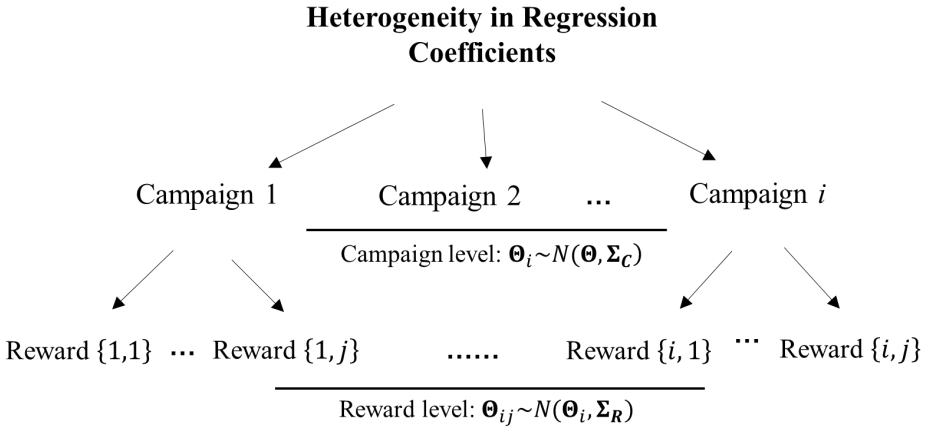


Figure 4.6: Structure of Random Effects

Compared to other frequentist approaches, the above multi-level random effects model provides us with a more flexible framework to model the complex dependency structure. Moreover, the hierarchical model allows the variation of random effects at different levels (i.e., Σ_R and Σ_C) to be estimated from the observed data. Thus, this model should perform no worse than would an OLS approach with no random effects (i.e., $\Sigma_R, \Sigma_C = 0$), or a conventional fixed effects approach that assumes the variation of the random effects to be infinite (i.e., $\Sigma_R, \Sigma_C = \infty$) (Calvin 1994; Gelman et al. 1995). Finally, from a methodological point of view, Bayesian methods provide the posterior parameter estimation, using sampling techniques that do not rely on asymptotic theories and that could provide more reliable results when a large sample size is not available(Smith 1973).

Prevention-Motivated vs. Promotion-Motivated Backers

The quantity scarcity of a reward is a consequence of excess demand or limited supply. Depending on individuals' underlying motivational goal, their susceptibility to these two types of scarcity may differ. Prevention-motivated backers prefer demand-induced scarce rewards to obtain a sense of security, whereas promotion-motivated backers desire uniqueness and personal advancement, which drives them to invest in rewards that are scarce due to

limited supply. Note that JD crowdfunding makes the current demand for a reward option, as well as the total availability, visible to all of the potential backers, as shown in Figure 2. Such a website design makes both of the inference processes possible. Nevertheless, Equation (4.4), by itself, cannot uncover which behavioral mechanism dominates, as both mechanisms would lead to an observed increased subjective desirability among backers.

We distinguish these two alternative mechanisms through augmenting the first-level equation with an interaction term between the scarcity measure S_{ijt} and the lagged accumulative number of backers for each reward $Y_{ij,t-1}$. Specifically, the first level equation in our multilevel model becomes:

$$\begin{aligned}\log(\mu_{ijt}) = \theta_{0ij} + \theta_{1ij}S_{ijt} + \theta_{2ij}Q_{ij} + \theta_{3ij}Y_{ij,t-1} + \theta_{4ij}S_{ijt}\cdot Y_{ij,t-1} \\ + X_j\cdot\Phi + \Upsilon_{ij}\cdot\alpha + H_{ijt}\cdot\eta + \varepsilon_{ijt}.\end{aligned} \tag{4.5}$$

If backers are prevention-motivated, the momentum brought by the scarcity of the reward (i.e., S_{ijt}) would be amplified by the lagged total number of backers $Y_{ij,t-1}$, leading to a positive $\bar{\theta}_4$. If backers are promotion-motivated, however, the attractiveness of a limited offering would be attenuated by $Y_{ij,t-1}$, as it directly indicates the number of individuals who have already obtained that offer. Thus, in this case, we would expect $\bar{\theta}_4$ to be negative.

4.3 Results

4.3.1 Divergence Diagnostics and Posterior Predictive Check

We estimate our model, using a Markov Chain Monte Carlo (MCMC) sampler, Hamiltonian Monte Carlo, implemented in Rstan (Carpenter et al. 2017). This sampler converges more quickly than do other MCMC algorithms, such as Gibbs sampling and Metropolis-Hastings, and explores the parameter space more efficiently (Hoffman and Gelman 2014). An introduction to Hamiltonian Monte Carlo can be found in Betancourt et al. (2017). Our MCMC sampling process consists of four different Markov chains sampled for 10,000 iterations. The first half of the draws is discarded during the warm-up period, and we use the remaining 5,000 draws in each chain to construct the posterior samples. The result of our estimation is

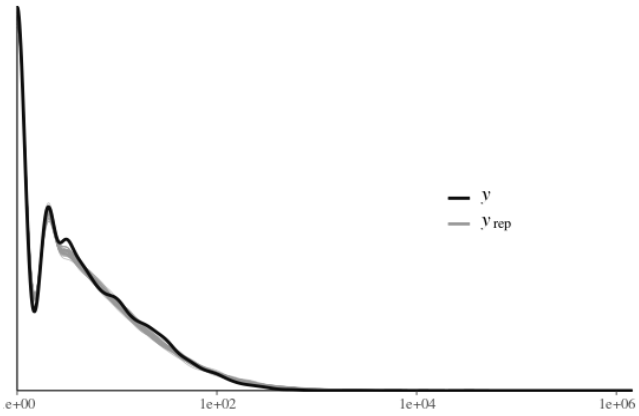


Figure 4.7: Posterior Predictive Distribution

a set of posterior distributions, approximated by the sample draws of each unknown parameter. To assess the convergence of our Markov chains, we calculate the within-chain variance and pooled variance across chains. The resulting potential scale reduction factors ($\hat{R}$) for all of the parameters are very close to 1[2], which is indicative of chain convergence (Gelman et al. 1995).

Given that our sampler has successfully converged, we then proceed to examine the fitness of our proposed model to ensure that the related assumptions are plausible and that it provides a good summary of the data at hand. To do so, we conduct a series of posterior predictive checks by simulating a replicated dataset (y_{rep}) under our fitted model and then comparing them to our observations (y). Any systematic discrepancies between the real and simulated data indicate the potential failure of the proposed model (Gelman et al. 1995). Because each estimated parameter is a distribution that involves uncertainty, we draw 10 different samples from the posterior parameter space, and each sample gives us a unique set of parameters to simulate y_{rep}. We plot the simulation for each draw in Figure 4.7. As seen, all of the sampled posterior distributions fit our observation, indicating the fitness of our model.

[2]Results will be provided upon request.

4.3.2 Model Comparison

Our full model allows the reward random effects to have a nested structure, and we include an interaction term to disentangle backers' motivational mechanism. Before we discuss the estimation results, we first demonstrate the fitness of our proposed model by comparing our full model with its three nested versions: Model 1, a homogeneous ZINB model without random effects or the interaction term; Model 2, a hierarchical ZINB model with campaign-specific random effects without the interaction term; and Model 3, a hierarchical ZINB model with nested reward-level random effects without the interaction term.

The model fit for the above three models and for our full model is evaluated using leave-one-out (LOO) cross-validation and the widely applicable information criterion (WAIC). These two measures estimate the pointwise *out-of-sample prediction accuracy* from a fitted Bayesian model, using the log-likelihood evaluated at the posterior simulations of the parameter values. LOO and WAIC have various advantages over simpler estimates, such as the Akaike information criterion (AIC) and deviance information criterion (DIC) that have been widely used in the previous literature (Vehtari and Gelman 2014, Vehtari et al. 2016). In this manner, as Table 4.4 shows, we get the best fit of our full model. The model performs worse when we exclude the interaction term (Model 3), indicating the need to account for the moderating effect. Moreover, the performance drops significantly if we ignore the hierarchical structure (Models 1 and 2). This result provides us with strong evidence of different levels of heterogeneity, suggesting the need to employ a multi-level model such as ours.

4.3.3 Discussion of Estimates

The results of our full model, as specified in Equation (4.5), is reported in Column (4) of Table 4.3. To show the robustness of our results, we also report the estimates for the three nested versions as described in the model comparison section (Columns (1)–(3) of Table 4.3). Because the coefficients are estimated, using a Bayesian approach, the significance levels of the parameters are based on the posterior intervals, which describe the information

Table 4.4: Model Fit and Comparison

Statistic	Model 1	Model 2	Model 3	Full Model
LOO	115,803.90	110,762.70	104,603.50	104,603.50
WAIC	115,822.80	110,679.70	104,183.40	104,117.10
ΔLOO w.r.t. full model	11,200.40	6,159.20	0	
ΔWAIC w.r.t. full model	11,705.70	6,562.60	66.3	

that we have about the location of the parameters after observing the data. A parameter is considered to be significant if the posterior credible interval does not contain zero (Gelman et al. 1995).

All of the estimates and their significance levels remain qualitatively the same across these specifications. As expected, the coefficient for the scarcity measure is negative and significant ($\bar{\theta}_1 = -1.25$), which confirms the positive effect of the scarcity strategy on the crowdfunding market, i.e., a reward option with a lower relative stock level (i.e., S_{ijt} decreases) helps it to attract more backers. Interestingly, the coefficients for the total availability dummies (the $\bar{\theta}_2$s) suggest that the supply level has no significant monotonic effect on backers' investment intention after we control for the unobserved reward heterogeneity and various potential confounders. One interpretation is that restricting the total supply itself does not add extra value to the rewards, especially when individuals are facing a wide range of available options with diversified features in the crowdfunding markets. The estimated coefficients for *Lagged Total Number of Backers* ($\bar{\theta}_3 = 1.25$) suggests a positive peers' effect; i.e., backers tend to perceive the rewards with higher prior capital accumulation to be more valuable. This finding echoes the evidence of individuals' observational learning behaviors in lending-based crowdfunding markets (Zhang and Liu 2012).

Our estimated coefficients for time-varying, environmental factors provide some interesting insights, e.g., individuals' investment intentions decrease over time during the fundraising

cycle ($\eta_1 = -0.09$), possibly reflecting a declining attention due to the ranking effects. The coefficient for the dummy variable success is positive and significant ($\eta_2 = 0.21$). This suggests that individuals exhibit higher investment momentum toward campaigns that already have reached their fundraising goals. This is expected, as the potential backers may interpret the success of a campaign as a quality signal, which, by reducing their uncertainty, motivates them to contribute. Interestingly, we observe a positive stock-out spillover effect ($\eta_3 = 0.95$). Our interpretation is that individuals' inability to participate in an offering will amplify their momentum to obtain the focal fundraising product, which leads to a higher contribution intention to other available reward levels within that campaign.

The coefficient for the within-campaign spillovers is positive and significant ($\eta_4 = 0.56$), implying that the popularity of a reward option can positively influence other rewards under the same campaign, which is intuitive and consistent with our expectation. The coefficient for across-campaign spillovers is significantly negative ($\eta_5 = -0.43$), which indicates a substitution demand pattern among campaigns on the platform: backers' attention will be directed to the campaigns that have attracted much traffic, causing a reduced demand for other fundraisers. Some additional insights can be drawn from our results. Backers' investment intention is positively related to the fundraising target ($\phi_1 > 0$) and negatively related to the fundraising duration ($\phi_2 < 0$). In addition, both a higher price and a higher relative price deviation are found to discourage contribution ($\alpha_1, \alpha_2 < 0$).

The above estimation results help us to confirm the existence of the scarcity effect in the crowdfunding market and to establish links between backers' investment intention and various influential factors. Next, we turn to the issue of the underlying motivational mechanism that gives rise to the observed scarcity effect in this market. This issue can be answered by examining the estimate of the interaction term between *Scarcity Measure* and *Lagged Total Number of Backers* in Column (4) of Table 4.3. The estimated coefficient is positive and significant ($\bar{\theta}_4 = 0.92$), which suggests that the persuasive power of the scarcity of a reward increases if the scarcity itself is due to the demand from a larger number of backers. This finding reveals that individual backers in this crowdfunding platform are mainly prevention-

motivated, or risk averse, and would prefer the rewards that have been chosen by more predecessors to reduce their sense of insecurity.

Finally, the estimated coefficients of the nested group-level random effects for our full model are reported in the appendix. As seen, all of the random effects are significant. These estimates suggest considerable heterogeneity both across and within campaigns; thus, it is necessary to employ a hierarchical model such as ours to take the nested unobserved heterogeneity into account.

4.3.4 Robustness Checks

Sub-Sample Analysis

Our estimation results suggest that individuals are prevention-motivated and prefer rewards that are scarce due to excess demand rather than limited supply. From a managerial perspective, however, it is also necessary to understand whether the magnitude of such an effect, as well as the identified dominant mechanism, varies across campaigns with different quality levels. Intuitively, high-quality campaigns would reduce individuals' uncertainty and, thus, make them become less risk averse. In this case, the effect of existing demand among these campaigns could be attenuated, and it is even possible that individuals would, instead, be more susceptible to supply-induced scarcity.

To investigate this possibility, we repeated the estimation of our full model, breaking down our data into two subsamples based on the mean of the fundraising target, which serves as a proxy for the potential quality of the campaign (Chakraborty and Swinney 2017). The estimation results are presented in the first two columns of Table 4.5. The new parameter estimates are not substantively different from those of the full model, suggesting that the findings are consistent across different types of campaigns.

Table 4.5: Robustness Checks

	Target <50,000		Target >50,000		Linear Model	
Variable	Mean	Est. Error	Mean	Est. Error	Mean	Est. Error
Population Level Coefficients						
Intercept	5.57∗	2.09	6.05∗	2.46	0.84∗	0.07
Scarcity	−0.81∗	0.4	−1.16∗	0.3	−1.05∗	0.01
Total Availability (100, 300]	0.96∗	0.32	0.67∗	0.31	0.04∗	0.01
Total Availability (300, 500]	0.77∗	0.33	1.08∗	0.32	0.04∗	0.01
Total Availability (500, ∞)	1.00∗	0.35	0.69∗	0.32	0.07∗	0.01
Not Limited (Yes = 1)	0.41	0.46	−0.81	0.4	0.09	0.01
Lagged Total Number of Backers	1.27∗	0.19	1.17∗	0.13	0.01	0
Fundraising Target	0.61∗	0.14	0.39∗	0.17	−0.01	0
Duration	−0.70∗	0.28	−0.45	0.27	0	0.01
Price	−1.00∗	0.1	−0.61∗	0.06	−0.01∗	0
Price Deviation	−0.17	0.11	−0.53∗	0.07	0	0
Estimated Delivery	−0.1	0.14	−0.15	0.13	0.01	0
Days Elapsed	−0.08∗	0	−0.09∗	0	0	0
Success (Yes = 1)	0.44∗	0.06	0.1	0.06	0	0
Stock-out Spillovers	1.56∗	0.31	0.39	0.27	0.12∗	0.02
Within Campaign Spillovers	0.43∗	0.04	0.61∗	0.03	0	0
Across Campaign Spillovers	−0.34∗	0.08	−0.45∗	0.08	0.01	0
Scarcity Measure × Lagged Total Number of Backers	0.86∗	0.18	0.90∗	0.12	1.02∗	0
Day-of-Week Fixed Effects	Yes		Yes		Yes	
Distribution Parameters						
Over-Dispersion Rate	0.54∗	0.01	0.37*	0.01		
Zero-Inflation Rate	0	0	0	0		
Sigma					0.14∗	0

†Note. *95% of the posterior interval does not contain 0.

Alternative Measure of Backers' Intention

In our main model, we use a count dependent variable, the total number of backers whom a reward has attracted within a day, as the measure for backers' intention and estimate the effects, using a non-linear model. To assess how sensitive our main results are to the choice of dependent variable and the model specification, we use the Total Amount of Funds that a reward has attracted within a day as an alternative measure of backers' intention. We re-conduct the analysis on the following multi-level linear model:

$$\begin{aligned}\log(y_{ijt}+1) = \theta_{0ij} + \theta_{1ij}S_{ijt} + \theta_{2ij}Q_{ij} + \theta_{3ij}Y_{ij,t-1} + \theta_{4ij}Y_{ij,t-1}\cdot S_{ijt} \\ + X_j\cdot\Phi + \Upsilon_{ij}\cdot\alpha + H_{ijt}\cdot\eta + e_{ijt}\end{aligned} \tag{4.6}$$

Because the amount of attracted funds (y_{ijt}) is non-negative and contains zero, we take the logarithm of it with an offset of 1. The included covariates and their notations are identical to those in our main model. We present the estimation results in the third column of Table 5. The signs of $\bar{\theta}_1$ stays negative, showing that there is, indeed, a statistically significant scarcity effect. Consistent with our previous findings, backers are more motivated to invest in campaigns that have attracted a larger number of backers, which is indicated by the positive and significant $\bar{\theta}_4$, confirming the robustness of our main findings.

4.3.5 Predictive Ability

To assess the predictive ability of our model, we randomly partition our data set into an 80% calibration set and a 20% holdout set. We first re-estimate the model on the calibration sample to get the posterior parameter distributions. The estimation results are reported in the appendix as an additional robustness check. Then, we draw 10 independent samples from the estimated posterior parameter space, which we use to make predictions on the holdout sample. We evaluate the predictive power by comparing the zero-inflation and over-dispersion rates in the predicted outcomes with the true observation, as seen in Figure 4.8. As seen, our model is particularly good at predicting these two salient features in new data.

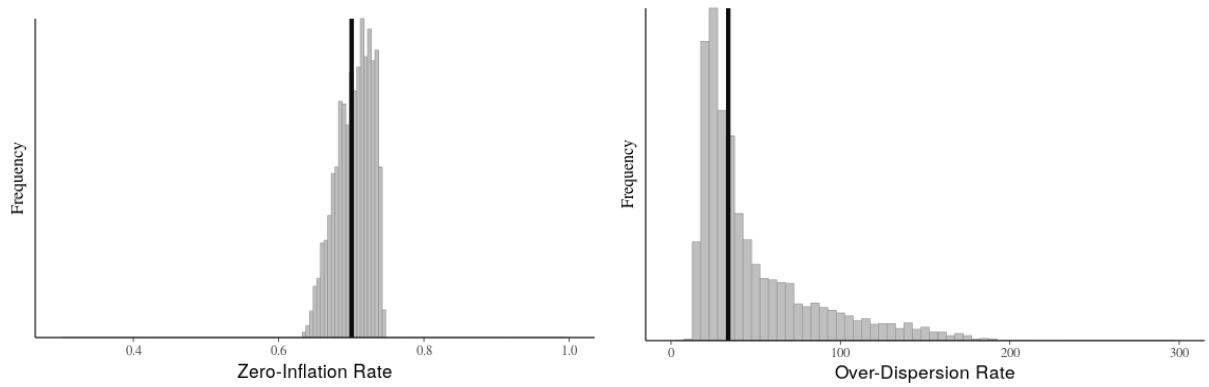


Figure 4.8: Statistics for the Predictions on the Holdout Sample

4.3.6 Managerial Implications

This study provides important managerial implications for platform managers, website designers, and fundraisers who attempt to facilitate their fundraising success through an effective reward scheme design. Several actionable strategies can be drawn from our empirical results. For example, we find that the relative scarcity of a reward, after controlling for unobserved heterogeneity, individuals' observational learning behavior, and other potential confounders, significantly influences backers' investment intention. This suggests the persistence of the positive effect of the scarcity strategy in the online crowdfunding market, and, thus, it is desirable to use this strategy to promote sales at the platform level. Our results also reveal that, given the relative scarcity level, individuals are more driven to obtain demand-induced scarce rewards. An important implication for platform managers and website designers is that, in addition to emphasizing the current stock level, it pays to highlight the existing demand for each open reward option to best harness the persuasive power of scarcity.

Designers of crowdfunding marketing websites also can benefit from the findings on the interactions among rewards and campaigns. The negative across-campaign spillover suggests

a substitution effect across different campaigns. In light of this result, platform managers should be cautious when there are some campaigns that have attracted much traffic. When they do, platforms should help create awareness for the fundraisers who face reduced demand in order to maintain the efficacy and the efficiency of the entire market.

Our estimation results also allow us to evaluate the magnitude of the scarcity effect. Suppose we add a quantity restriction of 500 for an unlimited reward that has already attracted 200 backers. Column (4) in Table 4.3 suggests that, other things held constant, the implementation of this strategy will, on average, increase backers' mean arrival rate by $((500 - 200)/500 - 1) \times (-1.25) = 0.5$, suggesting a marginal effect. Thus, for fundraisers who anticipate their ability to attract a large number of backers, this result indicates that the effectiveness of putting a tight limitation on the supply may be harmful, as it weakens the attractiveness signal that could be conveyed once the pre-set limit has been reached. However, for narrow-appeal campaigns that have limited ability to attract sufficient contributors at an early stage, it pays to put a quantity restriction on their rewards options. Although the daily effect is marginal, it could still yield significant influence on fundraising performance in the long run, as this momentum will be compounded in the presence of herding behaviors (Zhang and Liu 2012).

4.4 Discussion and Conclusion

The scarcity strategy, which is popular among advertisers and retailers in various offline markets to promote sales, has also become popular with fundraisers on the online crowdfunding platforms. These fundraisers intentionally put a quantity restriction on some of their reward options, with an attempt to increase individuals' subjective desirability toward their products. Despite the eagerness of these fundraisers to integrate this strategy into their reward scheme design to facilitate the success of their fundraising campaigns, its effectiveness has not yet received rigorous examination. Nor do we know much about how this strategy influences individuals' underlying evaluation process on which they rely to make investment decisions.

