TOWARDS A POST-TAYLORISM THEORY OF CROWDSOURCING

by

ZHOU JUNJIE
(BMgmt, Zhejiang University)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF INFORMATION SYSTEMS AND ANALYTICS

NATIONAL UNIVERSITY OF SINGAPORE

2025

Supervisor:
Professor Hahn Jungpil

Examiners:
Associate Professor Tan Chuan Hoo
Assistant Professor Lim Shi Ying

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Zhou Junjie

Zhou Junjie

November 2025

Acknowledgments

This dissertation could not have been completed without the help and support of many people. I take this opportunity to express my gratitude to them.

I am eternally indebted to my advisor, Jungpil Hahn, who started me on this journey and has shaped how I approach research. I hope I can one day fill your shoes.

Special thanks to my committee members Chuan Hoo Tan and Shi Ying Lim for their insightful feedback at every stage of this thesis and for being there when I needed help. Many thanks also go to the brilliant minds I've had the pleasure of working with. David Lehman, Lusi Yang, Gwanhoo Lee, Junyi Li, Yifei Wei, Christoph Mueller-Bloch, and Jonas Andersen, for inspiring me to grow as a professional.

Thanks also to my friends and colleagues in the Department of Information Systems and Analytics for fostering a collegial and enjoyable research environment. I would also like to thank Yaxi Jin for her generous support during the early years of my doctoral journey. I am also grateful to Jinya Hu for her continuous encouragement and support.

Last but not least, I would like to thank my family. My father Longyun Zhou, my mother Yamei Ding, and my sister Ying Zhou, who stood by me all the way with unconditional love and support.

Contents

Abstract

Crowdsourcing is rapidly becoming a mainstream innovation tool for firms. However, despite its wide adoption, crowdsourcing often fails to deliver the desired outcomes in practice. This dissertation seeks to examine how to fully leverage the innovativeness of crowdsourcing. I argue that existing studies and theories of crowdsourcing have been deeply rooted in the management paradigm of Taylorism and overlook the roles of knowledge (re)combination, contextual complexity, and measures of crowdsourcing effectiveness. While this theoretical perspective is valid in certain contexts, it might be ill-suited to guide the use of crowdsourcing for complex problem-solving.

This dissertation consists of two essays that, combined, aim to develop a post-Taylorism theoretical perspective of crowdsourcing that incorporates these important but neglected elements. The first essay entitled "Making the Crowd Wiser: (Re)-combination through Teaming in Crowdsourcing" examines the (re)combination among solution solvers (i.e., the crowd) via teaming – i.e., team formation on the fly. Although teams, when properly structured and composed, are expected to generate higher quality solutions compared to individual solvers, simply allowing teaming in crowdsourcing may not necessarily enhance global crowdsourcing effectiveness as team formation during crowdsourcing would shift the distribution of quality of all discovered solutions. I develop an agent-based simulation model, calibrated by empirical data from an online crowdsourcing platform, to examine whether and to what extent teaming could enhance the likelihood of finding high quality solutions for the solution seeker. I present a theoretical framework that identifies the underlying mechanisms of how teaming influences overall crowdsourcing effectiveness. The framework is used to explore the impact of teaming in generalized crowdsourcing environments.

The second essay entitled "The Cave of Crowdsourcing: A Systems Investigation of Crowdsourcing Effectiveness" focuses on the knowledge re(combination) between solution solvers (i.e., the crowd) and the solution seeker (i.e., the firm). To utilize crowdsourcing beyond its existing knowledge base, the firm must evaluate and select from the crowd's solutions. However, this evaluation can be biased by the firm's existing understanding. I conceptualize crowdsourcing as a system constrained by the seeker firm's existing knowledge – the system is incentivized to produce

satisfactory solutions for the firm but (potentially) ineffective solutions for the crowdsourced problem. I extend the agent-based simulation model and examine how the knowledge flows between the solution solvers and the seeker firm within the crowdsourcing system influences the system's performance in discovering solutions with absolute high quality. I identify two paradoxes in crowdsourcing. First, the crowdsourcing system is most likely to overlook high quality solutions developed by the solvers when they are most needed. Second, when the seeker firm provides feedback, although the seeker firm perceives the adopted solution as of higher quality, the absolute performance may be poor. To resolve these paradoxes, I examine how the seeker firm could leverage the wisdom of the crowd and learn from the crowd about what design choices constitute a good solution.