We address this research gap by investigating backers' behavior patterns in a leading reward-based crowdfunding platform in China. Utilizing a unique dynamic panel data set that consists of reward-level fundraising dynamics and static campaign characteristics, we validate the positive effect of the scarcity strategy in the crowdfunding market, using a hierarchical Bayesian framework. This proposed model allows us to flexibly capture the nested heterogeneity across rewards to yield a more eloquent empirical analysis. We also empirically unravel the underlying behavior mechanism that governs individuals' evaluation processes by examining the moderating role of previously accumulated backers on the scarcity effect. We find evidence that individual backers are prevention-motivated, i.e., are more susceptible to demand-induced scarcity. Given the relative scarcity level, individuals are more attracted to invest in rewards that achieve this scarcity due to excess demand rather than limited supply. The results of our study also reveal the interactions among campaigns as well as different reward levels within a campaign.

Although one of our main findings echoes some prior research that documents the positive effect of the scarcity strategy as a marketing tool, this study provides a novel and valuable contribution to the existing literature in the following ways. First, we extend prior work on the influence of the scarcity strategy to a novel context, the crowdfunding market, wherein the social aspect of individuals' online behavior is prominent, and their decisions involve uncertainty in the presence of information asymmetry. Second, we empirically uncover the dominant motivational mechanism that governs backers' evaluation process of the JD crowdfunding market. This finding is critical to an effective reward scheme design that facilitates fundraising success and adds to the literature on the optimal way to communicate information with potential backers to best harness the power of scarcity. Finally, our study contributes to the existing crowdfunding literature from a methodological perspective. All of the analyses in this study are conducted at the reward level. This novel perspective, which has not been implemented in prior work, which allows us to control for another level of heterogeneity and significantly improves the model fitness and prediction power.

Despite the contributions presented above, there are several limitations of this study that

provide avenues for future research. First, although we provide insight into the trade-off between the momentum brought by imposing quantity limitation and the possibility of restricting upcoming backers once the pre-set limit has been reached, we did not look into how to design an optimal reward scheme that takes into account the totality of reward options with different levels of popularity. Second, due to the limited sample size, we assumed a homogeneous consumer type in the crowdfunding market to reduce the complexity of our model. Thus, although our work can be seen as the first attempt to empirically quantify the effect of scarcity with behavioral explanations, more can be done to analyze whether this behavior pattern is consistent across different campaign categories. It also would be interesting to look into whether this behavior pattern varies dynamically over the fundraising life cycle. Finally, the limiting of our analysis to a certain platform may reduce the applicability of our results to other reward-based crowdfunding sites and to other types of crowdfunding markets. Thus, future research could include a variety of platforms. Overall, however, we believe that our results provide insight for crowdfunding platform managers, as the basic principles that drive individuals' decisions in the presence of social influence, under information and investment uncertainty, are similar.

Chapter 5

HOW DO TUMOR CYTOGENETICS INFORM CANCER TREATMENTS? DYNAMIC RISK STRATIFICATION AND PRECISION MEDICINE USING MULTI-ARMED BANDITS

5.1 Empirical Setting and Data

In this essay, we are particularly interested in automated and personalized decision support among patients with relapsed multiple myeloma, which is the second most common hematologic malignancy, with over 30,000 new cases diagnosed annually in the United States. Myeloma is a cancer of plasma cells that reside primarily in the bone marrow. Patients with multiple myeloma may present with anemia, renal failure, lytic bone lesions, and elevated calcium levels (Kyle and Rajkumar 2008).

While a cure remains elusive, the treatment regimen for patients with MM has endured dramatic changes within the past decade. At present, at least nineteen drugs are available, spanning eight pharmacologic classes. This extensive treatment arsenal provides a new possibility of improving patients' expected overall survival, making it possible to turn myeloma into a manageable condition (Kumar et al. 2008). However, it has also resulted in physicians' difficulties in deciding the best treatment combination to give next for each patient, as no head-to-head comparison has been made among all alternatives (Van Beurden-Tan et al. 2016). This calls for a systematic way to evaluate the efficacy of all combinations and facilitate medical decision-making, which motivates this study.

We work with Seattle Cancer Care Alliance (SCCA) and use the clinical dataset comprising the medical history for 803 patients with MM treated at their facility. When patients come to SCCA for cancer treatment, they will be cared by a team of specialists, including surgeons, medical oncologists, fellows, and advanced practice providers. Thus, our clinical

dataset consists of the EMRs for these 803 patients across the whole system. Since many of the patients came to the center seeking an autologous stem cell transplant (SCT), most of the patients were transplant eligible. Informed consent was obtained from all patients in accordance with protocols approved by the Institutional Review Board at the Fred Hutchinson Cancer Research Center. The dataset was de-identified before analysis to protect patient privacy. The full dataset contains a complete set of follow-up records for the 803 patients over seventeen years from 2000 to 2017.

Patients with MM usually receive sequential lines of therapies (LOT) during their treatment journey. Each LOT consists of one or more drugs given cyclically over time. Following the initial therapy, when there is clinical sign of relapse, a new line of therapy is initiated, which is signaled by a break with the previous treatment plan and adoption of a new regimen. This happens when the cancer becomes resistant to the treatment or the patient experiences unacceptable toxicity (Kyle and Rajkumar 2009). Over time, the periods of remission generally become shorter in duration until the patient has exhausted all available therapy or has died from the disease (Figure 5.1). A new line of therapy is flagged with the occurrence of any of the following events: 1) starting a new combination of treatment; 2) the addition of a new drug after the previous LOT; 3) a regimen is interrupted or discontinued, and then restarted later with a modified regimen with the addition of new drugs, or one or more regimens were administered in between. However, if a dosage changes, or a regimen is discontinued, but the same combination is restarted without any intervening mechanism, it is considered the same line. In this way, we end up identifying 3,762 LOTs in total for the 803 MM patients.

For all the 803 patients, we abstracted their demographic information, including gender and the age at diagnosis. For each patient per clinic visit, three categories of information is observed: 1) laboratory testing results, 2) treatment decision, and 3) the treatment efficacies. A patient visit is first described by their latest laboratory testing results. Patients with multiple myeloma underwent a broad range of clinical testing at different time points between visits for physicians to monitor disease progression or treatment-related toxicity.

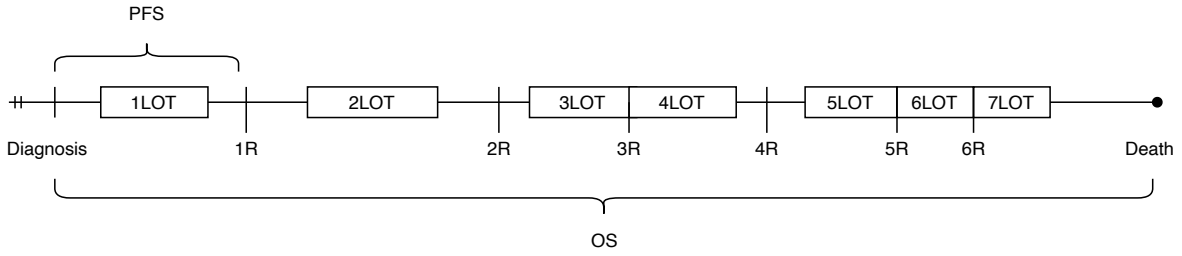


Figure 5.1: Event Endpoints. [a]

[a]The patient's initial visit establishes a diagnosis, which is indicated by the end point 1 labelled as "Diagnosis". A re-visit following the first diagnosis occurs when there is a clinical sign of relapse, which usually indicates the previous treatment fails to control patient's disease progression. This event indicates a relapse, and is indicated by the following events points after the diagnosis labelled as "1R" – "6R". The time interval between two visits is called line of therapy (LOT), which is measured by the PFS. All the PFSs sum up to OS, which is defined by the time from diagnosis until death from any cause.

These tests include common blood and urine tests, percentage of bone marrow plasma cells, and the number of bony lytic lesions detected on skeletal X-rays or magnetic resonance imaging (MRI). We also collected fluorescence in situ hybridization (FISH) results, a type of cytogenetic test that identifies chromosomal structural abnormalities within the malignant plasma cells that contribute to the disease pathogenesis and risk stratification (Rothman et al. 2012).

Second, we observe the treatment combination administered to patients at each visit. We group all the drugs received within 30 days after the first prescribed drug into a treatment combination following the MM literature (Hari et al. 2018). Among all the existing treatment combinations, we identified nine combinations that appear most frequently in our dataset, and group all other combinations whose occurrence is fewer than ten times as a benchmark group. [1] In this way, we end up having ten different combinations in our dataset, including the benchmark group.

Third, we observe the effectiveness of treatment given to each patient. Among the various

[1]All the combinations in the benchmark group only accounted for 3.33% of the total number of treatment assignments in our record

survival measures for cancer patients, the overall survival (OS), which measures time from diagnosis until death from any cause, has been widely accepted as the most reliable and available survival measure (Mariotto et al. 2014). While the OS is the most direct measurement for patients' survival, it is oftentimes not the best choice to use it in the medical research due to several reasons. First, OS uses death as the endpoint; thus, the demonstration of an increased OS due to a new treatment requires large patient populations, several years of follow-up, and it could be confounded with other causes unrelated to the disease of interest (Dimopoulos et al. 2017). Second, each patient only has one OS; therefore, it does not allow researchers to measure the efficacy for each individual intervention given throughout the entire treatment course.

Given the above caveats, most of the researchers and physicians use progression-free survival (PFS) as an event-driven surrogate for OS (Driscoll and Rixe 2009). PFS is defined as the number of days after the treatment until a disease progression. For example, if a patient i started his second LOT on April 1, 2017 and returned with a relapse on April 28, 2017 to initialize the third LOT, then the PFS corresponding to the treatment given in the second LOT is 27 days. In contrast to OS which measures overall modalities, PFS measures efficacy a selected treatment in improving the disease diagnosis for each individual LOT. PFS has been well accepted as a valid outcome to demonstrate OS benefit in the multiple myeloma community (Marshall et al. 2016), and hence is the survival outcome measure we use in this paper. [2]

In Table 5.1, we summarize the patients' demographic features and their disease response. As we can see, patients with multiple myeloma are very heterogeneous in terms of the OS, total number of LOTs they went through, and their treatment responses (i.e., PFS) for each LOT. These observed heterogeneities highlights the needs for a personalized treatment policies learnt for each individual patient, which we will show in the following section.

---

[2]Since the date of relapse cannot always be measured directly, we approximate it as the time to next treatment (TTNT), which is defined as the time interval in days between the start of two consecutive LOTs (Hari et al. 2018). However, we acknowledge that this assumption may not always be valid, because some patients may change to an alternative LOT due to intolerance, toxicity, or financial costs.

Table 5.1: Descriptive Statistics

Statistics	**Mean**	**Std.Dev**	**Min**	**Max**	**Median**
Age at diagnosis	58.10	8.88	22	82	59
Patient gender (male = 1)	0.60	0.49	0	1	1
PFS	163.23	268.69	14	4551	93
Total LOTs	4.22	3.44	1	25	3
Overall Survival	2078.76	1318.10	0	7291	1762.5

5.2 Design of Personalized Treatment Strategy

For each patient i, we use k to denote the k-th LOT when a relapse or diagnosis occurs and a treatment decision is needed. At each decision time, the physician is presented with a set of alternative treatment combinations (i.e., the arms), given the patient i's context at the time of the visit.

We consider the physician's sequential strategy in a discrete time frame, indexed by $n = 1, 2, \cdots, N, \cdots$, consisting of visits across all the patients sorted in the chronological order. Instead of maximizing the the survival outcome for each visit myopically, the physician's objective is to sequentially select the arms to maximize the cumulative survival outcome for all the patient visits over time horizon N. Thus, this objective requires the physician's treatment strategy having a good balance between exploitation – treating a patient as efficiently as possible given the current information, and exploration – choosing a seemingly sub-optimal option to gather more information regarding alternatives and achieve long-term benefits.

We formulate the physicians' decision dilemma as a Bayesian contextual bandit problem. Before describing our proposed treatment policy in detail, we first introduce two key components that define our MAB policy: 1) a parametric likelihood function f of the survival outcome; and 2) a treatment selection rule π. These two components work together to generate a treatment selection in the following way. The model f defines the predictive

distribution of each action's expected reward (i.e., the PFS) given the patients' context; the treatment allocation rule π then maps the predictive distribution into a recommended treatment action.

To begin describing a treatment policy in the bandit language, suppose that we have set $\mathcal{A} = \{1, 2, \cdots, M\}$ of M different treatment combinations (i.e., the arms), and a time horizon of N patient visits. In trial $n \in [N]$, the following sequence of events occurs:

1. The patient i arrives with relapse, and a treatment decision is needed to initiate his k-th LOT, with k being i-specific.

2. The physician holds a prior belief on the parameters of model f denoted by $p_n(\theta)$. The prior belief was formed based on information set $\mathcal{D}_n$, which consists of the previous patients' context, treatment selection histories, and the realized survival outcomes.

3. Given the patient's context at the time of visit, a treatment $a_n \in \mathcal{A}$ is selected based on reward model $f(\theta)$ where $\theta \sim p_n(\theta)$, and the treatment selection rule π.

4. The physician observes the reward of trial y_{n+1} for treatment action a_n, and his historical information set will be enhanced from $\mathcal{D}_n$ to $\mathcal{D}_{n+1}$, which will then be used to update the belief on model parameters from $p_n(\theta)$ to $p_{n+1}(\theta)$ for the next round of decision-making.

After the N trials, we obtain a cumulative reward denoted by $R_N = \sum_{n=1}^{N} y_n$. The objective is to sequentially select treatment actions $(a_0, \cdots a_{N-1})$ to maximize R_N – the cumulative progression-free survival measured in days across all patient visits. In the following sections, we first describe our proposed model to calibrate patients' responses in Section 5.2.1; then we present our TS based treatment selection strategy in combination with the multilevel Bayesian linear model to address the challenge of balancing the exploration-exploitation trade-off to produce effective yet rapid learning.

5.2.1 A Multilevel Bayesian Linear Model for Patients' Response

MM is a heterogeneous disease with various disease courses and responses to therapy (Munshi et al. 2011) – patients' response to treatment could be very heterogeneous across individuals, and also could change rapidly throughout each patient's disease course. Therefore, a promising treatment recommendation strategy should be able to learn such heterogeneities, and provide personalized treatment recommendations that are tailored for *each* individual patient, at *each* of their visit throughout their entire disease course.

To incorporate such heterogeneities and account for dependencies across visits, we allow the treatment efficacy to vary both at the patient level, and the LOT level. Specifically, for each trial n, we map it to a patient index i, and a patient-specific LOT index k, and we consider a multilevel Bayesian linear model as the following:

$$y_{i,k} \sim \mathcal{N}(\mu_{i,k}, \sigma^2). \tag{5.1}$$

where

$$\mu_{i,k} = \omega\varphi_{i,k} + u_i\varphi^u_{i,k} + \tau_k\varphi^\tau_{i,k} \tag{5.2}$$

In the above model, the survival outcome $y_{i,k}$ is influenced by the population-level effect ω that is shared across all visits; as well as the heterogeneous effects that are dissimilar across different grouping dimensions, which are the individual-level effect u_i, and the LOT-level effect τ_k. Denote by $\varphi_{i,k}$ the context feature vector that has shared influence ω for each patient visit, we use $\varphi^u_{i,k}$ to represent the feature vector that has heterogeneous influence across different patients, and $\varphi^\tau_{i,k}$ to denote the features that have dissimilar effects across different LOTs.

It is worth noting that while the above model is flexible in capturing the heterogeneity in multiple grouping dimensions, learning a reward distribution separately for each patient i, and for each LOT k is challenging, and it could instead significantly reduce the model performance when the number of observations are sparse within each grouping level. Thus, instead of assuming the effects are disjoint for units within each group, we adopt a partial-pooling approach (Gelman et al. 2013), which allows the learning from one patient (or LOT)

to improve the learning of treatment policies for all other patients (or LOTs). This can be done by constructing hyperpriors on the distribution of the individual-level effect u_i and the LOT-level effect τ_k. Specifically, we consider τ_k as a random, LOT-specific noise normally distributed around zero mean with covariance Σ_τ. For the individual level effect u_i, we assume $u_i \sim \mathcal{N}(\bar{u}_i, \Sigma_u)$, and model the individual mean $\bar{u}_i$ as a function of their demographic features:

$$\bar{u}_i = \gamma_1 z_{i1} + \gamma_2 z_{i2} + \cdots \gamma_d z_{id}. \tag{5.3}$$

Here, for example, in our focal context z_{i1} and z_{i2} can represent individual patients' gender and age at diagnosis, respectively.

We can summarize the proposed multilevel Bayesian linear model mathematically by the following hierarchical conditional distributions:

$$\begin{aligned} y_{i,k} \mid \omega, u_i, \tau_k &\sim \mathcal{N}(\omega\varphi_{i,k} + u_i\varphi^u_{i,k} + \ \tau_k\varphi^\tau_{i,k}, \sigma^2); \\ u_i \mid \gamma &\sim \mathcal{N}(\gamma_1 z_{i1} + \gamma_2 z_{i2} + \cdots \gamma_d z_{id}, \Sigma_u); \\ \tau_k &\sim \mathcal{N}(0, \Sigma_\tau); \\ \omega &\sim \mathcal{N}\left(\bar{\omega}, \Sigma_\omega\right); \\ \gamma &\sim \mathcal{N}\left(\bar{\gamma}, \Sigma_\gamma\right). \end{aligned} \tag{5.4}$$

We denote by $\theta = \{\bar{\omega}, \bar{\gamma}, \Sigma_\omega, \Sigma_\gamma\}$ the set of unknown parameters to be learnt from the data, with $\Sigma = \{\sigma^2, \Sigma_u, \Sigma_\tau\}$ as hyper-parameters of the model. To further present our model structure, we visualize the hierarchy of our multilevel Bayesian linear framework in Figure 5.2. As we can see, although we have separate notations for u_i for each patient i, those values are normally distributed around mean $\bar{u}_i$, which is computed from γ and the demographic feature z_i. Thus, we are not obtaining the distribution of u_i separately for each individual; instead, we leverage data points from all the visits to learn each patient's response. As a result, we are able to obtain predictive reward distribution for patient even in the absence of a large number of observations. In addition, an added benefit of our multilevel Bayesian model is that we are able to calibrate the $\bar{u}_i$ as a function of their demographic features z_i.

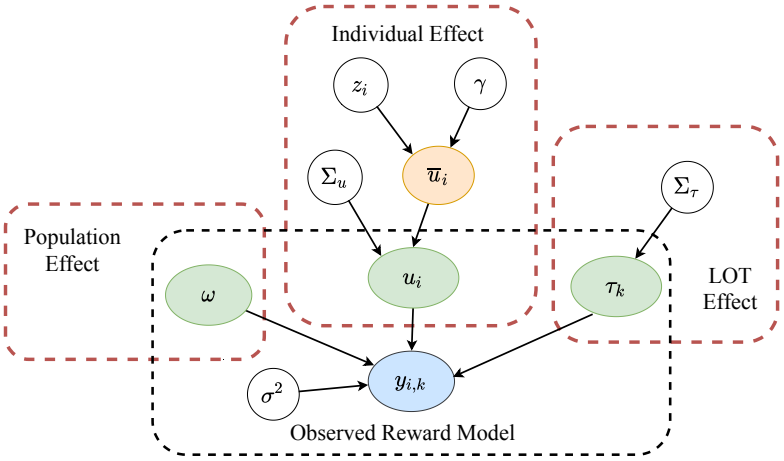


Figure 5.2: The Multilevel Bayesian Linear Model

Thus, patients with little data will be shrunk toward mean parameter vector $\bar{u}_i$, which is the average individual effect across patients with similar demographic feature.

5.2.2 Thompson Sampling and Posterior Updates

In this section, we describe our TS based treatment allocation rule for the multilevel Bayesian linear model described in Section 5.2.1. TS is a well-known bandit algorithm that balances the exploration-exploitation by randomly selecting an arm according to its posterior probability of being optimal, at any time period. The TS has experimentally been shown to provide optimal solutions and achieves logarithmic expected regret for the stochastic MAB problem (Agrawal and Goyal 2012). Recently, TS has attracted considerable attention due to its ease in implementation, flexibility in accommodating a wide range of structured decision problems in various practical settings, yet competitive performance compared to other bandit algorithms (Ferreira et al. 2018, Schwartz et al. 2017).

The TS works with our proposed multilevel Bayesian linear model in the following way. At each decision time n, we randomly sample the unknown model parameters $\hat{\theta}_n$ from the posterior distributions $p_n(\theta \mid \mathcal{D}_n)$, where $\mathcal{D}_n$ represents all the historical information

available till trial n. Given the sampled coefficients $\hat{\theta}_n$, we calculate the expected reward $\{\mu_1, \mu_2, \cdots, \mu_a, \cdots \mid \mathcal{D}_n\}$ for each action $a \in \mathcal{A}$, where $\mu_a \mid \mathcal{D}_n = E[y \mid \varphi_n, \hat{\theta}_n, a]$ given by the multilevel Bayesian liner model we defined in Section 5.2.1. Then, the TS policy chooses the action that yields the highest expected reward under the sampled parameter $\hat{\theta}_n$. Under TS, each arm will be chosen according to the probability that it maximizes the expected reward. As more patients' responses are observed and the posterior distribution $p_n(\theta \mid \mathcal{D}_n)$ shrinks to its true mean, the best arm will be chosen with higher and higher probability, thus achieving the exploitation objective, which guarantees the algorithm's long-term performance.

Starting with the multivariate normal prior $p_0(\theta)$ we defined in (5.4), our beliefs on the distribution of the unknown parameters will be updated as we receive more data throughout the trials. Denote by $\phi_{i,k} = (\varphi_{i,k}, \varphi^u_{i,k}, \varphi^\tau_{i,k})$ the set of feature vectors we receive for patient i's visit k. Given the prior distribution $p_0(\theta) \sim N(\mu_\theta, \Sigma_\theta)$ with prior mean μ_θ and covariance Σ_θ, and the model hyper-parmeters $\Sigma = (\sigma^2, \Sigma_u, \sigma_\tau)$, the joint posterior distribution for the unknown parameters $p_n(\theta \mid \mathcal{D}_n)$ can be obtained using the Bayes rule:

$$p_n(\theta \mid \mathcal{D}_n, \Sigma) \propto \prod_{i=1}^{i_n} \prod_{k=1}^{k_n} p(y_{i,k} \mid \phi_{i,k}, \omega, \eta_i, \tau_k, \sigma^2) p(\omega, \eta_i, \tau_k \mid \Sigma). \tag{5.5}$$

Denote the dimension of context feature $\varphi_{i,k}$, $\varphi^u_{i,k}$, $\varphi^u_{i,k}$ and $\varphi^\tau_{i,k}$ by q, q_u and q_τ, respectively. To provide the closed-form solutions for the posterior updates, we first combine Equation (5.2) and (5.3) and re-express $y_{i,k}$ as

$$y_{i,k} = \omega\varphi_{i,k} + (\gamma_1 z_{i1} + \gamma_2 z_{i2} + \cdots + \gamma_d z_{id} + \eta_i)\varphi^u_{i,k} + \ \tau_k \varphi^\tau_{i,k} + \varepsilon_{i,k}. \tag{5.6}$$

Then, we can re-write the multilevel equation (5.6) as

$$y_{i,k} = \tilde{\omega}\tilde{\varphi}_{i,k} + \ \eta_i \varphi^u_{i,k} + \tau_k \varphi^\tau_{i,k} + \varepsilon_{i,k}. \tag{5.7}$$

where $\tilde{\omega} = [\omega, \gamma_1, \gamma_2, \cdots, \gamma_d]$ consists of the population-level coefficients across multi-level models, and $\tilde{\varphi}_{i,k} = [\varphi_{i,k}; z_{i1}\varphi^u_{i,k}; z_{i2}\varphi^u_{i,k}; \ldots; z_{id}\varphi^u_{i,k}]'$ is the corresponding feature vectors.

Since in the future model updating, we sort the clinical data chronologically, based on the date of each patient visit. For the ease of exposition, we thus suppress the patient index i

and LOT index k in the above equation, and use index n to present the response and context features received at each trial in the rest of the manuscript. Let $y = (y_1; y_2; \ldots; y_n)'$ be a column vector for all observed survival outcomes till trial n. For each trial i, we construct an indicating column vector $d_n^u = (d_{n,1}^u, d_{n,2}^u, \ldots, d_{n,i_n}^u)$ with zeros and ones, where $d_{n,i}^u$ is a unit vector 1_{q_u} if trial n is related to the visit for the ith patient, and i_n the total number of unique patients observed till trial n. Similarly, we have $d_n^\tau = (d_{n,1}^\tau, d_{n,2}^\tau, \ldots, d_{n,k_n}^\tau)$, and $d_{n,k}^\tau$ is a unit vector 1_{q_τ} if trial n is corresponding to a patient visit for the kth LOT, and k_n the largest LOT among all the patient visits till trial n.

Given the above definitions, we can construct a feature matrix $X = (x_1; x_2; \ldots; x_n)'$, with each row $x_n = (\tilde{\varphi}_n, d_n^u \varphi_n^u, d_n^\tau \varphi_n^\tau)$. To illustrate the notions, suppose at trial $n = 3$, we observe two unique patients – the first patient visited in trial $n = 1$ on January 1st, and in $n = 3$ on March 1st, for total two LOTs, and the second patient visited for his first LOT on February 1st, at trial $n = 2$. Then, design matrix X is given by

$$X = \left[\begin{array}{c|cc|cc} \tilde{\varphi}_1 & \varphi_1^u & 0 & \varphi_1^\tau & 0 \\ \tilde{\varphi}_2 & 0 & \varphi_2^u & \varphi_2^\tau & 0 \\ \tilde{\varphi}_3 & \varphi_3^u & 0 & 0 & \varphi_3^\tau \end{array}\right] \begin{array}{l} \\ \\ \\ \end{array} \quad \begin{array}{cc} & \\ \underbrace{\qquad\qquad} & \underbrace{\qquad\qquad} \\ u_i & \tau_k \end{array} \tag{5.8}$$

Given the design matrix X, the posterior mean and variance of θ_n can be calculated using the following closed-form solution:

$$\mu_\theta^n = \Sigma_\theta^n \left(\Sigma_\theta^{-1}\mu_\theta + (\sigma^2)^{-1} X'y\right) \tag{5.9}$$

$$\Sigma_\theta^n = \left(\Sigma_\theta^{-1} + (\sigma^2)^{-1} X'X\right)^{-1}. \tag{5.10}$$

Since after each patient visit, the posterior distribution can be updated to benefit future decision making, yet a complete re-computation using all the historical data collected up to this point is clumsy, and computational inefficient. Hence, we rewrite equation (5.10) using the Sherman-Morrison formula (Sherman and Morrison 1950) to obtain a recursive updating formula for Σ_θ^n that does not require matrix inversion, and substitute it into equation (5.9)

to obtain a new set of updating formulas. Denoting the posterior mean and variance for θ at trial $n-1$ is μ_θ^{n-1} and Σ_θ^{n-1}, we have:

$$\mu_\theta^n = \mu_\theta^{n-1} + \frac{y_n - x_n{}'\mu_\theta^{n-1}}{\sigma^2 + x_n{}'\Sigma_\theta^{n-1}x_n}\Sigma_\theta^n x_n \tag{5.11}$$

$$\Sigma_\theta^n = \Sigma_\theta^{n-1} - \frac{\Sigma_\theta^{n-1}x_n x_n{}'\Sigma_\theta^{n-1}}{\sigma^2 + x_n{}'\Sigma_\theta^{n-1}x_n}. \tag{5.12}$$

5.2.3 Parametric Empirical Bayes for Updating Hyper-parameters

So far, our posterior updating was based on the hyper-parameter set $\Sigma = (\sigma^2, \Sigma_u, \sigma_\tau)$. Now we turn to discuss our strategy for tuning the hyper-parameters Σ. In this paper, we use the empirical Bayes approach to update the distribution of the hyper-parameters. That is, the objective of the empirical Bayes is to estimate the hyper-parameter $\hat{\Sigma} = (\hat{\sigma}^2, \hat{\Sigma}_u, \hat{\Sigma}_\tau)$ from observed data Y at every time point n that belongs to the updating schema $\mathcal{H}$[3], until trial N. Let

$$\mathcal{L}(\Sigma) = \log p(\mathcal{D} \mid \Sigma) = \log \int p(\mathcal{D} \mid \theta)p(\theta \mid \Sigma)d\theta \tag{5.13}$$

be the log-marginal likelihood of the hyper-parameter Σ given history $\mathcal{H}$.

We use the Expectation-Maximization (EM) algorithm (Dempster et al. 1977) to solve the above optimization problem. Instead of optimizing the original objective function $\mathcal{L}(\Sigma)$, the EM algorithm iteratively maximizes a lower-bound $\mathcal{F}_{EM}$ fo $\mathcal{L}(\Sigma)$:

$$\mathcal{F}_{EM}(q_\theta, \Sigma) = \mathbb{E}_{q_\theta}\left[\log p(Y \mid \theta) + \log p(\theta \mid \Sigma)\right] \tag{5.14}$$

where q_θ is the probability distribution over the parameter θ. At the l-th iteration, we have E-step and M-step proceeds as follows. At E-Step, we solve the posterior distribution given the hyper-parameters $\Sigma^{(l)}$:

$$\begin{aligned} q_\theta^{(l)} &= \text{argmax}_{q_\theta}\mathcal{F}_{EM}(q_\theta, \Sigma^{(l)}) \\ &= p(\theta \mid D, \Sigma^{(l)}), \end{aligned}$$

[3]We use the annealing optimization schema for the parameters updating. The details of the updating schema is presented in Appendix D.2.

where $p(\theta \mid D, \Sigma^{(l)})$ is given by the posterior mean and co-variance given by Equations (5.9), and (5.10), respectively. At the M-step, we can solve for the hyper-parameters $\Sigma^{(l+1)}$ using the derived closed-form solutions. Define

$$\mu_{p(u)} = [\mu'_{p(u_1)}, \mu'_{p(u_2)}, \cdots \mu'_{p(u_m)}]', \quad \mu_{p(\tau)} = [\mu'_{p(\tau_1)}, \mu'_{p(\tau_2)}, \cdots \mu'_{p(\tau_k)}]',$$

and

$$\Sigma_{p(u)} = diag\left(\Sigma_{p(u_1)}, \Sigma_{p(u_2)}, \cdots, \Sigma_{p(u_m)}\right), \quad \Sigma_{p(\tau)} = diag\left(\Sigma_{p(\tau_1)}, \Sigma_{p(\tau_2)}, \cdots, \Sigma_{p(\tau_k)}\right).$$

Then the updated hyper-parameters $\Sigma^{(l+1)} = \left\{\hat{\Sigma}_\tau^{(l+1)}, \hat{\Sigma}_\eta^{(l+1)}, (\hat{\sigma}^2)^{(l+1)}\right\}$ at the l-th iteration can be computed by the following equations:

$$\hat{\Sigma}_\tau^{(l+1)} = \frac{1}{k}(\mu'_{p(\tau)}\mu_{p(\tau)} + A_k \Sigma_{p(\tau)} A'_k), \tag{5.15}$$

$$\hat{\Sigma}_\eta^{(l+1)} = \frac{1}{m}(\mu'_{p(u)}\mu_{p(u)} + A_m \Sigma_{p(u)} A'_m), \tag{5.16}$$

$$(\hat{\sigma}^2)^{(l+1)} = \|Y - X\theta\|^2 + \mathrm{tr}(X\Sigma_\theta X'), \tag{5.17}$$

where A_m and A_k are sum vectors with dimension m and $k \times 1$. We repeat the above iterations until the marginal likelihood $\mathcal{L}(\hat{\Sigma})$ converges. We present the EM algorithm for empirical Bayes estimates of the hyper-parameters in Algorithm 3.

Algorithm 2 presents the details of our personalized treatment recommendation strategy, illustrating how it combines the multilevel Bayesian linear model, with the TS treatment selection rule to generate personalized treatment recommendation.

5.3 Causal Offline Evaluation

Unlike the standard supervised machine learning model, evaluating a bandit policy's performance is especially challenging due to the "partial-label" nature of bandit data (Langford et al. 2008, Li et al. 2011). To effectively evaluate the performance of our proposed policy, we need to know the survival outcomes for each patient visit under all possible treatment scenarios. However, we can only observe the factual survival outcome (i.e., PFS) associated

Algorithm 1: E-M algorithm for empirical Bayes estimates of hyper-parameters

Data Inputs: X, Y

$\hat{\mathbf{\Sigma}}^{(0)} \leftarrow \mathbf{\Sigma}_0$; $ll \leftarrow 0$; $l \leftarrow 0$;

repeat

  E-step: Given $\hat{\Sigma}^{(l)}$, compute conditional posterior distribution $q_\theta^{(l)} = p(\theta \mid \hat{\Sigma}^{(l)})$ from Equations (5.11) and (5.12).;

  M-step: Calculate $\hat{\Sigma}^{(l+1)} = \left(\hat{\Sigma}_\eta^{(l+1)}, \hat{\Sigma}_\tau^{(l+1)}, (\hat{\sigma}^2)^{(l+1)}\right)$ via Equations (5.15), (5.16), and (5.17) ;

  Compute log-likelihood $ll' = \mathcal{L}(\hat{\Sigma}^{(l+1)})$;

  $ll \leftarrow ll'$; $l \leftarrow l + 1$;

until *ll converges*;

return $\hat{\Sigma} = \left(\hat{\Sigma}_u, \hat{\Sigma}_\tau, \hat{\sigma}^2\right)$

with the treatment combination assigned to that patient in the historical data. The ideal way to unbiasedly evaluate our proposed policy is to run a clinical trial on real patients and collect their live responses. However, in practice, this approach is not feasible in our rare disease medical setting, due not only to the challenges of limited sample size and variations that guarantee a significant comparison, but also to the severe ethical issues of exposing patients to an untested policy.

The offline policy evaluation paradigm becomes valuable when we have only historical data available (Strehl et al. 2010, Li et al. 2011). Instead of running the target policy directly on the environment and getting online data, this approach relies on the historical dataset obtained from a different policy to approximate the expected reward using the target policy. Among the literature, the simplest and the most common practice among the bandit community is to estimate a reward function directly from the data, and use the reward estimated function to predict the missing labels used in policy evaluation (Beygelzimer and Langford 2009, Dudík et al. 2011).

Although the regression approach is valuable in proving ways for offline evaluation using

Algorithm 2: Thompson Sampling (TS) for Multilevel Bayesian Contextual Bandit

Inputs: $\mu_\theta, \Sigma_\theta$

$P(\theta) \leftarrow \mathcal{N}(\mu_\theta, \Sigma_\theta)$;

for $n = 1$ **to** N **do**

  Receive patient index i and LOT index k;

  Collect state variables $\mathcal{D}_n$;

  Draw $\hat{\theta}_n \sim p_{n-1}(\theta)$;

  Select arm $a_n = \text{argmax}_a \mathbb{E}[y \mid \phi_n, \hat{\theta}_n, a_n]$;

  Collect reward y_{n+1};

  Update $\mathcal{D}_{n+1} = \mathcal{D}_n \cup (\varphi_n,\, y_n, a_n)$;

  if $n \in \mathcal{H}$;

    Update hyper-parameters $\hat{\Sigma}$ via parametric empirical Bayes described in Section 5.2.3;

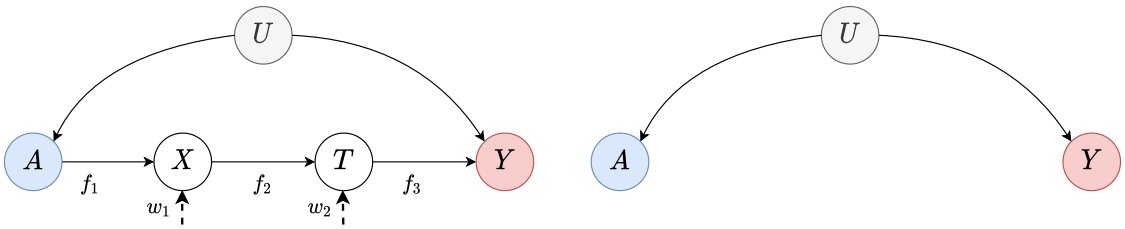
  Update posterior $p_{n+1}(\theta \mid \mathcal{D}_{n+1})$ via Equations (5.9) and (5.10)

end

historical datasets, a significant drawback of this approach is that it implicitly assumes that the probabilistic conditional distribution $P(Y \mid A = a)$ that *filters out* the sub-population of $A = a$ is the same as the causal distribution $P(Y \mid do(A = a))$ of *altering* the action $A = a$ on the counterfactual population (Bareinboim et al. 2015). [4]

We presented two causal graphs in Figure 5.3 to show how these two distributions differ under different causal relationships. In Figure 2(a), we have causal pathways from A to Y through f_1 to f_3 ($A \rightarrow X \rightarrow T \rightarrow Y$). Besides, we have unobserved confounders U that influence both A and Y($A \leftarrow U \rightarrow Y$). In our context, these confounders could include patients' preferences, lifestyles that may influence both the treatment selection and the survival outcome observed in the data. In this scenario, the estimated conditional distribution

[4]The *do* operator, introduced by Pearl's causal inference framework Pearl (1995), has been widely used to represent causal conditioning or to produce a new ensemble by the interventions.



(a) Causal graph with confounded causality (b) Causal graph with no causality

Figure 5.3: Example causal graphs. [a]

[a] An arrow indicates a causal relationship from the precedent to the subsequent factor. In Figure (a), the conditional distribution of $P(Y|A = a)$ reflects both the causal effect $P(Y|do(A = a))$ from the pathways f_1 through f_2 and the indirect relationship caused by U. In Figure (b), there is no causal path between A and Y, but we observe conditional dependence due to their common cause U

of $P(Y|A = a)$ reflects both the causal effect $P(Y \mid do(A = a))$ and the indirect relationship due to common cause U. Thus, we would not be able to separate the causal relationship $P(Y \mid do(A = a))$ from the estimated probabilistic conditional distribution $P(Y \mid A = a)$ without having knowledge about the confounders U. In Figure 2(b), we present a scenario in which there is no causal path from A to Y. Under the presence of the unobserved common cause U, we would see *correlation* between A and Y from the conditional distribution $P(Y \mid A = a)$, but changing the choice of A would not have any *causal impact* on Y, indicating $P(Y \mid do(A = a_1)) = P(Y \mid do(A = a_2))$ under different counterfactual choices a_1 and a_2.

As can be seen, the assumption $P(Y \mid A = a) = P(Y \mid do(A = a))$ fails whenever there exist un-controlled, hidden variables U that are correlated with both the treatment selection A and target outcome Y. While the external validity of the regression method heavily relies on this assumption, it does not hold true for most of the real-world applications. In such a case, the bandit algorithms actually maximize the rewards on the observed distribution $P(Y \mid A = a)$ estimated from the observed data, which does guarantee its performance on the counterfactual outcomes when changing the action that is different from the historical

data.

The replay offline evaluation policy proposed by Li et al. (2011), unlike the previously mentioned approaches, relies on log data directly without building a simulator. While this method provides a reliable way to unbiasedly estimate the policy without suffering the modeling and distribution biases it requires infinite, or at least large-scale logging data stream to obtain unbiased estimation, thus is also not feasible in our setting.

Facing these challenges, we propose a generative causal modeling approach for offline policy evaluation. Instead of building a naive regression model on high-dimensional raw features, we developed a causal graph from the treatment and covariates to predict survival outcomes based on medical literature, clinical evidence, and domain expert knowledge. The established causal chain allows us to estimate the corresponding generative model from the historical data, providing us a way to measure the effect of the treatment actions through the causal pathways, making our estimates hold true even in the presence of unmeasured confounders.

In Figure 5.4, we illustrate the building blocks of our policy evaluation framework. In the following sections, we first describe our causal generative model in detail, providing theory support for the modal calibration in Section 5.3.1. In Section 5.5, we present a series of validity checks to evaluate its validity as the causal counterfactual generator. Finally, in Section 5.4, we present the empirical performance of our proposed algorithm using causal offline evaluation.

5.3.1 A Generative Causal Hidden Markov Model

Before going into the modeling details, it is necessary for us to discuss some unique features of patients with multiple myeloma that motivate our subsequent model selection and calibration. Multiple myeloma is a heterogeneous disease with various disease courses, responses to therapy, and survival outcomes ranging from a few weeks to more than fifteen years. With the same treatment, some patients can have outstanding survival outcomes with stable disease status, whereas other patients may have a poor prognosis with aggressive

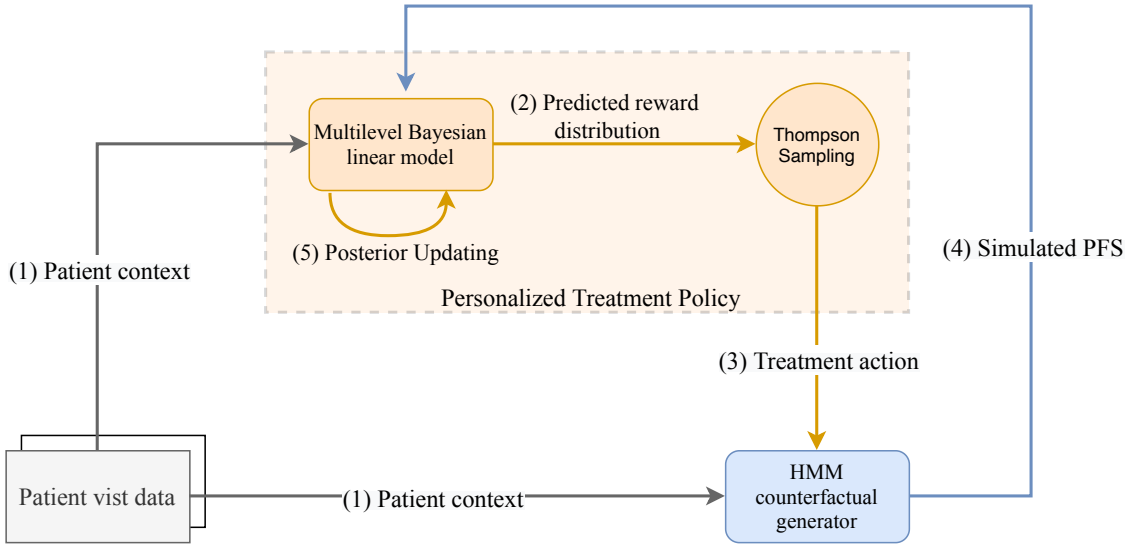


Figure 5.4: Causal Offline Evaluation Framework.

disease progression. Moreover, emerging clinical evidence suggests that some novel drugs could overcome the unfavorable prognosis of genetic abnormalities, leading to a change in risk level for relapse or disease progression (Munshi et al. 2011).

Thus, an ideal generative causal model that can capture patients' heterogeneity and the stochastic nature of treatment response should have the following properties: it can 1) model PFS as a result of previously given treatment therapies, 2) control for the relevant observed prognostic factors and unobserved risk levels that influence patients' observed survival outcomes, and 3) incorporates the dynamic and evolving nature of patients' risk level throughout the treatment journey. To operationalize these aspects, we propose using the Hidden Markov Model (HMM) to model patients' disease dynamics.