Taken together, these two essays contribute to the literature on crowdsourcing by proposing and developing a post-Taylorism perspective on crowdsourcing. The overall contributions are twofold. First, this dissertation enhances our understanding of how two distinct types of knowledge (re)combination influence crowdsourcing outcomes. Second, this dissertation proposes a methodology for measuring and comparing global crowdsourcing effectiveness across varying contingencies.

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background and Motivation

Since the inception of the open innovation model (Chesbrough 2003), crowdsourcing, where firms broadcast their internal problems to external solvers for innovative solutions, has enabled firms to harness the minds of more people than they can employ in their internal R&D function. Firms have developed their own crowdsourcing initiatives (Lykourentzou et al. 2021, Piezunka and Dahlander 2015) and partnered with dedicated crowdsourcing platforms (Lee et al. 2018, Jiang et al. 2021) to support innovation activities.

However, despite such wide adoption, recent industry analyses suggest that crowdsourcing might be problematic in solving complex problems, the raison d'être of innovation. For example, during the Mexico Gulf crisis, BP hosted a crowdsourcing campaign to collect solutions for oil spills. The crowdsourcing campaign ended up with "a lot of effort for little result" (Goldenberg 2011). Another interview with NASA's Center of Excellence for Collaborative Innovation also revealed that most of NASA's attempts to seek solutions from crowdsourced communities failed (Nunley 2020).

This thesis is concerned with how to utilize crowdsourcing for innovative solutions to complex problems. Before presenting the specific research questions, this thesis further proposes that one of the major obstacles to using crowdsourcing for solving complex problems is the earmarks of the management paradigm of Taylorism implicit in existing studies and theory of crowdsourcing.

1.2 How Taylorism Limits Crowdsourcing

Taylorism refers to the management paradigm that focuses on increasing workers' work efficiency through specialization in designated and repetitive tasks. This idea is deeply ingrained in the contemporary theory of crowdsourcing. First, the dominant theoretical perspective of crowdsourcing, the parallel path effect (Boudreau et al. 2011, Boudreau and Lakhani 2013, Terwiesch and Xu 2008), is consistent with Taylorism's assumption of specialization. The parallel path effect suggests there exists substantial uncertainty in approaching and solving a crowdsourced problem. As a result, having a large number of members in the crowd with specialized expertise from diverse domains increases the likelihood of matching the problem with the required knowledge and skills. While Taylorism allows minimal interaction among workers to prevent social loafing, the parallel path effect also implies the search process should be as independent as possible since interactions among crowds may merge separate searches for solutions and consequently hinder problem-solving by reducing the number of independent solutions.

Second, the choices of research context of prior crowdsourcing studies have reflected Taylorism's assumption of repetitive work. Despite the claim that crowdsourcing is one of the most widely adopted paradigms of the open innovation model (Chesbrough 2003), crowdsourcing research hasn't really aimed at addressing crowdsourcing effectiveness for complex problems that require innovative solutions. Many studies have focused on relatively simple tasks such as logo design (Jiang et al. 2021, 2020, Jian et al. 2019), image selection (Koh and Cheung 2022), naming (Koh 2019) and language translation (Liu et al. 2014).

Last, the main variables of interest in contemporary crowdsourcing studies align with Taylorsm's focus on efficiency. Prior research has extensively investigated various ways to encourage participation and contribution from the crowd within a crowdsourcing campaign. Different incentive mechanisms (Yang et al. 2009, Jian et al. 2019) and design elements (Jiang et al. 2021, 2020, Nagaraj 2021, Koh and Cheung 2022, Manshadi and Rodilitz 2022) have been shown to be associated with the extent to which the crowd will put effort into crowdsourcing. Another stream of literature on crowdsourcing has attempted to study crowdsourcing with formal economic models (Terwiesch and Xu 2008, Mihm and Schlapp 2019, Korpeoglu et al. 2020, Hu and Wang 2020). The fundamental equilibrium for the crowd is to maximize the likelihood of obtaining rewards with minimal effort. These studies have implied that crowds' dedication can guarantee crowdsourcing effectiveness.