In HMMs, at each time step t, the observed outcome y_t is generated conditioned on the hidden state s_t, and the hidden state will then transit to a new state s_{t+1} (Figure 5.5). HMMs have been generally accepted as a statistical graphical model for casual inference (Spirtes 2010, Kelly et al. 2012, Spirtes et al. 2013), due to their two key assumptions. First, the observation y_t is conditionally independent of all other variables given the hidden state

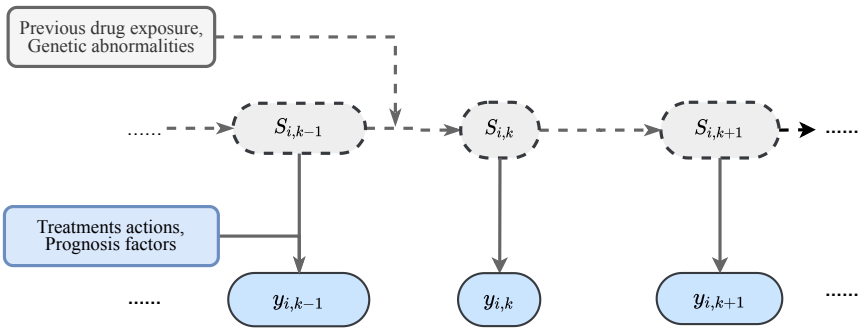


Figure 5.5: The Hidden Markov Model.

s_t's. This conditional independence property provides a structural basis for causal reasoning (Dawid 1979, Janzing and Schölkopf 2010). Second, the transitions of the hidden states are correlated through a first-order Markovian process. While an important defect of the causal model is to assume particular causal relationships between variables to be known a priori, the HMM does not suffer from this defect as it relies on the temporal relationship of the events to derive causality rather than reasoning (Wu 2010).

HMMs have been widely used as a structural framework to infer causality in various areas such as marketing (Montgomery et al. 2004, Netzer et al. 2008), online healthcare (Yan and Tan 2014), social media content creation (Huang et al. 2015), and user engagement (Zhang et al. 2019b). Due to their flexibility in accounting for the temporal causal structure, they are also used in medical literature for modeling chronic disease, which involves dynamic disease progression, stochastic in treatment outcomes, and unobserved health status changes over time (Dai et al. 2012, Denton 2018).

The Model.

In our HMM, we identify a patient's risk level as the unobserved latent state $s \in \{1, 2, \cdots, S\}$, with 1 being the lowest and S being the highest. All other things being equal, each risk level corresponds to a unique treatment response pattern. At any given point, a patient is associated with a latent risk level, which could change from time to time. Such a transition

can be attributed to natural disease progression, genetic abnormalities identified by the FISH test, or exposure to certain drugs given in previous treatments (Munshi et al. 2011). A patient's series of hidden states $S(i) = S_{i,1}, S_{i,2}, \cdots, S_{i,k_i}$ generates a sequence of observed outcomes $y(i) = y_{i,1}, y_{i,2}, \cdots, y_{i,k_i}$, which is defined by the patient's PFS along with their k_i LOT (Figure 5.5).

State transition model. We assume the state transition process is probabilistic and is governed by individuals' underlying transition propensity. In our HMM, a transition is allowed between any two states, as the patients' risk level can change quite dramatically. Conditioning on being at state s, a transition to a higher state s_H occurs if patients' transition propensity is higher than the threshold $\bar{\kappa}_{s \to s_H}$; similarly, a transition to a lower state s_L occurs if the transition propensity is lower than the threshold $\bar{\kappa}_{s \to s_L}$. At each time point, the transition probabilities from state s to any other state $j \in \{1, 2, \cdots, S\}$ sum to 1.

For patient i at his kth LOT, we define the transition propensity at the state s as following:

$$prop_{s,i,k} = r_s T_{i,k} + e_{s,i,k}. \tag{5.18}$$

Here, T_{ik} is a vector of influential factors that affect patients' hidden risk level transition; r_s is the state-specific coefficients, measuring the effect of T_{ik} on the underlying transition propensity. We consider two sets of factors that affect a patient's hidden state: 1) the detected genetic abnormalities, and 2) patients' previous exposure to treatment agents. The general consensus has been that the results of patients' FISH test, a genetic test that identifies cytogenetic abnormalities such as chromosome deletions or mutations, play an important role in risk stratification (Kapoor et al. 2009). It is documented that the detection of some genetic features, such as translocation of chromosomes 14 and 16, gain of 1q, or deletion of 17p may suggest patients' high risk for a poor prognosis (Rothman et al. 2012). Hence, in $T_{i,k}$, we include a set of dummy variables indicating the abnormalities detected from each genetic feature.

State dependent outcome model. The state-dependent outcome of interest is patients' PFS, which is approximated by the time duration between the start of two consecutive

LOTs. Following the previous medical literature (Zhang et al. 2012), we model the log transformation of $y_{i,k}$, as a linear function of the explanatory variables:

$$\log(y_{i,k}|S_{i,k} = s) = \lambda_s O_{i,k} + \varepsilon_{i,k}. \tag{5.19}$$

In the above equation, the vector $O_{i,k}$ includes the prognostic factors that directly influence patients' progression outcomes, λ_s is the state-dependent parameters, and $\varepsilon_{i,k}$ is the error term which follows the normal distribution. In this paper, we include two groups of factors can directly affect a patient's treatment outcome in vector $O_{i,k}$: 1) treatment choices, and 2) patient prognosis factors. The treatment a patient is given for each LOT can be a single drug or administered in the form of treatment combinations involving one or more individual drugs. Thus, we include a set of dummy variables indicating the choice of a specific treatment combination to consider the nonlinear interaction effects among drugs within a treatment combination, as well recognized by the previous literature (Thorlund and Mills 2012).

Factors that affect a patient's survival outcome are called prognosis factors. Among all possible factors for patients with MM proposed by previous studies, some are directly related to the disease progression, such as how advanced the disease is in its spread in the body, or which organs are involved (Kyle and Rajkumar 2009). Other factors depend on the individual's demographics, including gender, age at diagnosis, or disease burden, which affects the individual's capacity to tolerate intensive treatment. A detailed list of included laboratory tests with explanations is available in Appendix **??**. [5]

Model estimation.

Our HMM parameters consist of three components: 1) the initial state distribution π, 2) the parameters for state transition probability distribution $[\kappa, r]$, and 3) the parameters for the observed outcome probability distribution λ. Consider a patient i with an observed outcome

[5]Given the high-dimensionality of our prognosis feature set, we adopt the LASSO method to select the features that have the highest power in predicting patients' treatment outcomes. Our LASSO regression results suggest eighteen significant features, including the degree of bone lesions detected by MRI tests and ten other features identified by blood or urine tests.

sequence $y(i) = y_{i,1}, y_{i,2}, \cdots, y_{i,k_i}$. Given the parameter set $\upsilon = [\kappa, r, \lambda]$, the conditional probability of observing this sequence, given a fixed set of states $S(i) = S_{i,1}, S_{i,2}, \cdots, S_{i,k_i}$ is

$$P(y(i)|\upsilon, S(i)) = \prod_{t=1}^{k_i} P(y_{i,k}|S_{i,k}) = \eta_{i,1}(y_{i,1}|S_{i,1}) \cdot \eta_{i2}(y_{i,2}|S_{i,2}) \cdots \eta_{i,k}(y_{i,k}|S_{i,k_i}), \tag{5.20}$$

where the $\eta_{ik}(y_{ik}|S_{ik})$ is the probability of observing the survival outcome given the state according to Equation (5.19). The likelihood of observing a state sequence $S(i)$ is given by

$$P(S(i)|\upsilon) = \pi(i) \prod_{t=1}^{K_i} P(S_{i,k}|S_{i,k-1}) = \pi(i) q_{i,1}(S_{i,1}, S_{i,2}) \cdots q_{i,2}(S_{i,k}, S_{i,k+1}) \cdots q_{i,k-1}(S_{i,k_i-1}, S_{i,k_i}). \tag{5.21}$$

Here, $\pi(i)$ is the initial state distribution probability. $q_{i,k}(S_{i,k}, S_{i,k+1})$ is the state transition probability. To estimate the model, we start with using the latent class model to estimate the initial distribution for the latent risk level π; then, we use the Maximum Likelihood Estimation (MLE) to estimate the model parameters $\upsilon = [\kappa, r, \lambda]$ by maximizing the likelihood in (5.21).

One remaining parameter still to be determined is the number of states. While this parameter is taken as given when we derive the likelihood function for our full model, we select the one with the best model performance among different alternatives using the high-dimensional Bayesian Information Criteria (HD-BIC) proposed by Gao and Song (2010), the result of which suggests that the three-state HMM outperforms other specifications. Specifically, for each number of states, we calculate

$$\text{BIC} = c \ln(k) N - 2 \times \ln(L), \tag{5.22}$$

where L is the estimated likelihood of the model using MLE, k is the number of parameters, and N is the sample size, which is the number of patients in our data. c is a multiplicative factor with various choices. In empirical studies, using $c = 1$ is shown to have good empirical results (Gao and Song 2010).

First, we estimate the HMM with $s = 2, 3, 4$ separately. Then, we calculate the HD-BIC for each model. The model with a smaller HD-BIC will be selected. As can be seen from Table C.1, the model with the number of hidden states $s = 3$ has the best performance.

Table 5.2: Model Fit and Comparison

Number of States	**Log-likelihood**	**Variables**	**HD-BIC**
2	-2131.76	90	13263.13
3	-1689.21	138	13232.92
4	-1583.75	188	13640.38

5.3.2 Estimation Results

Initial state distribution probabilities π estimated using the latent class model are $[0.250, 0.419, 0.33]$. We show the estimation results for the state transition parameters in Table C.2. The parameters that influence patients' risk level transition can be classified into three groups: 1) patients' demographic information that is fixed over time periods, 2) patients' genetic characteristics indicated by their FISH test results, 3) previous exposure to various drugs.

We present the estimation results for the transition parameters in Table C.2. First, we notice that although the transition patterns are dissimilar across patients at different risk levels, the manifestations of some cytogenetic features do have a statistical significant influence on patients' underlying risk levels. More interestingly, our findings reveal that a patient's previous exposure to certain drugs will trigger a transition as well, which provides an empirical support for the hypothesis proposed by Munshi et al. (2011) mentioned in the previous section. The change in genetic features along a patient's disease course and the variations in the treatment strategies are well recorded in our dataset, making it necessary to continuously monitor patients' genetic profile and reassess their risk status to deliver better medical care. The estimation results provide evidence of the dynamic pattern of patients' unobserved risk levels.

The estimated parameters for state-dependent treatment outcomes are presented in Table C.3. In the result table, we categorize the factors that influence patients' PFS into four groups: 1) assigned treatment, 2) patients' demographics that are fixed over time, 3) patients' health condition, which is correlated to the disease burden, and 4) various laboratory test

Table 5.3: Estimation Results for Transition Parameters

Variable Group	Parameters	State 1 (High Risk)		State 2 (Medium Risk)		State 3 (Low Risk)	
		Variables that affect state transition					
Genetic Characteristics	del17p	3.6409	(2.5283)	−0.1599	(0.8600)	0.5280	(4.0923)
	del13q	−12.0843 ∗∗	(5.0011)	−0.4978	(0.7326)	0.5553	(2.2656)
	t11_14	1.7676	(3.5872)	−0.6640 ∗∗	(0.3705)	0.1755	(4.4622)
	del1p	−2.2867	(4.5256)	1.7592	(2.9013)	−5.1307∗	(3.1934)
	t14_16	4.1979	(2.8780)	−1.5924	(1.6593)	−6.4308	(8.4684)
	t4_14	−13.5757 ∗∗	(4.5318)	0.1563	(0.8824)	−3.5334∗	(2.3051)
Treatment Mechanism	Chemotherapy	8.6563 ∗∗	(2.9107)	0.3465 ∗∗	(0.1699)	−0.6849	(2.7491)
	Immunotherapy	0.9102	(2.5481)	−1.1867	(0.7043)	−0.0342	(2.5548)
	Proteosome_inhibitor	0.5505	(2.9495)	−0.2173	(0.6061)	0.7340	(2.6767)
	Stem_Cell_Transplant	0.2646	(3.0360)	4.1854∗∗∗	(1.1861)	0.6321	(3.6858)
	Steroid	0.3266	(2.6040)	1.5261 ∗∗	(0.5644)	−1.2160	(2.8015)
		Thresholds					
	State 1 (High Risk)			1.0063	(2.4509)	6.4272	(4.9086)
	State 2 (Medium Risk)	−1.3429 ∗∗	(0.5035)			3.2816	(4.0658)
	State 3 (Low Risk)			−1.0656	(1.7737)	0.7049	(2.2583)

results that indicate the current disease progression.

The variations in the coefficients of parameters across different states indicate a systematic difference in treatment response patterns when patients change to a different risk level. This finding suggests a high-level heterogeneity in treatment response across patients, and even for the same patient across lines of therapy, due to risk level transitions. We also identify several significant prognostic factors that may affect myeloma patients' survival outcomes besides the treatments, such as patients' sex and their overall health conditions as indicated by various laboratory results.

In sum, we can draw two main insights from the HMM estimation results. First, they provide evidence of the dynamic pattern of patients' unobserved risk levels. Second, the state-dependent outcome parameters suggest that the effectiveness of different treatment combinations and the influence of various prognosis factors do vary significantly among patients at different risk levels. This heterogeneity is usually not well considered in the conventional treatment strategy, making the efforts of developing an adaptive treatment strategy

Table 5.4: Estimation Results for Outcome Parameters

Variable Group	Parameters	State 1 (High Risk)		State 2 (Medium Risk)		State 3 (Low Risk)	
Treatments	Comb.01	−0.4475	(0.3117)	0.1253***	(0.1253)	0.1253	(0.1069)
	Comb.02	0.7867 **	(0.2421)	0.1950	(0.1950)	−0.2267	(0.2322)
	Comb.03	3.4978***	(0.6568)	0.2557	(0.2557)	−0.2767	(0.3425)
	Comb.04	3.8557***	(0.4985)	0.2152	(0.2152)	−0.6529	(0.3962)
	Comb.05	0.2617	(0.4579)	0.2171	(0.2171)	−0.8455	(0.6251)
	Comb.06	−0.1881	(0.4827)	0.2767	(0.2767)	−0.9135 **	(0.4518)
	Comb.07	3.5601***	(0.6061)	0.2024	(0.2024)	−0.0478	(0.5053)
	Comb.08	1.0148***	(0.2925)	0.2381	(0.2381)	−1.0363	(1.0403)
	Comb.09	−1.7351***	(0.2572)	0.0849***	(0.0849)	−0.3332	(0.3916)
	Comb.10	−1.6193***	(0.2427)	0.1183 **	(0.1183)	0.7422	(0.5987)
	Num of Agents	0.4033	(0.2415)	0.1304	(0.1304)	−2.0729 **	(0.7306)
	Num of Agents (Quadratic)	−0.7241***	(0.2096)	0.1111 **	(0.1111)	−3.0238***	(0.5235)
	Drug Cost	−1.2546***	(0.2153)	0.0757	(0.0757)	−0.3710	(0.3646)
Demographics	Age at Diagnosis	−0.0413	(0.0480)	0.0389	(0.0389)	0.0265	(0.3563)
	PatientSex	0.4345 **	(0.1891)	0.0654	(0.0654)	−0.0487	(0.1221)
	Lag PFS	0.0640	(0.0400)	0.0562 **	(0.0562)	−0.1558 **	(0.0616)
Disease Progression	Line	0.2591***	(0.0465)	0.0236	(0.0236)	0.1723	(0.1973)
	Line (Quadratic)	−0.7493***	(0.1244)	0.0640 **	(0.0640)	−0.0297	(0.1360)
Laboratory result	HCT	0.1604	(0.1110)	0.0401	(0.0401)	−0.0837	(0.1678)
	IGA	0.0308	(0.0416)	0.0503	(0.0503)	−0.1291	(0.1000)
	IGG	0.1122	(0.0899)	0.0326	(0.0326)	−0.1023	(0.1225)
	IGM	−0.0634	(0.0900)	0.0822 **	(0.0822)	−0.0163	(0.1301)
	ALB	−0.1604	(0.0820)	0.0264	(0.0264)	0.1171	(0.1624)
	CRE	0.0896*	(0.0482)	0.0292	(0.0292)	0.0347 **	(0.0112)
	Bone PLAZ	−0.1510	(0.0911)	0.0368	(0.0368)	−0.4604 **	(0.1628)
	MRI (Abnormal)	−0.1016	(0.1282)	0.0701	(0.0701)	0.3173	(0.2120)
	FLCR (High)	−0.0237	(0.1526)	0.0600	(0.0600)	−0.3908	(0.2717)
	LD (High)	−0.4607	(0.2948)	0.0732	(0.0732)	0.8770***	(0.1571)
	LD (Low)	0.1934	(0.2640)	0.0662	(0.0662)	0.6888 **	(0.2914)
	MPROTEIN (High)	−0.1835	(0.1256)	0.0678	(0.0678)	−0.4613	(0.3699)
	BETA2 (High)	0.4042 **	(0.1530)	0.1100	(0.1100)	−0.3156	(0.4424)
Variance	Sigma	0.4852	(0.3509)	0.3255	(0.3255)	0.7555***	(0.2119)

that takes into account individual-level real-time dynamics both urgent and indispensable.

5.4 Empirical Results

In this section, we demonstrate the effectiveness of our proposed personalized treatment recommendation strategy through our causal offline evaluation framework (Figure 5.4). We first compare the empirical performance of our proposed policy with several benchmark strategies and then conduct an ablation analysis that studies the performance of the proposed recommendation strategy by removing certain components, to understand the contribution of each component to the overall performance.

5.4.1 Experiment Setup

We sort the clinical data chronologically, based on the date of each patient visit. To reduce the noise in the data, we use only the raw variables whose support (the fraction of records having that feature) is at least 20%, and also eliminate the variable with zero variance. After these steps, each follow-up episode for a patient is represented by a raw feature vector of over 385 variables, which comprises: 1) patient ID and current LOT, 2) demographic information including gender and age at diagnosis; 3) degree of bone lesions; 4) genetic information indicated by FISH test results; and 4) the results of all other laboratory tests recorded in EMRs that measure patients' overall health conditions. It is worth noting that in the medical setting, many tests tend to be run together as a panel. Furthermore, the natural clinical correlations among some of the laboratory results lead to multi-collinearity among our high-dimension feature vectors. To alleviate this issue, we adopt principle component analysis (PCA) to transform the patients' features into a vector of 30 principal components. [6]

The simulation proceeds as follows. At each time period when a patient visits, we assign the treatment to the patient based on different treatment selection rules. Then, we observe a treatment outcome generated by the HMM counterfactual simulator, based on which we

[6]These 30 principal components can explain more than 90% of the total variance in the original variable space.

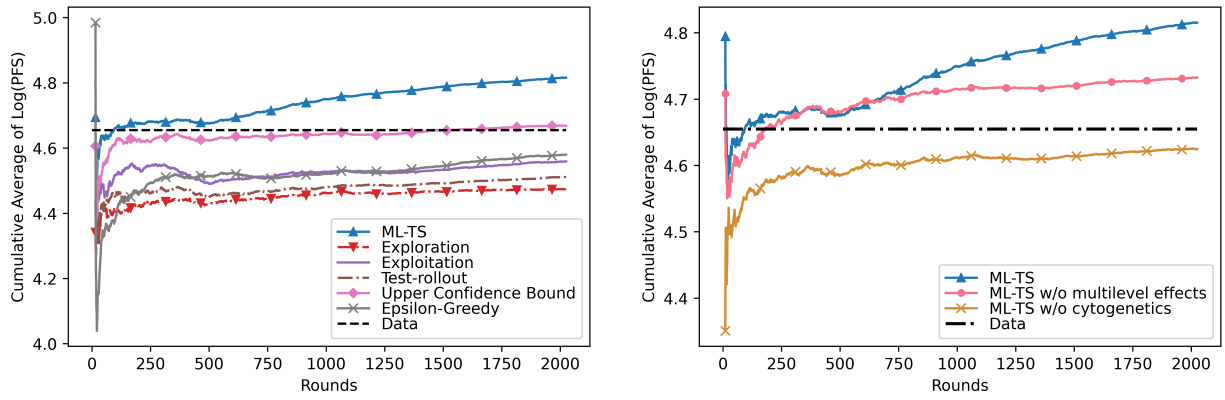
will update the belief model parameters. It is worth noting that the feature vector of each patient is as in the real world setting for each clinic visit, and the simulated length of the treatment course of each patient is thus the same as the actual LOTs presented in clinical dataset. We repeat this procedure for 500 times and evaluate the performance on patients' treatment outcomes averaged over all repeated experiments.

We compare the performance of our proposed policy with several benchmark treatment selection strategies. These strategies differ from the TS policy only in the way the treatments are selected in each trial; all the rest, including the belief updating, is identical to the policy we described in Algorithm 1. The first set of strategies concerns simple treatment selection heuristics, including (1) *pure-exploration* (which randomly selects the treatment for all patient visits); (2) *pure-exploitation* (which chooses the option that has the best expected treatment outcome based on the patient's context information); and (3) *test-rollout strategy*. When conducting this policy, we first perform the pure-exploration strategy for the first $N/2$ trials; then, we switch to the pure-exploitation and implement the best option for the remaining $N/2$ trials. This strategy has a sharp transition from exploration to exploitation, as opposed to our proposed model that balances the two goals when making every single decision.

The second set of strategies we consider are alternative MAB policies that are commonly used in the bandit literature. The first strategy is the *epsilon-greedy* policy, which is a randomized policy that mixes exploration and exploitation by setting a pre-determined portion of exploration ε. For any $\varepsilon \in [0, 1]$, the policy randomly allocates ε of patient visits to take the treatment selected by the random strategy, while it allocates $1 - \varepsilon$ patient visits to the seemingly best treatment at that time. The second MAB algorithm we consider is the upper confidence bounding (UCB) algorithm. Originating from the work of Auer et al. (2002), at each round, the UCB algorithm assigns each arm a UCB index, which is the sum of the expected reward and an upper bound on the probability that the above expected mean is an underestimate of the true mean. The arm that has the highest UCB value will be chosen.

5.4.2 Comparison with Benchmark Policies

In Figure 5.6(a), we plot out the evolution of the cumulative average reward (i.e., logarithm of PFS) by trials, which is defined by $\frac{1}{n}\sum_{n'=1}^{n}\log y_{n'}$. This allows us to examine how the average performance of these strategies varies in the process of the experiment. First, we notice that none of the simple heuristics has a satisfying performance in our application, while pure-exploration has the worst performance. The second insight comes from the above results in the learning path of our proposed multilevel linear Bayesian TS (ML-TS) algorithm. As seen in Figure 5.6, the performance of our proposed model (blue line with triangle dots) is continuously improving through its active learning, as indicated by increasing the average PFS over time. In contrast, other policies have a much slower learning rate, and some of them tend to converge to a suboptimal point. These results suggest the merit of our proposed policy in the clinical setting, where the cost of errors is exceptionally high.



(a) Benchmark Model Comparison Results

(b) Ablation Studies Results

Figure 5.6: Experiment Results

We also report the cumulative average reward at the end of the simulation trials in Table 5.5. For each policy under comparison, we present the logarithm of PFS averaged over 500 simulation runs reward. We also report the improvements of our proposed model over each

of the benchmark strategy, both in terms of the absolute PFS improvement (ΔPFS), and the percentage lift compared to each benchmark models (Δ%PFS).We further conducted bootstrap T-test to test whether the differences between our proposed model, and each of the benchmark models are statistically significant. As we can see, while our model does not outperform other benchmark model by a significant margin under the logarithm scale, the improvement is substantial if we consider it in normal scale, and the difference is indeed significant at the 99% confidence level. As can be seen from Table 5.5, for example, compared to the current clinical practice, our proposed algorithm is predicted to increase the PFS of each LOT by an average of 17 days.

Table 5.5: Empirical Performance Comparison†

		Model	Average of Log(PFS)	ΔPFS w.r.t. Benchmark Models	Δ%PFS w.r.t. Benchmark Models
Proposed model	Multilevel Bayesian linear Thompson Sampling (ML-TS)		4.8044		
Benchmark models	Simple heuristics	Pure-exploration	4.4724^{***}	34.4796	39.38%
		Pure-exploitation	4.5584^{***}	26.6156	27.89%
		Test-rollout	4.5106^{***}	31.0699	34.15%
	MAB strategies	Epsilon-greedy	4.5769^{***}	24.8338	25.55%
		UCB	4.6670^{***}	15.6681	14.73%
	Observed PFS in dataset		4.6548^{***}	16.9580	16.14%

†Note 1. ΔPFS significance level: $\cdot\ p < 0.1$; $*\ p < 0.05$; $**\ p < 0.01$; $***\ p < 0.001$

†Note 2. Significant levels are marked relatively to Multilevel Bayesian liner contextual bandit

Comparing our proposed model performance the simple heuristic strategies, we find that using TS as the treatment allocation rule improves the the model performance by around 30%. This result shows that the efficiency of balancing the exploration-exploitation trade-off translates into satisfactory model performance. Meanwhile, compared to other bandit strategies, our proposed policy can increase the PSF by an additional 15-25 days, and the increase is significant at the 99% confidence level. This may be because UCB is known to over-explore more than necessary. Thus, in our empirical healthcare setting, in which the sample size is limited, and each observation is very expensive, our proposed TS algorithm

offers much more value.

Table 5.6: Ablation Analysis Results†

Model		Average of Log(PFS)	ΔPFS w.r.t. Benchmark Models	Δ%PFS w.r.t. Benchmark Models
Proposed model	Multilevel Bayesian linear Thompson Sampling (ML-TS)	4.8044		
Ablation analysis	Remove multilevel effects	4.7118^{***}	10.7940	9.70%
	Remove cytogenetics feature	4.6242^{***}	20.1250	19.75%

†Note 1. ΔPFS significance level: $\cdot$ $p < 0.1$; $*$ $p < 0.05$; $**$ $p < 0.01$; $***$ $p < 0.001$

†Note 2. Significant levels are marked relatively to Multilevel Bayesian liner contextual bandit

We also note that without directly modeling heterogeneous across patients and/or different disease stages, a good balance between exploration and exploitation can increase the average logarithm of PFS by 7 days, compared to the patients' actual responses to the clinical treatment decisions observed in the data. We also find that incorporating the patient and LOT level fixed effects in the model increases the algorithm performance, in terms of the average logarithm of PFS, by 9.7%, endorsing the necessity of a richer, more granular-level model that takes into account heterogeneity across patients and at different disease stages.

For a better illustration of the learning curve across all the patients' visits over time, we show the evolution of the average logarithm of PFS of the ablation analysis in Figure 5.6(b). As we can see from the results, the full model has a larger learning rate compared to the model without multilevel effects. Such difference in performance could be due to the mis-specification of the benchmark model – as we will show later in Section 5.4.4, we observe significant heterogeneity in treatment response across patients, and also in different LOTs. Together with the discussion of performance of alternative treatment allocation rules we presented Section 5.4.2, we conclude that each of our model component, i.e., Thompson sampling used as the decision policy, cytogenetics features and response heterogeneity modeled with multilevel effects, is effective in the treatment recommendation procedure.

5.4.3 Ablation Analysis of Model Components

In this section, we conduct an ablation analysis to show that each of our model components (i.e. the multilevel Bayesian specification, patients' cytogenetic features) is effective in improving treatment decisions. First, we removed the multilevel effects (i.e., u_i and τ_k)and only kept the coefficients that are shared at the population level (i.e., ω). This model hints at the merit of considering heterogeneity of patient responses when designing a personalized treatment regimen. Second, we use the same multilevel linear Bandit specification in Equation (5.2), but we remove patients' cytogenetic features from the context vector. This benchmark model allows us to examine the value of including cytogenetic information in the personalized treatment recommendation.

The results of the ablation analysis are presented in Table 5.6. A superscripted asterisk denotes that a benchmark model performs significantly worse than the proposed model. An inferior performance of the model without cytogenetics feature compared to the current clinical practice implies that with medical advances, cytogenetic tests that identify specific cancer biomarkers (that are associated with response, resistance, or lock of response to certain treatment approaches) is still the key to ensure that patients are receiving the proper treatment (Rothman et al. 2012). Including patients' cytogenetic information in the personalized recommendation strategy significantly improves the model performance by 19.75%.

5.4.4 Subgroup Analysis

So far, we have shown the merits of our TS policy by comparing its performance with various algorithms at the population level. To get more insights from a more granular perspective, we now perform a set of posterior analyses to illustrate how our proposed personalized treatment regimen would benefit different subgroups of patients. First, we group patients into four segments based on their gender and age at diagnosis, to uncover the value of our personalized strategy for different demographic segments. We summarize their average achieved PFS after each visit using our treatment strategy and report the percentage lift

Table 5.7: Segments by Demographics

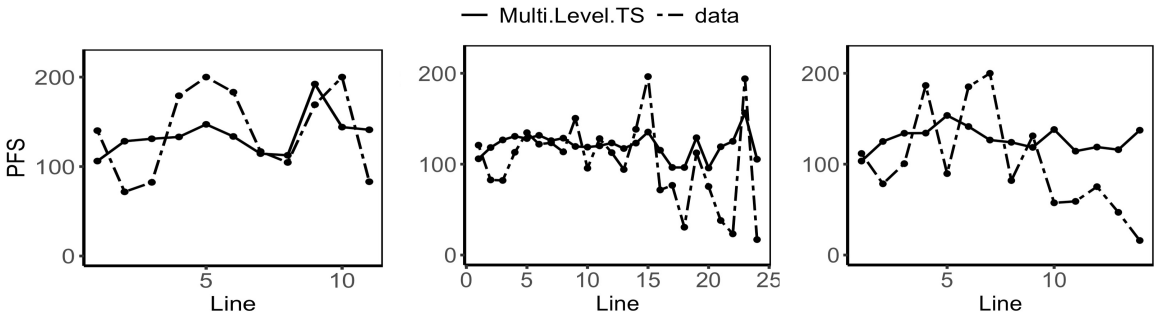
Segment	Age At Diagnosis	Gender	Clinical Practice	TS	%Lift
1	30 to 55	Female	99.03	117.68	19%
2	30 to 55	Male	105.01	126.70	21%
3	55 to 80	Female	105.13	143.91	37%
4	55 to 80	Male	107.41	139.96	30%

compared to that in the clinical practice in Table 5.7. Overall, we can achieve a higher PFS per LOT for patients from all demographic segments. However, the advantage is much more significant for elderly patients diagnosed over age 55, for both gender groups (37% for segment 3 and 30% for segment 4). This result suggests limitations for the current clinical practice in treating aging populations who are likely to have an increased risk of frailty and decreased tolerance for available therapeutics (Willan et al. 2016). Due to the limited treatment evidence for such sub-populations with high uncertainty in treatment responses, elderly patients can benefit the most from a data-driven analytics model like ours, which can achieve efficient and fast information collection while maintaining a good balance between exploration and exploitation. The results also suggest that if clinical trials are not feasible for all patients, elderly patients should be targeted first to see the merits of adaptive, personalized treatment recommendations in real medical practice.

The second attribute we consider is the earlier-mentioned static clinical risk stage, which stratifies patients at diagnosis based on agreed criteria. This metric helps physicians make treatment decisions in clinical practice, as physicians agree that patients categorized into various stages have systematic differences in prognosis and overall disease progression. To examine how the dynamics and treatment outcomes differ among different cancer stages, we plot the logarithm of the PFS against time separately for patients classified into each risk stage group in Figure 5.7. In each subplot, the dashed line represents the observed PFS for patients in different cancer stages over time, and the solid line represents the PFS obtained

by our proposed personalized treatment strategies.

Overall, the clinical practice did relatively well in treating patients at the early stages. However, as LOT increase, patients' survival outcomes start to drop, especially those categorized in a higher clinical stage. The causes of this sharp decrease in PFS are multiple. First, it could be attributed to disease progression. Second, the sequential and stochastic nature of chronic diseases makes patients highly heterogeneous in their disease conditions and individual treatment needs in their later stages. Thus, the decreased PFS may result from physicians' imprecise knowledge in tailoring treatments most appropriate to individual patients. In contrast, our proposed treatment strategies are highly personalized based on the patients' characteristics and past medical history. As seen in Figure 5.7, there is no significant drop in the PFS achieved by TS, even in later LOT. We conclude that our strategy allows more effective treatment and maximizes patients' survival outcomes by keeping the disease controlled at a relatively stable condition.



(a) Clinical Risk Stage I (b) Clinical Risk Stage II (c) Clinical Risk Stage III

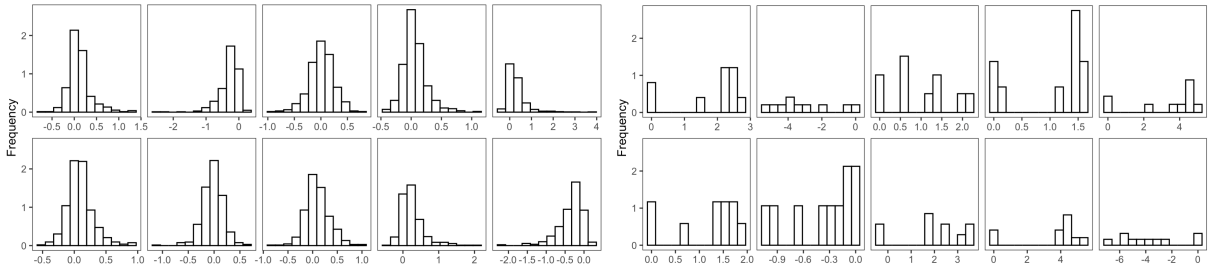
Figure 5.7: Posterior Analysis By Clinical Risk Stage

In Figure 5.7a and Figure 5.7b, we plot out the parameter distribution for the individual fixed effects (u_i's), and LOT fixed Effects (τ_k's), respectively, to examine the subgroup heterogeneity at a more granular level. Each u_i and τ_k is a 10-dimensional vector; thus, each subplot represents the distribution of the parameter in that specific dimension. From these

plots, we can see significant variation in both individual and LOT level coefficients. This suggests the treatment responses are heterogeneous across patients and at different disease stages, endorsing the necessity of a richer, more granular-level model that takes into account both levels of heterogeneity like ours.

5.5 Robustness Checks

In this section, we conduct additional analyses to evaluate the validity of our HMM as a counterfactual generator in our setting. We split the patients visits into a training and a testing set, and re-estimated the HMM using the training set data. In Section 5.5.1, we examine the performance of our HMM at the aggregate-level by comparing it with several benchmark strategies; in Section 5.5.2, we examine the robustness of HMM from a more granular perspective, checking its performance at the cohort, and at the individual-patient level; finally, in Section 5.5.3, we conduct a simulation study to test the robustness of HMM against an assumed omniscient HMM to validate the counterfactuals generated from the estimated HMM.



(a) Individual Fixed Effects

(b) LOT Fixed Effects

Figure 5.8: Fixed Effects Distribution

5.5.1 Aggregate-Level Robustness Checks

To start with, we examine the performance of the HMM simulator on our clinical data, aggregating across all the patient visits. In particular, we examine the HMM simulator's goodness-of-fit on the training dataset, and its predictive ability on the testing dataset by comparing its error rate with several baseline models that have been commonly used in the medical setting for survival analysis. These baseline models include the Cox regression, accelerated failure time model (AFT), log-normal, and log-logistic model. We also compare and include a non-parametric Random Forest (RF) machine learning strategy in survival analysis in high dimensional settings (Mogensen et al. 2012). We report the performance of each model in term of the MSE (Mean Squared Error), MAE (Mean Absolute Error), LMAE (Log Mean Absolute Error), and LRMSE (Log Root Mean Square Error) in Table 5.8. Additionally, we conducted a bootstrap T-test to check whether the performance difference is statistically significant between HMM and other benchmark models.[7] As we can see in Table 5.8, the HMM outperforms most of the benchmark models on both training and testing dataset, and the improvements are significant. Among all the baseline models, AFT is the most promising benchmark. Occasionally, it outperforms HMM in terms of MAE in the training dataset; however the good training error does not entail good performance on the testing set.

5.5.2 Segment-level Robustness Checks

Next, we evaluate the validity of the simulator at the patient cohort level.. Following the literature that uses HMM model to fit the disease progression markers of 430 human immunodeficiency virus (HIV) patients (Satten and Longini Jr 1996), we generated a simulated PFS from our estimated HMM using the training dataset, and we group these simulated outcomes across patients based on their line-of-therapy (LOT) at the visit. By comparing

[7]For example, the MSE for AFT model on the training set is 0.8314*** in Table 5.8, which suggests that the MSE of the AFT model is 0.8314, and is statistically different from the MSE of the HMM model (i.e., 0.764) at the 99.5% confidence level.

Table 5.8: HMM Performance against Benchmark Models†

Models	Training Set				Testing Set			
	MSE	MAE	LMAE	LRMSE	MSE	MAE	LMAE	LRMSE
Cox Regression	7.5187***	2.6306***	0.8469***	0.9018***	7.8624***	2.6424***	0.8917***	0.9379***
AFT	0.8314*	0.6838*	0.1545.	0.3887*	0.8933*	0.7986**	0.1910*	0.3884.
RF	4.4638***	1.6000***	0.4349***	0.5275***	4.8606***	1.6568***	0.4364***	0.5139***
Log-normal	0.8564*	0.7132*	0.1656*	0.3597.	0.877*	0.8022*	0.2151*	0.3794*
Log-logistic	0.8159*	0.7158*	0.1916**	0.3881**	0.8757*	0.7201**	0.2165*	0.3994*
HMM	0.7640	0.7010	0.1327	0.3394	0.8213	0.7408	0.2409	0.3781

†Note 1. $\cdot$ $p < 0.1$; $*$ $p < 0.05$; $**$ $p < 0.01$; $***$ $p < 0.001$

†Note 2. Significant levels are marked relatively to HMM

the simulated outcomes with observed outcomes in the data at the cohort level, we are able to directly evaluate the model's fitness and predictive ability for patients at different disease progression stages. As can be seen in Table 5.9, for LOTs with a relatively large cohort size (e.g., Line 1 to Line 6), the observed and the expected PFS are in agreement, indicating that our model is accurate in predicting patients' disease progression dynamically, at least at the population level. However, some lack of fit and accuracy after the 7th line could be due to the limited sample size for those cohorts.

Table 5.9: Observed and expected PFS in each successive LOT.†

Line	Training Set			Testing Set		
	Observed	Expected	Cohort Size	Observed	Expected	Cohort Size
1	4.7904	4.7800	120	4.7645	4.8156	31
2	4.4539	4.4258	139	4.3113	4.4114	35
3	4.5156	4.4901	154	4.3989	4.3991	38
4	4.6538	4.6309	158	5.1917	4.8891	39
5	4.8716	4.8564	158	5.1175	4.9712	40
6	4.9592	4.9287	108	4.7938	4.7792	27
7	4.8822	4.8929	74	5.0466	4.9168	17
8	4.6711	4.7134	51	4.8416	5.0635	12
9	4.9669	4.9840	37	5.2967	5.3053	10
10	4.5284	4.6167	24	4.8959	5.0540	6

†PFS values are presented in log-scale.

Besides assessing the model at the segment level, we take a more granular view and see

whether the estimated HMM can capture the dynamics of patients' disease progression at the individual level. To do so, we first recover patient's unobserved risk level across periods using the Viterbi decoding algorithm (Viterbi 1967), which is a dynamic programming algorithm for finding the optimal sequence of hidden states. Given an observation outcome sequence (i.e., the PFS) and an estimated HMM $\hat{\pi}$ and $\hat{\upsilon} = [\hat{\kappa}, \hat{r}, \hat{\lambda}]$, the algorithm returns the most likely hidden state paths $S(i) = S_{i,1}, S_{i,2}, \cdots, S_{i,s_i}$ through the HMM that assigns maximum likelihood to the observed sequence $y(i) = y_{i,1}, y_{i,2}, \cdots, y_{i,s_i}$. These recovered states allow us to examine whether these recovered hidden states are able to capture the dynamics in disease progression.

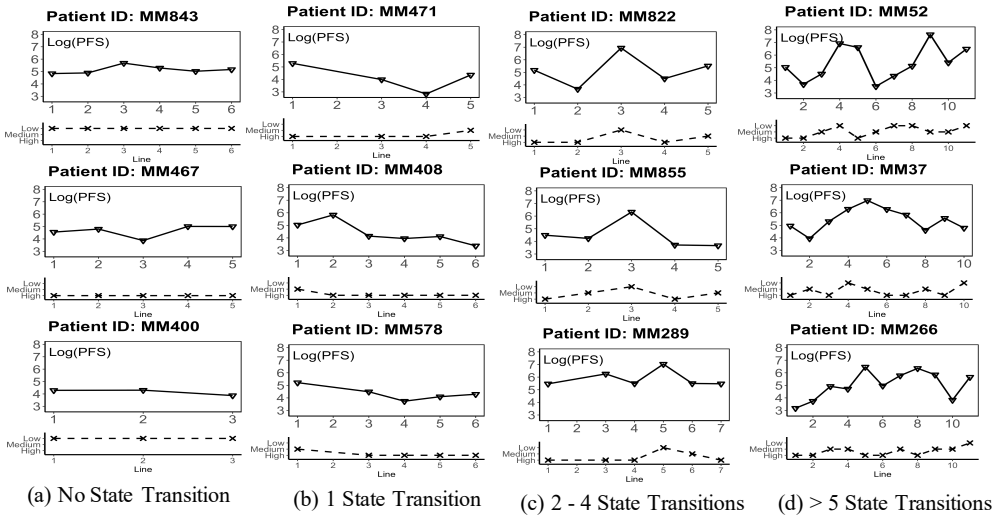


Figure 5.9: HMM Posterior Analysis of Patient Dynamics.

We group all the individuals based on the number of hidden state transitions in their treatment journey to illustrate the performance of HMM on individual patients with different disease progression patterns. Precisely, we classify patients into four groups: (1) no state transition; (2) one state transition; (3) two to four state transitions; and finally (4) greater than five state transitions. We randomly select three patients within each group and plot the observed PFS along with our predicted hidden risk levels in Figure 5.9. Each column in

Figure 5.9 plots the dynamics for the three randomly selected samples in each subgroup. In each sub-figure, the observed PFS over time is shown in the top panel, and the predicted risk level is presented underneath. As we can see, our predicted hidden states are concurrently fluctuating with the observed PFS throughout the patients' entire treatment course, indicating that our estimated HMM is able to capture the dynamics of patients' disease progression and risk level changes.

5.5.3 Robustness against an Omniscient HMM Simulator

Finally, we conducted a simulation study to test the robustness of HMM against an assumed truth generator to validate the counterfactuals generated by the estimated HMM, compared to the true HMM generator. In particular, we take one HMM as the ground truth and generate "factual" outcomes. The parameters of the true HMM could have been chosen arbitrarily, but to make the true HMM more realistic, we fit the HMM on the real clinical data and use the estimated parameters as the ground truth. We use the true HMM to generate "factual" data points, based on which we first re-fit our HMM model, and then use it as the counterfactual generator to repeat our entire MAB decision-making process. By comparing the empirical performance between the "ground-truth" HMM and the estimated HMM, we would like to examine to what extent our estimated HMM can approximate the underlying data-generating process.