Taken together, the theoretical foundation, the contexts studied, and the variables of interest in crowdsourcing research have been deeply rooted in Taylorism. However, as firms adopt crowdsourcing for complex problems that require innovative solutions, implicit assumptions consistent with Taylorism, albeit valid, might not be suitable as theoretical guides to fully utilize the innovativeness of crowdsourcing. First, given the multidisciplinary nature of innovation, innovative solutions will likely involve interdependent pieces of knowledge from disparate domains. A single specialized agent (e.g., a firm or a solver of the crowd) might not have the complete set of knowledge and skills to tackle all of the interdependencies and independently develop an innovative solution. There is still a paucity of theory that underpins the efficient (re)combination of expertise during crowdsourcing. Second, the shift from utilizing crowdsourcing for relatively simple tasks to complex problems raises the importance of including problem complexity and interdependence as essential contingencies for crowdsourcing theory. It is not clear whether conclusions drawn from the study of simple tasks will also apply to complex problems. Last, merely increasing the efficiency of the crowd (e.g., encouraging participation and contribution) might not guarantee crowdsourcing effectiveness for complex problems. Crowdsourcing effectiveness is jointly determined by the crowd's endeavor and expertise. While the former decides to what extent the crowd's expertise has been harnessed, the latter decides the quality of the solutions when expertise from the crowd is fully exploited. As crowdsourced problems become more complicated, expertise possessed by specialized agents might not be sufficient for innovative solutions. In other words, even if the crowd participants have dedicated themselves to crowdsourcing, the developed solutions may still be of low quality when expertise has become the primary constraint. An important but somewhat overlooked aspect in the extant theoretical perspectives of crowdsourcing is how to enhance crowdsourcing effectiveness to complement the view of efficiency for complex problems.

Given these limitations, the next section will provide an suggestive outline of a post-Taylorism theory of crowdsourcing.

1.3 Towards a Post-Taylorism Theory of Crowdsourcing

In developing a post-Taylorism theory of crowdsourcing, this thesis draws on the concept of search on rugged landscapes to lend structure and coherence. The broad intellectual tradition of landscape search spans biology (Kauffman 1993), economic sociology (Gavetti et al. 2017, Levinthal 1997), and technology and innovation management (Fleming and Sorenson 2003). From organisms to organizations, landscape search explains how various entities (e.g., species, organization forms, technology, and products) evolve through adaptation – an entity is characterized as consisting of some attributes (e.g., genome, organization elements, product features). Each location on the landscape represents a particular configuration of these attributes and is associated with a certain fitness level to the environment. Entities adapt to different locations on the landscape by changing the configurations of their attributes. Depending on the magnitude of change, entities' adaptation is restricted within areas with different proximity. Kauffman (1993) showed that the fitness of mapping an entity to its environment depends on not only the attributes of the entity independently but also on how those attributes interact with each other. When there are few interdependencies among attributes, entities can evolve on a smooth surface, but as the number of interdependencies increases, entities must adapt on a rugged landscape with multiple summits and valleys. This thesis instantiates entities as solutions / innovations to complex problems, so the metaphor of landscape search provides an apt theoretical framework to investigate complex adaptive problem-solving (Afuah and Tucci 2012).

Consider a firm that navigates the rugged landscape in search of a solution to a problem (i.e., seeks the solution with highest fitness on the landscape). It cannot search all possible alternatives and immediately move to the highest location on the landscape due to limited information and resources. Rather, the firm is constrained by bounded rationality (Cyert and March 1963). Alternatives within the firm's neighboring areas are incremental changes from the firm's current location and are thus relatively more predictable. The firm searches its neighborhood on the landscape in a "satisficing" manner – whenever an alternative from a neighboring area is perceived to be better than the current position, the firm moves to that location, and so forth.