We repeat the above simulation process for 30 rounds, and thus estimate 30 different HMMs from the simulated observations generated from the "ground-truth" HMM. For each estimated HMM, we use it as the counterfactual generator in our offline evaluation framework, and repeat the entire MAB decision-making process for 500 runs. In Table 5.10, we presented the average $\log(PFS)$ across all patients and across all LOTs achieved by the multilevel Bayesian linear TS policy from those simulated HMMs over 15,000 runs, and compare it with the results using the "ground-truth" HMM. Given our sample size is relatively large (15,000 samples) and the differential impact between the True HMM and the simulated HMMs is very close to zero and highly insignificant, based on the informativeness of statistical insignificance

Table 5.10: Simulation Results Comparison

Model			Average of Log(PFS) from True HMM	Average of Log(PFS) from Simulated HMMs	P-value
Proposed model	Multilevel Bayesian linear contextual bandit		4.8044	4.8132	0.5998
Benchmark Models	Simple heuristics	Pure-exploration	4.4724	4.4713	0.5896
		Pure-exploitation	4.5584	4.5623	0.5169
		Test-rollout	4.5106	4.5056	0.5351
	MAB strategies	Epsilon-greedy	4.5769	4.5680	0.5736
		UCB	4.6670	4.6644	0.5472
	Ablation analysis	Remove multilevel effects	4.7118	4.7277	0.5023
		Remove cytogenetics feature	4.6242	4.6263	0.5841

†P-value <0.05 indicates a statistically significant difference between the mean from True HMM and the mean from the Simulated HMMs.

emphasized by Abadie (2020), we claim that there is no statistically significant difference between the average results from the True HMM, and that from the simulated HMMs.

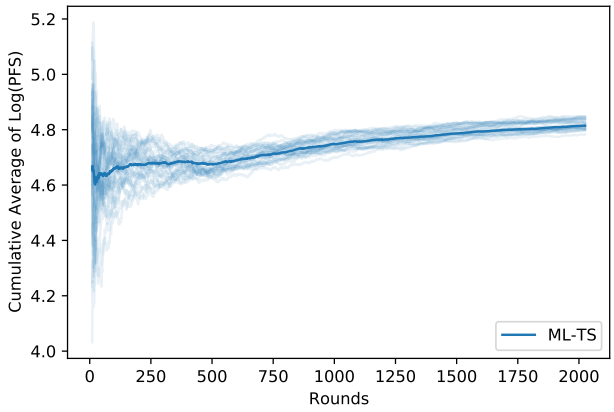


Figure 5.10: Simulated Experiment Results

We also plot out the evolution of the average logarithm of PFS by rounds for each of the simulation runs, along with the outcome obtained from the "ground-truth" HMM in Figure 5.10. As we can see from the results, the simulated HMMs can approximate the ground-truth quite well, indicating the performance of the estimated HMM is comparable with the

omniscient HMM simulator.

5.6 Conclusion

In this study, we consider the problem of designing personalized treatment recommendations for patients with multiple myeloma using a data-driven analytics method. We modeled the problem as a Bayesian contextual bandit that combines the multilevel Bayesian linear model with TS based treatment selection rule for generating personalized recommendation. This framework allows us to flexibly account for patient and line-of-therapy (LOT) level heterogeneity even in the absence of a large number of observations.

We evaluate the performance of proposed model on clinical data collected from 803 patients treated at SCCA. Our experiment results suggest that by balancing the exploration-exploitation trade-off using our proposed TS policy, we are able to achieve a more efficient treatment recommendation at the individual level by considering patient disease dynamics, cytogenetic variants, and unobserved heterogeneity in treatment responses. This advantage is significantly more remarkable for high-risk sub-populations, whose disease usually involves more complications that cannot be appropriately treated with conventional therapies. Moreover, through the counterfactual analysis from the structural econometric model, we validate the value of the TS algorithm in fast learning by comparing it to other widely-used MAB algorithms. This is critical in healthcare applications where the clinical data is usually sparse and limited and where ethical issues prevent arbitrary clinical trials without careful empirical evidence from observational data.

We have intended to make three key methodological contributions. The first contribution is in addressing a long-standing issue faced by physicians when treating chronic disease by framing it as a multi-armed bandit problem. There have been discussions in the healthcare community on whether there are better options than clinical trials when there are a large number of alternative treatment combinations available. One major reason is that the clinical trial only allows the head-to-head comparison between two different treatment alternatives. Thus, it is difficult to synthesize all efficacy evidence for direct comparisons of

all available treatment combinations (van Beurden-Tan et al. 2017). Given that there has been an increasing number of cancer treatment agents have emerged due to the scientific advances, it becomes imperative to find better ways to quickly identify the treatment that gives the best patient response from a large number of alternatives. This study addresses the above concern. The experiment results show the power of implementing a machine learning technique, TS, for practical use in developing personalized treatment strategies that can best balance the exploration-exploitation dilemma.

Second, we developed a multilevel Bayesian linear model to learn patients' heterogeneous response on treatment decisions even in the absence of a large number of observations; in addition, we obtained a closed-form online recursive updating formula on the posterior that is applicable to general cases, to facilitate computation and empirical application of our proposed method by future research. As such, we position our model development as a methodological contribution, both for the multi-armed bandit literature and to the healthcare analytics literature.

Third, we develop a HMM-based causal offline evaluation framework to estimate patients' counterfactual survival outcomes. Grounded on medical literature, clinical evidence, and domain expert knowledge, this framework allows us to measure the treatment actions through the established causal pathways, excusing us from the potential covariate shift issues that are present in the conventional regression methodologies.

Beyond the methodological contribution, our study also offers important implications to the healthcare community from the practical perspective. While the evidence from actual patients is needed to see the real improvement, it is worth noting not all forms of strategies can be undergone a clinical trial; and any strategy that will be studied in clinical trial needs to be somewhat proved effective before implementing it on real patients. In this study, we have conducted various robustness checks and simulation analysis on the HMM suggest that our HMM has good fit and predictive performance both at the population, and the individual patient level. Therefore, we believe the HMM model is a good approximate of the reality, and we would expect the actual improvement should be in a similar magnitude

of our empirical findings. Thus, our study contributes to the community by taking the very first step to provide empirical evidence for the effectiveness of MAB personalized treatment recommendation strategy, and convince physicians and medical researchers to move forward.

We believe there are several possible extensions of this study, including ones that would overcome the limitations of this study. One line of future investigation could focus on extending the TS under the presence of medical constraints. In the clinical practice, physicians may face a series of constrains when designing the treatment strategies for individual patients, such as their tolerance to the side effects, concerns on the medical cost, or the patients' own preference. Thus, further research in this regard would be valuable to both the theory and practice of a personalized treatment recommendation strategy tailored to the needs of each individual patient.

Second, we assume a stationary reward generating process, where the statistical properties remain constant over the time horizon of interest. However, the non-stationary property of the reward distribution has been observed in many real-world personalized healthcare settings, in which specific action chosen influences the reward distributions of each action in subsequent time horizon (Mintz et al. 2020). For example, in the context of using adherence–improving intervention to increase patients' physical activity, the specific action chosen influences the reward distributions of each action in subsequent time horizon. Thus, in the future research, we could extend our work to non-stationary bandit, to provide personalized healthcare solutions to more general real-word application scenarios.

Third, due to the limitation of data and constraints on medical ethics, we evaluate the empirical performance of our algorithm through the proposed causal offline evaluation approach. While the various robustness checks and simulation analysis suggests our HMM provides a good approximate of the reality, one caveat must be aware of is that our reported model performance is a model-based prediction, and the evidence from actual patients is needed to see the real improvements. In the future research, we plan to collaborate with SCCA on designing field experiment on patients with multiple myeloma to assess the value of our proposed model in real-world medical practice.

In closing, although we focus on the specific case of designing personalized treatment strategies for patients with multiple myeloma, we envision that this proposed framework could be easily applicable in general application domains that aim to provide personalized recommendations. It could enable researchers to pinpoint concrete and rigorous dynamic business solutions when they encounter the exploration/exploitation trade-off and when they do not have the ability to conduct clinical trials.

Chapter 6

CONCLUDING REMARKS

The rapid advancements of digital technologies have fostered transformation of various online and offline environments. In this dissertation, I look into three related topics that are related to such transformation: Business model innovation under digital transformation, marketing strategy in digital two-sided markets, and medical decision-making with healthcare information systems.

My first essay aims to understand investors' behavioral mechanisms in the novel Fintech initiative empowered by digital transformation – the crowdfunded Supply Chain Finance (SCF). As required by crowdfunded SCF platforms, the unique presence of loan guarantors in the financing process alters how fundraisers and investors interact, which gives rise to investor learning. By developing a Bayesian learning model, wherein we conceptualize individual perception of guarantor reliability as a subjective attitude underlying the perceived risk of a loan listing, we confirm the existence of investor learning: an individual's incidence decision and amount decision are both driven by her perception of guarantor reliability. In addition, we observe that this latent perception has different moderating effects on investor responses to listing attributes, such as interest rate and loan duration. Our empirical analysis generates useful implications for both platform managers,and for supply chain members.

My second essay studies the marketing strategies in the digital markets. In particular, I study a commonly used offline marketing strategy, the scarcity strategy in the context of reward-based platform, and its effectiveness in motivating individual engagement. Using a unique dynamic panel dataset from a leading reward-based crowdfunding platform, we develop a multi-level Bayesian model to examine the scarcity effect with controls for potential confounders as well as rewards' nested unobserved heterogeneities. Moreover, we uncover the

motivational mechanism that underlies backers' investment behavior that gives rise to the observed effect. We find that backers are preventionmotivated on the crowdfunding platform. Scarcer rewards tend to attract more backers, and this positive effect is amplified by a larger number of existing backers, as obtaining demand-induced scarce products provides these risk-averse backers with a sense of security. This study offers valuable insight into the website and reward scheme design, providing useful implications for both platform managers and crowdfunding fundraisers.

In my third essay, I looked into the digitization of healthcare systems, and see how we can utilize Electronic Medical Records to improve health-care decision making. In particular, we consider the problem of designing personalized treatment recommendations for patients with multiple myeloma using a data-driven analytics method. We formulate the treatment recommendation problem as a Bayesian contextual bandit, which sequentially selects treatments based on contextual information about patients and therapies, with the goal of maximizing overall survival outcomes. We developed a multilevel Bayesian linear Thompson sampling to learn patients' heterogeneous response on treatment decisions, which allows us to flexibly account for patient and line-of-therapy level heterogeneity even in the absence of a large number of observations.

Facing the difficulty of evaluating the performance of the policy with only observational data, we propose a causal offline evaluation approach to measure the effect of the treatment in the presence of unmeasured confounders. We evaluate the performance of our policy on clinical data collected from 803 patients treated at Seattle Cancer Care Alliance. Our policy achieved an 19.75% predicted improvement compared to the current clinical practice, and outperforms other benchmark strategies. Moreover, our policy achieves higher improvement for aging or high-risk patients with more complications by keeping the disease controlled at a relatively stable condition.

My dissertation intends to serve two purposes. First, I aim to study the important aspects related to digital transformation, providing insights in business practice across various online and offline markets. Second, I combine different empirical, and optimization methodologies

to transform raw data into business insights and enhance decision-making quality. Given the speed of advancements in digital technology, those methodologies and decision-making strategies can help researchers and to embrace the new possibilities ahead.

BIBLIOGRAPHY

Abadie A (2020) Statistical nonsignificance in empirical economics. *American Economic Review: Insights* 2(2):193–208.

Agrawal A, Catalini C, Goldfarb A (2014) Some simple economics of crowdfunding. innovation policy and the economy, 14 (1), 63-97.

Agrawal S, Goyal N (2012) Analysis of thompson sampling for the multi-armed bandit problem. *Conference on learning theory*, 39–1 (JMLR Workshop and Conference Proceedings).

Allon G, Babich V (2020) Crowdsourcing and crowdfunding in the manufacturing and services sectors. *Manufacturing & Service Operations Management* 22(1):102–112.

Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA (2012) Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. *Statistics in Medicine* 31(17):1887–1902.

Alousi A, Logan B, Couriel D, Krakow EF, Pidala J, Hemmer M, Last M, Arora M, Lachance S, Spellman S, Wang T, Moodie EEM (2017) Tools for the precision medicine era: How to develop highly personalized treatment recommendations from cohort and registry data using q-learning. *American Journal of Epidemiology* 186(2):160–172.

Auer P, Cesa-Bianchi N, Fischer P (2002) Finite-time analysis of the multiarmed bandit problem. *Machine learning* 47(2-3):235–256.

Balachander S, Liu Y, Stock A (2009) An empirical analysis of scarcity strategies in the automobile industry. *Management Science* 55(10):1623–1637, ISSN 0025-1909, URL `http://dx.doi.org/10.1287/mnsc.1090.1056`.

Bardhan I, Oh Jh, Zheng Z, Kirksey K (2014) Predictive analytics for readmission of patients with congestive heart failure. *Information Systems Research* 26(1):19–39.

Bareinboim E, Forney A, Pearl J (2015) Bandits with unobserved confounders: A causal approach. *Advances in Neural Information Processing Systems* 28:1342–1350.

Bastani H, Bayati M, Khosravi K (2017) Mostly exploration-free algorithms for contextual bandits.

Belavina E, Marinesi S, Tsoukalas G (2020) Rethinking crowdfunding platform design: mechanisms to deter misconduct and improve efficiency. *Management Science* 66(11):4980–4997.

Benna UG, Benna AU (2018) *Crowdfunding and Sustainable Urban Development in Emerging Economies*. Advances in E-Business Research (Hershey, PA: IGI Global).

Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. *Management Science* 62(5):1511–1531.

Bertsimas D, Vayanos P (2017) Data-driven learning in dynamic pricing using adaptive optimization. *Optimization Online* .

Betancourt M, Byrne S, Livingstone S, Girolami M (2017) The geometric foundations of hamiltonian monte carlo. *Bernoulli* 23(4A):2257–2298, ISSN 1350-7265.

Beygelzimer A, Langford J (2009) The offset tree for learning with partial labels. *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 129–138.

Bikhchandani S, Sharma S (2000) Herd behavior in financial markets. *IMF Staff papers* 47(3):279–310, ISSN 1020-7635.

Brinkley J (2014) A doubly robust estimator for the attributable benefit of a treatment regime. *Statistics in Medicine* 33(29):5057–5073.

Brock TC (1968) *Implications of commodity theory for value change*, 243–275 (New York: Academic Press, Inc).

Burtch G, Ghose A, Wattal S (2016) Secret admirers: An empirical examination of information hiding and contribution dynamics in online crowdfunding. *Information Systems Research* 27(3):478–496, ISSN 1047-7047.

Burtch G, Gupta D, Martin P (2020) Referral timing and fundraising success in crowdfunding. *Manufacturing & Service Operations Management* .

Burtch G, Hong Y, Liu D (2018) The role of provision points in online crowdfunding. *Journal of Management Information Systems* 35(1):117–144, ISSN 0742-1222.

Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P,

Riddell A (2017) Stan: A probabilistic programming language. *Journal of statistical software* 76(1):1–37, ISSN 1548-7660.

Chakraborty S, Swinney R (2021) Signaling to the crowd: Private quality information and rewards-based crowdfunding. *Manufacturing & Service Operations Management* 23(1):155–169.

Chapelle O, Li L (2011) An empirical evaluation of thompson sampling. *Advances in neural information processing systems*, 2249–2257.

Chen YJ, Dai T, Korpeoglu CG, Körpeoğlu E, Sahin O, Tang CS, Xiao S (2020) Om forum—innovative online platforms: Research opportunities. *Manufacturing & Service Operations Management* 22(3):430–445.

Cialdini RB, Trost MR (1998) *Social influence: Social norms, conformity and compliance*, volume 1-2 of *The handbook of social psychology* (New York, NY, US: McGraw-Hill), 4th edition edition, ISBN 0-19-521376-9 (Hardcover).

Collins FS, Varmus H (2015) A new initiative on precision medicine. *New England Journal of Medicine* 372(9):793–795.

Dai JY, Gilbert PB, Mâsse BR (2012) Partially hidden markov model for time-varying principal stratification in hiv prevention trials. *Journal of the American Statistical Association* 107(497):52–65.

Dawid AP (1979) Conditional independence in statistical theory. *Journal of the Royal Statistical Society: Series B (Methodological)* 41(1):1–15.

DeGraba P (1995) Buying frenzies and seller-induced excess demand. *The RAND Journal of Economics* 26(2):331–342, ISSN 07416261, URL `http://dx.doi.org/10.2307/2555920`.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39(1):1–22.

Denton BT (2018) *Optimization of Sequential Decision Making for Chronic Diseases: From Data to Decisions*, book section 13, 316–348. INFORMS TutORials in Operations Research (INFORMS).

Dimopoulos M, Sonneveld P, Nahi H, Kumar S, Hashim M, Kulakova M, Duran M, Heeg B, Lam

A, Dearden L (2017) Progression-free survival as a surrogate endpoint for overall survival in patients with relapsed or refractory multiple myeloma. *Value in Health* 20(9):A408.

Driscoll JJ, Rixe O (2009) Overall survival: still the gold standard: why overall survival remains the definitive end point in cancer clinical trials. *The Cancer Journal* 15(5):401–405.

Dudík M, Langford J, Li L (2011) Doubly robust policy evaluation and learning. *arXiv preprint arXiv:1103.4601* .

Fatehi S, Wagner MR (2019) Crowdfunding via revenue-sharing contracts. *Manufacturing & Service Operations Management* 21(4):875–893.

Ferreira KJ, Simchi-Levi D, Wang H (2018) Online network revenue management using thompson sampling. *Operations Research* 66(6):1586–1602.

Gao X, Song PXK (2010) Composite likelihood bayesian information criteria for model selection in high-dimensional data. *Journal of the American Statistical Association* 105(492):1531–1540.

Garraway LA (2013) Genomics-driven oncology: Framework for an emerging paradigm. *Journal of Clinical Oncology* 31(15):1806–1814.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) *Bayesian data analysis* (CRC press).

Gelman A, Carlin JB, Stern HS, Rubin DB (1995) *Bayesian data analysis* (New York: Chapman and Hall/CRC), 3rd edition edition, ISBN 1135439419, URL `http://dx.doi.org/https://doi.org/10.1201/b16018`.

Gurr TR (2015) *Why men rebel* (New York: Routledge), ISBN 1317248945, URL `http://dx.doi.org/https://doi.org/10.4324/9781315631073`.

Hari P, Romanus D, Palumbo A, Luptakova K, Rifkin RM, Tran LM, Raju A, Farrelly E, Noga SJ, Blazer M, Chari A (2018) Prolonged duration of therapy is associated with improved survival in patients treated for relapsed/refractory multiple myeloma in routine clinical care in the united states. *Clinical Lymphoma, Myeloma and Leukemia* 18(2):152–160.

Hemer J (2011) A snapshot on crowdfunding. Report, Fraunhofer ISI Working Papers Firms and Regions (No. R2/2011), URL `https://www.econstor.eu/handle/10419/52302`.

Higgins ET (1998) *Promotion and prevention: Regulatory focus as a motivational principle*, vol-

ume 30 of *Advances in experimental social psychology* (San Diego, CA, US: Academic Press), ISBN 0065-2601.

Hoffman MD, Gelman A (2014) The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research* 15(1):1593–1623.

Huang Y, Singh PV, Ghose A (2015) A structural model of employee behavioral dynamics in enterprise social media. *Management Science* 61(12):2825–2844.

Inman JJ, Peter AC, Raghubir P (1997) Framing the deal: The role of restrictions in accentuating deal value. *Journal of Consumer Research* 24(1):68–79, ISSN 1537-5277.

Janzing D, Schölkopf B (2010) Causal inference using the algorithmic markov condition. *IEEE Transactions on Information Theory* 56(10):5168–5194.

Jarvenpaa SL, Tractinsky N, Saarinen L (1999) Consumer trust in an internet store: A cross-cultural validation. *Journal of Computer-Mediated Communication* 5(2), ISSN 1083-6101, URL `http://dblp.uni-trier.de/db/journals/jcmc/jcmc5.htmlJarvenpaaTS99`.

Jiang Y, Ho YC, Yan X, Tan Y (2018) Investor platform choice: Herding, platform attributes, and regulations au - jiang, yang. *Journal of Management Information Systems* 35(1):86–116, ISSN 0742-1222, URL `http://dx.doi.org/10.1080/07421222.2018.1440770`.

Joenssen DW, Müllerleile T (2016) *Limitless Crowdfunding? The Effect of Scarcity Management*, 193–199 (Cham: Springer International Publishing), ISBN 978-3-319-18017-5, URL `http://dx.doi.org/10.1007/978-3-319-18017-5_13`.

Kapoor P, Kumar S, Fonseca R, Lacy MQ, Witzig TE, Hayman SR, Dispenzieri A, Buadi F, Bergsagel PL, Gertz MA (2009) Impact of risk stratification on outcome among patients with multiple myeloma receiving initial therapy with lenalidomide and dexamethasone. *Blood* 114(3):518–521.

Kelly D, Dillingham M, Hudson A, Wiesner K (2012) A new method for inferring hidden markov models from noisy time sequences. *PloS one* 7(1):e29703.

Kumar SK, Rajkumar SV, Dispenzieri A, Lacy MQ, Hayman SR, Buadi FK, Zeldenrust SR, Dingli D, Russell SJ, Lust JA (2008) Improved survival in multiple myeloma and the impact of novel therapies. *Blood* 111(5):2516–2520.

Kyle RA, Rajkumar SV (2008) Multiple myeloma. *Blood* 111(6):2962–2972.

Kyle RA, Rajkumar SV (2009) Treatment of multiple myeloma: a comprehensive review. *Clinical Lymphoma and Myeloma* 9(4):278–288.

Lai TL, Robbins H (1985) Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics* 6(1):4–22.

Langford J, Strehl A, Wortman J (2008) Exploration scavenging. *Proceedings of the 25th international conference on Machine learning*, 528–535.

Lekkakos SD, Serrano A, Ellinger A (2016) Supply chain finance for small and medium sized enterprises: the case of reverse factoring. *International Journal of Physical Distribution & Logistics Management* .

Li L, Chu W, Langford J, Wang X (2011) Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms. *Proceedings of the fourth ACM international conference on Web search and data mining*, 297–306 (ACM).

Lin M, Viswanathan S (2016a) Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science* 62(5):1393–1414, ISSN 0025-1909.

Lin M, Viswanathan S (2016b) Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science* 62(5):1393–1414.

Liu D, Brass DJ, Lu Y, Chen D (2015) Friendships in online peer-to-peer lending: pipes, prisms, and relational herding. *MIS Quarterly* 39(3):729–742, ISSN 0276-7783.

Mariotto AB, Noone AM, Howlader N, Cho H, Keel GE, Garshell J, Woloshin S, Schwartz LM (2014) Cancer survival: an overview of measures, uses, and interpretation. *Journal of the National Cancer Institute Monographs* 2014(49):145–186.

Marshall J, Schwartzberg LS, Bepler G, Spetzler D, El-Deiry WS, Xiao N, Reddy SK, Kim ES, Poste GH, Raghavan D (2016) Novel panomic validation of time to next treatment (tnt) as an effective surrogate outcome measure in 4,729 patients.

Mintz Y, Aswani A, Kaminsky P, Flowers E, Fukuoka Y (2020) Nonstationary bandits with habituation and recovery dynamics. *Operations Research* 68(5):1493–1516.

Moe WW, Schweidel DA (2012) Online product opinions: Incidence, evaluation, and evolution. *Marketing Science* 31(3):372–386.

Mogensen UB, Ishwaran H, Gerds TA (2012) Evaluating random forests for survival analysis using prediction error curves. *Journal of statistical software* 50(11):1.

Mollick E (2014) The dynamics of crowdfunding: An exploratory study. *Journal of business venturing* 29(1):1–16.

Montgomery AL, Li S, Srinivasan K, Liechty JC (2004) Modeling online browsing and path analysis using clickstream data. *Marketing science* 23(4):579–595.

Munshi NC, Anderson KC, Bergsagel PL, Shaughnessy J, Palumbo A, Durie B, Fonseca R, Stewart AK, Harousseau JL, Dimopoulos M, Jagannath S, Hajek R, Sezer O, Kyle R, Sonneveld P, Cavo M, Rajkumar SV, San Miguel J, Crowley J, Avet-Loiseau H (2011) Guidelines for risk stratification in multiple myeloma: report of the international myeloma workshop consensus panel 2. *Blood* .

Murphy SA (2005) An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine* 24(10):1455–1481.

Netzer O, Lattin JM, Srinivasan V (2008) A hidden markov model of customer relationship dynamics. *Marketing science* 27(2):185–204.

Pearl J (1995) Causal diagrams for empirical research. *Biometrika* 82(4):669–688.

Pineau J, Guez A, Vincent R, Panuccio G, Avoli M (2009) Treating epilepsy via adaptive neurostimulation: a reinforcement learning approach. *International journal of neural systems* 19(4):227–240.

Rivera DE, Pew MD, Collins LM (2007) Using engineering control principles to inform the design of adaptive interventions: a conceptual introduction. *Drug and alcohol dependence* 88 Suppl 2(Suppl 2):S31–S40.

Rothman R, Voskoboinik N, Herishanu Y, Orr-Urtreger A, Naparstek E, Trestman S (2012) Prognostic significance of the fish panel for patients with multiple myeloma. *Blood* 120(21):5001.

Satten GA, Longini Jr IM (1996) Markov chains with measurement error: Estimating the

'true'course of a marker of the progression of human immunodeficiency virus disease. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 45(3):275–295.

Schwartz EM, Bradlow ET, Fader PS (2017) Customer acquisition via display advertising using multi-armed bandit experiments. *Marketing Science* 36(4):500–522.

Sherman J, Morrison WJ (1950) Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics* 21(1):124–127.

Soden SE, Saunders CJ, Dinwiddie DL, Miller NA, Atherton AM, Alnadi NA, Leeder JS, Smith LD, Kingsmore SF (2012) A systematic approach to implementing monogenic genomic medicine: genotype-driven diagnosis of genetic diseases. *Journal of Genomes and Exomes* 2012(1):15–24.

Spirtes P (2010) Introduction to causal inference. *Journal of Machine Learning Research* 11(5).

Spirtes PL, Meek C, Richardson TS (2013) Causal inference in the presence of latent variables and selection bias. *arXiv preprint arXiv:1302.4983* .

Strausz R (2017) A theory of crowdfunding: A mechanism design approach with demand uncertainty and moral hazard. *American Economic Review* 107(6):1430–76.

Strehl A, Langford J, Li L, Kakade SM (2010) Learning from logged implicit exploration data. *Advances in neural information processing systems* 23:2217–2225.

Thompson WR (1933) On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika* 25(3/4):285–294.

Thorlund K, Mills E (2012) Stability of additive treatment effects in multiple treatment comparison meta-analysis: a simulation study. *Clinical epidemiology* 4:75.

Tomkins S, Liao P, Klasnja P, Murphy S (2020) Intelligentpooling: Practical thompson sampling for mhealth. *arXiv preprint arXiv:2008.01571* .

van Beurden-Tan CH, Franken MG, Blommestein HM, Uyl-de Groot CA, Sonneveld P (2017) Systematic literature review and network meta-analysis of treatment outcomes in relapsed and/or refractory multiple myeloma. *Journal of Clinical Oncology* 35(12):1312–1319.

Van Beurden-Tan CHY, Franken M, Blommestein H, Uyl-De Groot CA, Sonneveld P (2016) Systematic literature review and network meta-analysis of treatments for relapsed/refractory multiple myeloma patients. *Blood* 128(22):2144.

Van Buuren S (2007) Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical methods in medical research* 16(3):219–242.

van de Klundert J (2016) *Healthcare Analytics: Big Data, Little Evidence*, book section 14, 307–328. INFORMS TutORials in Operations Research (INFORMS).

Vehtari A, Gelman A (2014) Waic and cross-validation in stan.

Vehtari A, Gelman A, Gabry J (2016) Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing* 27(5):1413–1432.

Viterbi A (1967) Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory* 13(2):260–269.

Wang Y, Powell W (2016) An optimal learning method for developing personalized treatment regimes.

Wang Y, Wang C, Powell W (2017) Optimal learning for sequential decision making for expensive cost functions with stochastic binary feedbacks.

Weinmann M, Tietz M, Simons A, vom Brocke J (2017) Get it before it's gone? how limited rewards influence backers' choices in reward-based crowdfunding. *Proceedings of the International Conference on Information Systems* (Atlanta, GA: Association for Information Systems).

Willan J, Eyre TA, Sharpley F, Watson C, King AJ, Ramasamy K (2016) Multiple myeloma in the very elderly patient: challenges and solutions. *Clinical interventions in aging* 11:423.

Worchel S, Lee J, Adewole A (1975) Effects of supply and demand on ratings of object value. *Journal of Personality and Social Psychology* 32(5):906–914, ISSN 1939-1315(Electronic),0022-3514(Print), URL `http://dx.doi.org/10.1037/0022-3514.32.5.906`.

Wu Z (2010) A hidden markov model for earthquake declustering. *Journal of Geophysical Research: Solid Earth* 115(B3).

Wuttke DA, Blome C, Heese HS, Protopappa-Sieke M (2016) Supply chain finance: Optimal introduction and adoption decisions. *International Journal of Production Economics* 178:72–81.

Yan L, Tan Y (2014) Feeling blue? go online: An empirical study of social support among patients. *Information Systems Research* 25(4):690–709.

Yang L, Wang Z, Hahn J (2017) Scarcity strategy in crowdfunding: An empirical exploration.

Proceedings of the International Conference on Information Systems (Atlanta, GA: Association for Information Systems).

Zhang B, Tsiatis AA, Laber EB, Davidian M (2012) A robust method for estimating optimal treatment regimes. *Biometrics* 68(4):1010–1018.

Zhang J, Liu P (2012) Rational herding in microloan markets. *Management science* 58(5):892–912, ISSN 0025-1909.

Zhang T, Zhang CY, Pei Q (2019a) Misconception of providing supply chain finance: Its stabilising role. *International Journal of Production Economics* 213:175–184.

Zhang Y, Li B, Luo X, Wang X (2019b) Personalized mobile targeting with user engagement stages: Combining a structural hidden markov model and field experiment. *Information Systems Research* 30(3):787–804.

Appendix A

ESTIMATION STRATEGY FOR HIERARCHICAL BAYESIAN MODEL

A.1 Bayesian Estimation and Model Hierarchy

First, let $\theta_i = \bar{\theta} + \xi_i$, where ξ_i indicates the individual-level deviations. We thus can directly estimate the deviations using $\xi_i \sim N(0, \Sigma)$ and include $\bar{\theta}$ in the population parameter set ψ. As mentioned in the estimation strategy section, we use Monte Carlo Markov Chain (MCMC) methods to estimate parameters in our model. To be more specific, we combine Gibbs sampler and Metropolis algorithm to recursively make draws from the following conditional distributions of the model parameters:

$$\xi_i \mid Y_i, D_i, \psi, R_i;$$

$$\Sigma \mid \xi_i;$$

$$\psi \mid Y, D, \xi_i, R_i;$$

$$R_{ikt} \mid Y_{ikt}, D_{ikt}, R_{ik(t-1)}, \psi, \xi_i;$$

Step 1: generate $\xi_i \mid Y_i, D_i, \psi, R_i$. The conditional distribution of θ_i is

$$f(\xi_i \mid Y_i, D_i, \psi, R_i) \propto |\Sigma|^{-1/2} \exp[-1/2(\xi_i)'\Sigma^{-1}(\xi_i)] L(Y_i, D_i|\xi_i, R_i, \psi) \tag{A.1}$$

Where R_i denotes the vector of individual i's belief on K guarantors in all periods. Clearly, this posterior distribution does not have a closed-form; therefore, we use the Metropolis-Hasting algorithm to generate new draws with a random walk proposal density. Given the last draw θ_i^r in round r, the probability that proposed θ_i^* will be accepted in round $r+1$ is calculated using the following formula:

$$P(\text{accept}) \propto \frac{f(\xi_i^* \mid Y_i, D_i, \psi, R_i)}{f(\xi_i^r \mid Y_i, D_i, \psi, R_i)} \tag{A.2}$$

$$= \frac{|\Sigma|^{-1/2} \exp[-1/2(\xi_i^*)'\Sigma^{-1}(\xi_i^*)] L(Y_i, D_i|\xi_i^*, R_i, \psi)}{|\Sigma|^{-1/2} \exp[-1/2(\xi_i^r)'\Sigma^{-1}(\xi_i^r)] L(Y_i, D_i|\xi_i^r, R_i, \psi)} \tag{A.3}$$

Step 2: Generate $\Sigma \mid \xi_i$ from

$$\Sigma \mid \xi_i \sim IW(p + I + 2, G_0^{-1} + \sum_{i=1}^{I} (\theta_i - \bar{\theta})'(\theta_i - \bar{\theta}), \tag{A.4}$$

where p is the dimension of Σ and I is the number of individuals in the dataset, and G_0^{-1} is the identity matrix.

Step 3: Generate $\psi \mid Y, D, \xi_i, R_i$.The conditional distribution of ψ is

$$f(\psi \mid Y, D, \xi_i, R_i) = |\Sigma_\psi|^{-1/2} \exp\left[-\frac{1}{2(\psi - \psi_0)'\Sigma_\psi^{-1}(\psi - \psi_0)}\right] L(Y, D \mid \psi, \xi_i) L(R_i\psi). \tag{A.5}$$

Similar to what we have done for θ_i, we use the Metropolis-Hasting methods to make draws for α. The probability of acceptance is

$$P(\text{accept}) \propto \frac{f(\psi^* \mid Y, D, \xi_i, R_i)}{f(\psi^r \mid Y, D, \xi_i, R_i)} \tag{A.6}$$

$$= \frac{|\Sigma_\psi|^{-1/2} \exp\left[-\frac{1}{2(\psi^* - \psi_0)'\Sigma_\psi^{-1}(\psi^* - \psi_0)}\right] L(Y, D \mid \psi, \xi_i) L(R_i\psi)}{|\Sigma_\psi|^{-1/2} \exp\left[-\frac{1}{2(\psi^r - \psi_0)'\Sigma_\psi^{-1}(\psi^r - \psi_0)}\right] L(Y, D \mid \psi, \xi_i) L(R_i\psi)} \tag{A.7}$$

Note that among the variable set ψ, we've set the unobserved reliability level to be within $[-1, 1]$ for the identification purpose. Thus, we estimate the log odds transformation of it, which maps the support $[-1, 1]$ to a real line.

Step 4: Sequentially draw $R_{ikt} \mid Y_{ikt}, D_{ikt}, R_{ik(t-1)}, \psi, \xi_i$. Finally, we sequentially draw R_{ikt} for each individual i from $t = 1$ to T. The conditional distribution of R_{ikt}, given $R_{ik(t-1)}$ is

$$f(R_{ikt} \mid Y_{ikt}, D_{ikt}, R_{ik(t-1)}, \psi, \xi_i) \propto \tag{A.8}$$

$$|\nu_{ikt}^2|^{-1/2} \exp\left[-\frac{1}{2(R_{ikt} - \bar{R}_{ikt})'(\nu_{ikt}^2)^{-1}(R_{ikt} - \bar{R}_{ikt})}\right] L(Y_{ikt}, D_{ikt} \mid \psi, \theta_i, R_{ikt}) \tag{A.9}$$

The probability of acceptance is

$$P(\text{accept}) \propto \frac{f(R^*_{ikt} \mid Y_{ikt}, D_{ikt}, R_{ik(t-1)}, \psi, \xi_i)}{f(R^r_{ikt} \mid Y_{ikt}, D_{ikt}, R_{ik(t-1)}, \psi, \xi_i)} \tag{A.10}$$

$$= \frac{|\nu^2_{ikt}|^{-1/2} \exp\left[-\frac{1}{2(R^*_{ikt}-\bar{R}_{ikt})'(\nu^2_{ikt})^{-1}(R^*_{ikt}-\bar{R}_{ikt})}\right] L(Y_{ikt}, D_{ikt} \mid \psi, \theta_i, R^*_{ikt})}{|\nu^2_{ikt}|^{-1/2} \exp\left[-\frac{1}{2(R^r_{ikt}-\bar{R}_{ikt})'(\nu^2_{ikt})^{-1}(R^r_{ikt}-\bar{R}_{ikt})}\right] L(Y_{ikt}, D_{ikt} \mid \psi, \theta_i, R^r_{ikt})} \tag{A.11}$$

R_{ikt} and μ^2_{ikt} in the above equations are calculated based on Equation (3.14) and (3.15).

A.2 Identification

We now briefly discuss some intuition as to how the parameters in our models are identified. In our model, the individual investors make investment decisions based on their (perceived) utility. The variance of the two signals, σ^2_F and σ^2_D, are identified from the dynamics of individuals' investment behaviors over time. Every time an individual investor gets a signal from the repayment activities, she would update her belief, which in turns affects her investment behavior.

The individual's prior beliefs about the mean reliability of the guarantors R_0 is identified from the direction of change in investment behavior as she receives the repayment signals. If, after receiving a repayment signal from the money borrowers, her investment behavior, for example, the incidence probability has increased, then it implies that the individual's prior belief about mean reliability was lower than her updated perception. Thus, the direction of the change and the extent of the change from the initial investment behaviors identify the prior belief about the mean reliability of the guarantors. The true reliability level, R_k, where $k = 1, \cdots, K$, is identified from the likelihood of a project with a certain guarantor being invested and the amount of being invested. We have already seen that as the individual investor learns, the reliability belief, R_{ikt}, evolves to the true reliability level R_k. At the extreme, i.e., the steady-state, the quality belief is the true quality. Since we include R_{ikt} linearly into our utility model, this helps us separate it from the base utility level $\bar{\beta}_0$ and $\bar{\gamma}_0$, identify the population mean of the slope parameter $\bar{\beta}_1$, $\bar{\beta}_2$, $\bar{\gamma}_1$ and $\bar{\gamma}_2$; and also be able to make interpretations on the linear and quadratic interaction terms. Besides, we fix $\sigma_(R_0)^2$,

the variance of individuals' initial belief to 1, and the estimates for σ_F^2 and σ_D^2 then represent the relative magnitude of the signal variance compared to individuals' initial value.

Appendix B

VARIABLE AND TREATMENT SPECIFICATION IN CLINICAL DATASET

B.1 Treatment Combinations in Multi-Armed Bandit Problem

In our manuscript, we use a unique ID (e.g., Comb X) to refer to each distinct treatment combination. The detailed drug list of each combination is presented in the table below.

Table B.1: Explanations of Treatment Combinations

Variable	Drugs
Comb. 1	Stem Cell Transplant (SCT)
Comb. 1	Lenalidomide
Comb. 3	Bortezomib
Comb. 4	Bortezomib, Dexamethasone, Lenalidomide
Comb. 5	Bortezomib, Cyclophosphamide, Dexamethasone
Comb. 6	Dexamethasone
Comb. 7	Dexamethasone, Thalidomide
Comb. 8	Cyclophosphamide, Dexamethasone
Comb. 9	Cyclophosphamide, Dexamethasone, Etoposide
Comb. 10	Carfilzomib, Dexamethasone, Pomalidomide
Comb. Other	All other treatment combinations.

B.2 Hidden Markov Model Variable Specification

Factors that affect a patient's survival outcome are called prognosis factors. Among all possible factors for patients with multiple myeloma proposed by previous studies, some are directly related to the disease progression, such as how advanced the disease is in its spread

in the body, or which organs are involved. We use LASSO to select the most relevant prognostic factors in predicting patient's treatment outcome in our hidden Markov model (HMM). Among a large set of available features, we end up identifying eleven important factors, including the degree of bone lesions detected by MRI test, and other ten features identified by blood or urine tests. For each of the eleven laboratory tests, we present its lab code, full name, as well as the panel it belongs to in Table B.2.

Table B.2: Explanation of Lab Codes

Category	Abbreviations	Description	Panel
Oncology	FLCR	Free Light Chain Ratio	Tumor biomarker
	LD	Lactate Dehydrogenase	Tumor biomarker
	MPROTEIN	Monoclonal Protein	Tumor biomarker
	BETA2	Beta-2 Microglobulin	Tumor biomarker
Other laboratory tests	HCT	Hematocrit	Blood count
	IGA	Immunoglobulins A	Immunoglobulin test
	IGG	Immunoglobulins G	Immunoglobulin test
	IGM	Immunoglobulins M	Immunoglobulin test
	ALB	Albumin	Liver function
	CRE	Creatinine	Kidney function
Magnetic Resonance Imaging	MRI	Bone Lesions	Bone scan

Other factors depend on the individual's demographics, including gender, age at diagnosis, or disease burden, which affects the individual's capacity to tolerate intensive treatment. We have summarized the variables we included in the transition model, and the treatment outcome model in Table B.3.

Table B.3: Variable Specification for HMM

Variable Set	Variable Category	Variable	Description
Variables that affect risk transition (T_{ik})	Genetic Characteristics	del(17p)	Deletions of the short arm of 17
		del(13q)	Chromosome 13q deletion
		t(11;14)	Translocation of chromosome 11 and 14
		del(1p)	Deletions of the short arm of chromosome 1
		t(14;16)	Translocation of chromosome 14 and 16
		t(4;14)	Translocation of chromosome 4 and 14
	Past exposure to treatment agents	Chemotherapy	The mechanism of treatment agent belongs to
		Immunotherapy	
		Proteosome Proteasome Inhibitor	
		Stem Cell Transplant	
		Steroid	
	Treatments	Current treatment combination	A set of dummy variables indicating the treatment combination the patient was given
		Number of agents	The total number of agents included in the treatment combination
		Number of agents (quadratic)	
		Drug cost	The cost of the drugs included in the treatment combination
Variables that affect treatment outcomes (O_{ik})	Demographics	Age at diagnosis	
		Patient sex	
	Disease progression	Lag PFS	Patient's PFS for the last LOT
		LOT	Line of therapy
		LOT (quadratic)	Line of therapy
	Laboratory test results	Hematocrit (HCT)	A test for blood count
		Immunoglobulins A (IGA)	An immunoglobulin test
		Immunoglobulins G (IGG)	
		Immunoglobulins M (IGM)	
		Albumin (ALB)	A test for liver function
		Creatinine (CRE)	A test for kidney function
		Bone marrow plasma cells	An important indicator of multiple myeloma
		Bone lesions	Result from Magnetic Resonance Imaging (MRI) describing the degree of Bone lesions.
		Free Light Chain Ratio (FLCR)	Tumor biomarker

Appendix C

TECHNICAL SUPPLEMENTS ON THE HIDDEN MARKOV MODEL

In this section, we present the supplement details on our generative causal HMM model.

C.1 Estimation and Model Comparison

HMM parameters consist of three components: 1) the initial state distribution π, 2) the parameters for state transition probability distribution $[\kappa, r]$, and 3) the parameters for the observed outcome probability distribution λ. We start with using the latent class model to estimate the initial distribution for the latent risk level π; then, we use the Maximum Likelihood Estimation (MLE) to estimate the model parameters $\upsilon = [\kappa, r, \lambda]$.

One remaining parameter still to be determined is the number of states. While this parameter is taken as given when we derive the likelihood function for our full model, we select the one with the best model performance among different alternatives using the high-dimensional Bayesian Information Criteria (HD-BIC) proposed by Gao and Song (2010), the result of which suggests that the three-state HMM outperforms other specifications. Specifically, for each number of states, we calculate

$$\text{BIC} = c\ln(k)N - 2 \times \ln(L), \tag{C.1}$$

where L is the estimated likelihood of the model using MLE, k is the number of parameters, and N is the sample size, which is the number of patients in our data. c is a multiplicative factor with various choices. In empirical studies, using $c = 1$ is shown to have good empirical results (Gao and Song 2010).

First, we estimate the HMM with $s = 2, 3, 4$ separately. Then, we calculate the HD-BIC for each model. The model with a smaller HD-BIC will be selected. As can be seen from

Table C.1, the model with the number of hidden states $s = 3$ has the best performance.