The firm's search process on the landscape is deeply shaped by its knowledge about the landscape. The firm's knowledge base and capability determine what attributes it can change and as a result, which areas on the landscape it can explore. Additionally, the firm has its unique cognitive frame of problem-solving according to its knowledge base (Gavetti and Levinthal 2000, Afuah and Tucci 2012). A firm's cognitive frame determines how it perceives local alternatives and consequently affects the firm's search trajectory by modifying every satisficing decision. If the firm's knowledge and capability hinder the firm from exploring outside its neighborhood, or the cognitive frame does not capture the nuance within the firm's neighborhood, the firm might not be able to approach positions with high fitness on its own.

Now consider a focal firm that attempts to search the landscape via crowdsourcing and broadcasts its problem to the crowd – a group of agents (other firms, consulting companies, individual solvers). Agents from the crowd are with their diverse knowledge bases and unique cognitive frames. Consequently, agents will search different areas with various search trajectories. The focal firm only needs to motivate agents such that the landscape is explored as much as possible and then evaluate all discovered locations at the end of crowdsourcing. In this way, crowdsourcing enables the firm to effectively conduct distant search on the landscape and approach areas that were out of reach via merely internal search (Afuah and Tucci 2012).

However, as discussed earlier, any single specialized agent may be unable to approach and search areas of high fitness due to limited knowledge and imperfect cognitive frame. This is especially so when the landscape is rugged, and the interdependencies among attributes are unbeknownst to agents. Although crowdsourcing enables distant search, it does not necessarily increase the likelihood of finding the global optimum of the landscape if little distant search is conducted in areas with high fitness. Little is known about how to navigate agents to areas beyond their local knowledge bases via (re)combination to increase the likelihood of finding the global optima on rugged landscapes. The following sections use the metaphor of landscape search to outline these key elements of a post-Taylorism theory of crowdsourcing.

1.3.1 (Re)combination of Expertise

Although the agents in crowdsourcing search different areas on the landscape, it is not guaranteed that their explored areas are indeed areas with high fitness. The firm may manage to recruit a great number of agents but meanwhile, gain little exposure into innovative solutions because the recruited agents cannot approach areas of high fitness due to insufficient knowledge and skills. The efficient (re)combination of expertise possessed by different agents may expand agents' search areas by allowing for modifying more attributes and for inducing new search trajectories from the integration of different cognitive frames. Therefore, a post-Taylorism theory of crowdsourcing should complement the parallel path perspective with a (re)combination perspective to enable crowdsourcing to identify solutions that cannot be discovered merely by parallel search.

Several nuances emanate from the interplay between the parallel and (re)combination perspectives. First, the (re)combination perspective can enrich the parallel path perspective only if there is ex-ante independent search before the combination. (Re)combinations before crowdsourcing only change the unit of analysis in the framework of the parallel path effect. For example, if several individual solvers form a team and join a crowdsourcing contest, the specialized team would be the basic agent that searches a particular area on the landscape according to the portfolio of team members' expertise. Hence, a post-Taylorism theory of crowdsourcing should focus on in-progress (re)combinations. Second, the (re)combination of expertise inherently weakens parallel search and the extent to which the parallel path effect could be leveraged. (Re)combination merge separate search areas and trajectories and as a result, fewer independent attempts are made by the same number of agents. A post-Taylorism theory of crowdsourcing should put the same emphasis on both the opportunities and obstacles of (re)combination.

1.3.2 Complexity as Contingency

Parallel search and (re)combination may lead to different outcomes across varying degrees of the ruggedness of the landscape. When attributes of the solutions are relatively independent, a solution to the problem can be found by configuring each attribute separately. Put simply, agents search on a landscape with few peaks that are reachable as long as agents choose a direction with a positive gradient during the search process. Conversely, for complex problems with interconnected attributes, the problem-solving process may get trapped where no attributes can be further modified as it can reduce the value of other configured attributes due to their interdependencies. Agents search on a rugged landscape with numerous local peaks. Search along the directions with positive gradients may lead agents to one of the local peaks, but not necessarily the global peak. Whenever agents reach a local peak on the landscape, they can only cease the search process since they cannot find a better alternative in the immediate neighborhood. Given the fundamental differences between search on smooth and rugged landscapes, a post-Taylorism theory of crowdsourcing should take a contingency view concerning complexity and underpin the utilization of crowdsourcing for complex problems to complement the extant theory that develops from simple problems.