Table C.1: Model Fit and Comparison

Number of States	**Log-likelihood**	**Variables**	**HD-BIC**
2	-2131.76	90	13263.13
3	-1689.21	138	13232.92
4	-1583.75	188	13640.38

C.2 Estimation Results

Initial state distribution probabilities π estimated using the latent class model are $[0.250, 0.419, 0.33]$. We show the estimation results for the state transition parameters in Table C.2. The parameters that influence patient's risk level transition can be classified into three groups: 1) patients' demographic information that is fixed over time periods, 2) patients' genetic characteristics indicated by their FISH test results, 3) previous exposure to various drugs.

We present the estimation results for the transition parameters in Table C.2. First, we notice that although the transition patterns are dissimilar across patients at different risk levels, the manifestations of some cytogenetic features do have a statistical significant influence on patients' underlying risk levels. More interestingly, our findings reveal that a patient's previous exposure to certain drugs will trigger a transition as well, which provides an empirical support for the hypothesis proposed by Munshi et al. (2011) mentioned in the previous section. The change in genetic features along a patient's disease course and the variations in the treatment strategies are well recorded in our dataset, making it necessary to continuously monitor patients' genetic profile and reassess their risk status to deliver better medical care. The estimation results provide evidence of the dynamic pattern of patients' unobserved risk levels.

The estimated parameters for state-dependent treatment outcomes are presented in Table

Table C.2: Estimation Results for Transition Parameters

Variable Group	Parameters	State 1 (High Risk)		State 2 (Medium Risk)		State 3 (Low Risk)	
		Variables that affect state transition					
Genetic Characteristics	del17p	3.6409	(2.5283)	−0.1599	(0.8600)	0.5280	(4.0923)
	del13q	−12.0843 ∗∗	(5.0011)	−0.4978	(0.7326)	0.5553	(2.2656)
	t11_14	1.7676	(3.5872)	−0.6640 ∗∗	(0.3705)	0.1755	(4.4622)
	del1p	−2.2867	(4.5256)	1.7592	(2.9013)	−5.1307∗	(3.1934)
	t14_16	4.1979	(2.8780)	−1.5924	(1.6593)	−6.4308	(8.4684)
	t4_14	−13.5757 ∗∗	(4.5318)	0.1563	(0.8824)	−3.5334∗	(2.3051)
Treatment Mechanism	Chemotherapy	8.6563 ∗∗	(2.9107)	0.3465 ∗∗	(0.1699)	−0.6849	(2.7491)
	Immunotherapy	0.9102	(2.5481)	−1.1867	(0.7043)	−0.0342	(2.5548)
	Proteosome_inhibitor	0.5505	(2.9495)	−0.2173	(0.6061)	0.7340	(2.6767)
	Stem_Cell_Transplant	0.2646	(3.0360)	4.1854∗∗∗	(1.1861)	0.6321	(3.6858)
	Steroid	0.3266	(2.6040)	1.5261 ∗∗	(0.5644)	−1.2160	(2.8015)
		Thresholds					
	State 1 (High Risk)			1.0063	(2.4509)	6.4272	(4.9086)
	State 2 (Medium Risk)	−1.3429 ∗∗	(0.5035)			3.2816	(4.0658)
	State 3 (Low Risk)			−1.0656	(1.7737)	0.7049	(2.2583)

C.3. In the result table, we categorize the factors that influence patients' PFS into four groups: 1) assigned treatment, 2) patients' demographics that are fixed over time, 3) patients' health condition, which is correlated to the disease burden, and 4) various laboratory test results that indicate the current disease progression.

The variations in the coefficients of parameters across different states indicate a systematic difference in treatment response patterns when patients change to a different risk level. This finding suggests a high-level heterogeneity in treatment response across patients, and even for the same patient across LOTs, due to risk level transitions. We also identify several significant prognostic factors that may affect myeloma patients' survival outcomes besides the treatments, such as patients' sex and their overall health conditions as indicated by various laboratory results.

In sum, we can draw two main insights from the HMM estimation results. First, they provide evidence of the dynamic pattern of patients' unobserved risk levels. Second, the state-dependent outcome parameters suggest that the effectiveness of different treatment

Table C.3: Estimation Results for Outcome Parameters

Variable Group	Parameters	State 1 (High Risk)		State 2 (Medium Risk)		State 3 (Low Risk)	
Treatments	Comb.01	−0.4475	(0.3117)	0.1253***	(0.1253)	0.1253	(0.1069)
	Comb.02	0.7867 **	(0.2421)	0.1950	(0.1950)	−0.2267	(0.2322)
	Comb.03	3.4978***	(0.6568)	0.2557	(0.2557)	−0.2767	(0.3425)
	Comb.04	3.8557***	(0.4985)	0.2152	(0.2152)	−0.6529	(0.3962)
	Comb.05	0.2617	(0.4579)	0.2171	(0.2171)	−0.8455	(0.6251)
	Comb.06	−0.1881	(0.4827)	0.2767	(0.2767)	−0.9135 **	(0.4518)
	Comb.07	3.5601***	(0.6061)	0.2024	(0.2024)	−0.0478	(0.5053)
	Comb.08	1.0148***	(0.2925)	0.2381	(0.2381)	−1.0363	(1.0403)
	Comb.09	−1.7351***	(0.2572)	0.0849***	(0.0849)	−0.3332	(0.3916)
	Comb.10	−1.6193***	(0.2427)	0.1183 **	(0.1183)	0.7422	(0.5987)
	Num of Agents	0.4033	(0.2415)	0.1304	(0.1304)	−2.0729 **	(0.7306)
	Num of Agents (Quadratic)	−0.7241***	(0.2096)	0.1111 **	(0.1111)	−3.0238***	(0.5235)
	Drug Cost	−1.2546***	(0.2153)	0.0757	(0.0757)	−0.3710	(0.3646)
Demographics	Age at Diagnosis	−0.0413	(0.0480)	0.0389	(0.0389)	0.0265	(0.3563)
	PatientSex	0.4345 **	(0.1891)	0.0654	(0.0654)	−0.0487	(0.1221)
	Lag PFS	0.0640	(0.0400)	0.0562 **	(0.0562)	−0.1558 **	(0.0616)
Disease Progression	Line	0.2591***	(0.0465)	0.0236	(0.0236)	0.1723	(0.1973)
	Line (Quadratic)	−0.7493***	(0.1244)	0.0640 **	(0.0640)	−0.0297	(0.1360)
Laboratory result	HCT	0.1604	(0.1110)	0.0401	(0.0401)	−0.0837	(0.1678)
	IGA	0.0308	(0.0416)	0.0503	(0.0503)	−0.1291	(0.1000)
	IGG	0.1122	(0.0899)	0.0326	(0.0326)	−0.1023	(0.1225)
	IGM	−0.0634	(0.0900)	0.0822 **	(0.0822)	−0.0163	(0.1301)
	ALB	−0.1604	(0.0820)	0.0264	(0.0264)	0.1171	(0.1624)
	CRE	0.0896*	(0.0482)	0.0292	(0.0292)	0.0347 **	(0.0112)
	Bone PLAZ	−0.1510	(0.0911)	0.0368	(0.0368)	−0.4604 **	(0.1628)
	MRI (Abnormal)	−0.1016	(0.1282)	0.0701	(0.0701)	0.3173	(0.2120)
	FLCR (High)	−0.0237	(0.1526)	0.0600	(0.0600)	−0.3908	(0.2717)
	LD (High)	−0.4607	(0.2948)	0.0732	(0.0732)	0.8770***	(0.1571)
	LD (Low)	0.1934	(0.2640)	0.0662	(0.0662)	0.6888 **	(0.2914)
	MPROTEIN (High)	−0.1835	(0.1256)	0.0678	(0.0678)	−0.4613	(0.3699)
	BETA2 (High)	0.4042 **	(0.1530)	0.1100	(0.1100)	−0.3156	(0.4424)
Variance	Sigma	0.4852	(0.3509)	0.3255	(0.3255)	0.7555***	(0.2119)

combinations and the influence of various prognosis factors do vary significantly among patients at different risk levels. This heterogeneity is usually not well considered in the conventional treatment strategy, making the efforts of developing an adaptive treatment strategy that takes into account individual-level real-time dynamics both urgent and indispensable.

C.3 Robustness Check Using Multiple Imputation

In Table C.5 and Table C.4, we report the estimation results using the Multiple Imputation by Chained Equations (MICE) algorithm proposed by Van Buuren (2007) on our dataset. We found that the coefficients from the estimations are qualitatively similar to our main results. Most of the influential variables we identified in our main results are still statistically significant in this alternative model specification, indicating the robustness of our results to the missing data imputation method.

Table C.4: Robustness Check Using MICE: Estimation Results for Transition Parameters

Variable Group	Parameters	State 1 (High Risk)		State 2 (Medium Risk)		State 3 (Low Risk)	
		Variables that affect state transition					
Genetic Characteristics	del17p	−1.4273	(1.5035)	−0.0319	(0.4268)	0.0927	(7.7562)
	del13q	−0.3666***	(0.0594)	−1.0926	(0.8996)	−0.2988	(3.0420)
	t11_14	1.0044	(0.6785)	0.2623	(1.6429)	1.2433	(5.2077)
	del1p	−4.5298 **	(1.8841)	−0.4314	(2.5983)	7.0174	(4.3233)
	t14_16	0.0029	(1.0468)	4.2133	(3.9674)	1.7421	(5.2707)
	t4_14	−1.2143 **	(0.6540)	2.1128 **	(1.0145)	−10.2938	(6.7824)
Treatment Mechanism	Chemotherapy	0.3245*	(0.1895)	2.4630***	(0.5103)	6.2270	(4.7169)
	Immunotherapy	0.8810	(0.5064)	−1.2699 **	(0.5885)	9.2799	(5.8903)
	Proteosome_inhibitor	−1.2170	(1.2025)	−0.7269	(0.8827)	4.9279	(4.0454)
	Stem_Cell_Transplant	1.5510	(0.8774)	−2.5979 **	(0.8148)	−1.8322	(1.9504)
	Steroid	0.6769 **	(0.3335)	0.3237*	(0.3002)	−5.9156	(4.9008)
		Thresholds					
	State 1 (High Risk)			0.6945	(0.6147)	4.6160	(3.2686)
	State 2 (Medium Risk)	−1.9907	(1.1063)			10.9592	(9.0077)
	State 3 (Low Risk)			6.1091	(4.6918)	7.0883	(2.6228)

Table C.5: Robustness Checks using MICE: Estimation Results for Outcome Parameters

Variable Group	Parameters	State 1 (High Risk)		State 2 (Medium Risk)		State 3 (Low Risk)	
Treatments	Comb.01	1.0235***	(0.0563)	0.4064	(0.4064)	0.1762**	(0.0541)
	Comb.02	0.6057***	(0.1199)	0.2196	(0.2196)	0.0716	(0.1049)
	Comb.03	3.5266***	(0.3602)	0.7079**	(0.7079)	0.5844***	(0.1709)
	Comb.04	2.5879***	(0.2333)	0.5125	(0.5125)	−0.3686**	(0.1817)
	Comb.05	1.0465***	(0.2103)	0.3916	(0.3916)	−0.7578***	(0.2171)
	Comb.06	−0.2806**	(0.1018)	0.2289**	(0.2289)	−0.0095	(0.2580)
	Comb.07	3.0107***	(0.3834)	0.3916	(0.3916)	−0.5094***	(0.1003)
	Comb.08	1.5954***	(0.2409)	0.3091**	(0.3091)	−0.2664	(0.5582)
	Comb.09	−0.2465**	(0.1142)	0.2252***	(0.2252)	−0.5689	(1.8830)
	Comb.10	−0.3266**	(0.1264)	0.3412***	(0.3412)	−0.4431**	(0.1793)
	Num of Agents	0.2213**	(0.1007)	0.1619	(0.1619)	−1.3290	(1.1266)
	Num of Agents (Quadratic)	−0.2339**	(0.1056)	0.1721**	(0.1721)	−0.4600	(0.3614)
	Drug Cost	−0.9664***	(0.1119)	0.2299	(0.2299)	0.0569	(0.1035)
Demographics	Age at Diagnosis	0.0448	(0.0348)	0.0274	(0.0274)	−0.1159	(0.0755)
	PatientSex	−0.0474	(0.0313)	0.0837	(0.0837)	0.7122***	(0.1456)
Disease Progression	Lag PFS	0.0746**	(0.0361)	0.0264	(0.0264)	−0.3224***	(0.0198)
	Line	0.0447**	(0.0172)	0.0383	(0.0383)	0.5244***	(0.1153)
	Line (Quadratic)	−0.1497**	(0.0553)	0.1032**	(0.1032)	0.2513***	(0.0379)
Laboratory result	HCT	0.0185	(0.0388)	0.0882	(0.0882)	−0.5184***	(0.0675)
	IGA	0.0739**	(0.0299)	0.0229	(0.0229)	0.2500***	(0.0488)
	IGG	0.0607	(0.0344)	0.0992	(0.0992)	0.3216***	(0.0942)
	IGM	0.2392**	(0.0821)	0.0964	(0.0964)	0.4914**	(0.1830)
	ALB	−0.0369	(0.0191)	0.0514	(0.0514)	0.5370***	(0.0788)
	CRE	−0.0001	(0.0139)	0.0337	(0.0337)	1.1928***	(0.1619)
	Bone PLAZ	−0.0840**	(0.0364)	0.1067**	(0.1067)	0.0344	(0.0410)
	MRI (Abnormal)	−0.0546	(0.0735)	0.0578	(0.0578)	−0.6937***	(0.1045)
	FLCR (High)	0.0119	(0.0376)	0.0752	(0.0752)	−0.3540**	(0.1144)
	LD (High)	0.0386	(0.0333)	0.1344***	(0.1344)	0.1135	(0.1155)
	LD (Low)	0.1012***	(0.0195)	0.1112***	(0.1112)	0.5992**	(0.1940)
	MPROTEIN (High)	0.0502	(0.0477)	0.1263***	(0.1263)	0.0246	(0.1216)
	BETA2 (High)	0.0222	(0.0487)	0.4613**	(0.4613)	−0.3083	(0.2246)
Variance	Sigma	0.5730	(0.3191)	0.2412	(0.2412)	0.1389	(0.2742)

C.4 Robustness Check on Hidden Markov Model Adding Fixed Effects

In Table C.6 and Table C.7, we report the estimation results of the HMM model by adding individual fixed effects into the model. The patient-level heterogeneity parameters are insignificant after controlling for a comprehensive set of covariates, and as we can see from the results, the coefficients remain qualitatively the same, suggesting the robustness of our findings.

Table C.6: Robustness Check Adding Fixed Effects: Estimation Results for Transition Parameters

Variable Group	Parameters	State 1 (High Risk)		State 2 (Medium Risk)		State 3 (Low Risk)	
		Variables that affect state transition					
Genetic Characteristics	del17p	4.0499	(2.5283)	−1.0012	(0.8825)	−0.2674	(4.1600)
	del13q	−12.2697**	(5.0011)	−1.2504	(0.8275)	−0.3205	(2.3139)
	t11_14	1.6774	(3.5872)	−0.8047*	(0.4398)	0.8824	(4.4683)
	del1p	−1.6474	(4.5256)	1.6107	(2.9367)	−5.6138	(4.2537)
	t14_16	4.6609	(2.8780)	−1.4404	(1.7202)	−5.6804	(8.5430)
	t4_14	−13.2317*	(6.5318)	−0.304	(0.8927)	−3.4941	(2.3943)
Treatment Mechanism	Chemotherapy	8.1416**	(2.9107)	0.4091*	(0.2437)	−0.8479	(2.8448)
	Immunotherapy	1.8595	(2.5481)	−1.3021	(0.7967)	0.7965	(2.6167)
	Proteosome inhibitor	0.1323	(2.9495)	−0.5123	(0.6771)	0.6573	(2.6844)
	Stem Cell Transplant	0.4006	(3.0360)	3.3369***	(1.2396)	−0.1933	(3.6864)
	Steroid	0.182	(2.6040)	0.971	(0.6266)	−1.3458	(2.8325)
		Thresholds					
	State 1 (High Risk)			1.0366	(2.4790)	6.0284	(4.9086)
	State 2 (Medium Risk)	−1.6253**	(0.5046)			3.5009	(4.1613)
	State 3 (Low Risk)			−0.7355	(1.8547)	1.3837	(2.2719)

Table C.7: Robustness Checks Adding Fixed Effects: Estimation Results for Outcome Parameters

Variable Group	Parameters	State 1 (High Risk)		State 2 (Medium Risk)		State 3 (Low Risk)	
Treatments	Comb.01	−1.1213	(0.3117)	0.1860***	(0.1860)	0.0509	(0.2055)
	Comb.02	1.0203 **	(0.2421)	0.2307	(0.2307)	−0.0224	(0.2690)
	Comb.03	2.8114***	(0.6568)	0.2626	(0.2626)	−0.8511	(0.4074)
	Comb.04	4.4522***	(0.4985)	0.2298	(0.2298)	0.0411	(0.4279)
	Comb.05	0.4252	(0.4579)	0.2269	(0.2269)	−0.5974	(0.6336)
	Comb.06	−0.3451	(0.4827)	0.2879	(0.2879)	−0.8833 **	(0.5164)
	Comb.07	3.2793***	(0.6061)	0.2535	(0.2535)	−0.1345	(0.5396)
	Comb.08	1.4575***	(0.2925)	0.2674	(0.2674)	−0.8878	(1.0866)
	Comb.09	−0.6895 **	(0.9572)	0.1296***	(0.1296)	0.1789	(0.4668)
	Comb.10	−1.4747***	(0.2427)	0.1652	(0.1652)	0.4904	(0.6813)
	Num of Agents	0.8253	(0.2415)	0.1944	(0.1944)	−1.3174*	(0.7933)
	Num of Agents (Quadratic)	−0.7639***	(0.2096)	0.1474 **	(0.1474)	−3.3862***	(0.5239)
	Drug Cost	−1.3254***	(0.2153)	0.1387	(0.1387)	0.0776	(0.4536)
Demographics	Age at Diagnosis	−0.5306	(0.0480)	0.0873	(0.0873)	−0.3826	(0.3790)
	Patient Sex	0.4297 **	(0.1891)	0.1173	(0.1173)	−0.6009	(0.2161)
Disease Progression	Lag PFS	−0.1229	(0.0400)	0.1524*	(0.1524)	−0.2773 **	(0.0727)
	Line	0.1544***	(0.0465)	0.1173	(0.1173)	0.9721	(0.2627)
	Line (Quadratic)	−1.8142***	(0.1244)	0.0642 **	(0.0642)	0.1559	(0.1610)
Laboratory result	HCT	−0.0516	(0.1110)	0.0582	(0.0582)	0.047	(0.1926)
	IGA	0.4933	(0.0416)	0.1021	(0.1021)	−0.2834	(0.1917)
	IGG	0.0893	(0.0899)	0.1292	(0.1292)	0.7012	(0.1975)
	IGM	−0.0036	(0.0900)	0.1201 **	(0.1201)	−0.0028	(0.2170)
	ALB	0.4049	(0.0820)	0.0772	(0.0772)	−0.2481	(0.1643)
	CRE	−0.1851	(0.4822)	0.0326	(0.0326)	0.1297 **	(0.0175)
	Bone PLAZ	−0.1293	(0.0911)	0.1054	(0.1054)	−0.2021	(0.2193)
	MRI (Abnormal)	0.3259	(0.1282)	0.1278	(0.1278)	0.2319	(0.2971)
	FLCR (High)	−0.9063	(0.1526)	0.0868	(0.0868)	−1.0579	(0.2855)
	LD (High)	−1.0024	(0.2948)	0.1729	(0.1729)	1.2717***	(0.2273)
	LD (Low)	−0.6451	(0.2640)	0.0971	(0.0971)	0.9722 **	(0.3117)
	MPROTEIN (High)	−0.0822	(0.1256)	0.0818	(0.0818)	−0.8789	(0.4519)
	BETA2 (High)	0.4267 **	(0.1530)	0.1135	(0.1135)	−0.4593	(0.4733)
Variance	Sigma	0.5094	(0.3509)	0.2099	(0.2099)	0.2870***	(0.7009)

Appendix D

HYPER-PARAMETERS UPDATING IN MULTILEVEL BAYESIAN LINEAR THOMPSON SAMPLING

D.1 EM Algorithm for Updating Hyper-parameters

Algorithm 3: E-M algorithm for empirical Bayes estimates of hyper-parameters

Data Inputs: X, Y

$\hat{\mathbf{\Sigma}}^{(0)} \leftarrow \mathbf{\Sigma}_0$; $ll \leftarrow 0$; $l \leftarrow 0$;

repeat

  E-step: Given $\hat{\Sigma}^{(l)}$, compute conditional posterior distribution $q_\theta^{(l)} = p(\theta \mid \hat{\Sigma}^{(l)})$ from Equations (5.11) and (5.12).;

  M-step: Calculate $\hat{\Sigma}^{(l+1)} = \left(\hat{\Sigma}_\eta^{(l+1)}, \hat{\Sigma}_\tau^{(l+1)}, (\hat{\sigma}^2)^{(l+1)}\right)$ via Equations (5.15), (5.16), and (5.17) ;

  Compute log-likelihood $ll' = \mathcal{L}(\hat{\Sigma}^{(l+1)})$;

  $ll \leftarrow ll'$; $l \leftarrow l + 1$;

until ll *converges*;

return $\hat{\Sigma} = \left(\hat{\Sigma}_u, \hat{\Sigma}_\tau, \hat{\sigma}^2\right)$

D.2 Hyper-Parameters Updating Schema

We use the annealing optimization schema for the parameters updating. Starting from trial $n = 0$, we will update the hyper-parameters after every $\{t_0, t_1, t_2, \cdots, \}$ trials, where $t_0 = 1$, and

$$\frac{t_k}{t_{k-1}} = 1 + \alpha \log(1 + k). \tag{D.1}$$

In the above equation, we set the learning rate $\alpha = 0.02$. Thus, the updating schema is given by $\mathcal{H} = \{0, t_0, t_0 + t_1, \cdots, \}$.