1.3.3 Crowdsourcing Effectiveness

One last element of a post-Taylorism theory of crowdsourcing is the shift of focus from efficiency to effectiveness.

Although prior theoretical perspectives on crowdsourcing have mainly focused on whether agents from the crowd search their respective areas as extensively as possible, the outcomes of extensive search are constrained by agents' expertise and cognitive frames of problem-solving. When agents' expertise is only a subset of all required knowledge and skills, or their cognitive frames are inaccurate representations of the reality, the best possible solutions discovered via parallel search may not actually have high fitness. In other words, agents' search efficiency may not necessarily be positively associated with crowdsourcing effectiveness. While the mainstream efficiency-oriented theory attempts to utilize crowdsourcing to its full potential, a post-Taylorism theory should focus on how to improve the potential of crowdsourcing.

There are some hurdles to developing an effectiveness-oriented theory. First and foremost, there aren't any appropriate and observable counterfactuals of crowdsourcing outcomes. Even if the firm has adopted an innovative solution via crowdsourcing, it may not be the most effective solution that could have been developed from the same crowd with the same levels of efficiency, especially when the search process entails the impacts of both parallel search and (re)combination. Second, the firm's solution adoption in crowdsourcing per se does not necessarily imply effectiveness. Consider the insufficient expertise and inaccurate cognitive frames that drive the firm from internal R&D to crowdsourcing for distant search. It is not guaranteed whether the firm can choose the most effective solution from crowdsourcing. Developing a post-Taylorism theory requires additional methodological consideration to overcome these problems.

Given these aforementioned elements, this thesis develops a post-Taylorism theoretical perspective of crowdsourcing through simulation methods.

1.4 Research Focus and Contributions

This thesis is comprised of two essays that combined, aim to develop a post-Taylorism theoretical perspective of crowdsourcing that improves crowdsourcing effectiveness with a focus on the (re)combination of expertise across varying levels of problem complexity. This thesis uses a simulation-based method that enables creative experimentation to produce novel theory (Davis et al. 2007). Specifically, this thesis first formalizes the conceptual landscape of problem-solving with the *NK* fitness landscapes model (Kauffman 1993) and then models the solution seeker (i.e., the firm) and solution solvers (i.e., the crowd) as agents with different search areas and cognitive frames due to limited knowledge. This thesis runs a set of simulation experiments to unravel the impact of (re)combination on the likelihood of finding the global optima of landscapes with varying levels of ruggedness. Essay I, and essay II focuses on the (re)combination among solution solvers (i.e., the crowd), and the (re)combination between solution solvers (i.e., the crowd) and the solution seeker (i.e., the firm), respectively.

Taken together, my dissertation offers several implications for research and practice. First, the two essays extend the literature on crowdsourcing by providing insights into how knowledge (re)combination influences crowdsourcing effectiveness. Given the predominance of focus on the parallel path effect in the existing crowdsourcing literature, they have the potential to enrich our understanding by highlighting alternative mechanisms through which crowdsourcing can generate value. Second, the research design in my dissertation capture global crowdsourcing effectiveness that has been largely overlooked due to the limitations of archival data. My dissertation, therefore, extends the literature by foregrounding this core construct. Third, this dissertation establishes problem complexity as a boundary condition and therefore contribute to a contingency view of crowdsourcing.

For practical implications, this dissertation offers insight into how firms can more effectively leverage the innovativeness of crowdsourcing. The findings help firms decide whether to adopt a crowdsourcing approach should be adopted and if so, how to proactively intervene in the process to steer the massive parallel search. This dissertation also provides guidance for crowdsourcing platforms on how to improve their design to better support the solution seeker firm in discovering high quality solutions.

1.5 Thesis Organization

This opening chapter has provided an overview of the study context and motivations based on the theoretical gaps. It highlights the importance of a post-Taylorism theory to fully utilize the innovativeness of crowdsourcing. Chapter 2 describes Essay I in detail. It first reviews theoretical foundations of (re)combination via teaming in crowdsourcing and then presents how different theoretical constructs are formalized into different components of a computational model. The computational model is calibrated using data from an online crowdsourcing platform and is used to derive theoretical mechanisms. The identified theoretical mechanisms are then applied to interpret the impact of teaming in generalized crowdsourcing environments. Implications to both research and practice are then elaborated. Chapter 3 describes Essay II in detail. It begins by reviewing and conceptualizing crowdsourcing as a system constrained by the seeker firm's existing knowledge. The identified theoretical foundations are translated into a computation model. Four experiments that capture different knowledge flows between the solution seeker and solvers within the crowdsourcing system are conducted. Implications to both research and practice are then provided. I conclude this thesis in Chapter 4.

1.6 Abstracts of the Two Essays

This section presents the abstract of the two essays.

1.6.1 Essay 1: Making the Crowd Wiser: (Re)combination through Teaming in Crowdsourcing

Firms have widely adopted crowdsourcing because of its advantages in conducting parallel search by recruiting a large number of solvers to generate innovative solutions to challenging problems. Formerly independent solvers may join forces to collaboratively develop solutions via teams to co-create solutions that would not have emerged from any parallel yet independent path alone. However, when teams are formed on the fly during crowdsourcing (teaming), they inevitably alter the distribution of the quality of all solutions via reducing the number of independently developed solutions. There is scant research examining whether and to what extent teaming could compensate for the loss in the parallel search and eventually increase the likelihood of firms obtaining high quality solutions, namely crowdsourcing effectiveness. In this paper, we posit that solvers are likely to make use of a variety of publicly available information that crowdsourcing platforms provide as quality signals in determining potential teammates for teaming. In turn, this reliance on observable quality signals will shape the self-selected teaming process and subsequent crowdsourcing effectiveness. Using simulation experiments, we find that the impact of teaming on crowdsourcing originates from the immediate returns from identifying other solvers (or their solutions) to integrate with and potential returns from teamwork-based collaboration, albeit conditionally depending on problem complexity and the timing of teaming. Moreover, under certain conditions, teaming may not compensate for the loss in parallel search and will hamper global crowdsourcing effectiveness. The findings shed new light on crowdsourcing effectiveness and design and point to exciting new lines of inquiry.

1.6.2 Essay 2: The Cave of Crowdsourcing: A Systems Investigation of Crowdsourcing Effectiveness

Firms increasingly adopt crowdsourcing to solve complex problems by leveraging external solvers. However, as solvers with diverse knowledge partaking in crowdsourcing, the firm's existing knowledge may constrain its ability to evaluate and steer solvers' parallel search. This study conceptualizes crowdsourcing as a system constrained by the firm's knowledge, where solvers' efforts may shift from effective problem solving to merely fitting the firm's existing understanding. We examine the ability of the crowdsourcing system to obtain high quality solutions, namely crowdsourcing effectiveness, by analyzing the interactions between the firm and solvers. Using simulation experiments, we find two paradoxes. First, when crowdsourcing is most needed to develop solutions beyond the firm's knowledge, the crowdsourcing system is also most likely to overlook high quality solutions. Second, feedback from the firm may make the firm more satisfied at the cost of worsening crowdsourcing effectiveness, leading to a system collapse. To resolve these paradoxes, we examine how the firm might learn from the wisdom of the crowd. We find that learning from the crowd enhances crowdsourcing effectiveness, contingent on the firm's knowledge level and problem complexity. These findings shed new light on crowdsourcing effectiveness and design and point to exciting new lines of inquiry.

Chapter 2 Making the Crowd Wiser: (Re)-combination through Teaming in Crowd-sourcing

2.1 Introduction

With the advancement of information technology, firms, as solution seekers, are able to post their problems on crowdsourcing platforms for external solvers from all over the world to develop innovative solutions (Malhotra et al. 2021, Boudreau and Lakhani 2013). The effectiveness of this crowdsourcing approach lies in the ability to discover the extreme-value outcomes from the underlying distribution of possible quality of outcomes (Boudreau et al. 2011, Loch et al. 2001). Complex problems entail considerable uncertainty in what knowledge is required to solve them, and as a result, engaging a crowd of solvers with as diverse knowledge as possible addresses this uncertainty by increasing the likelihood that at least one participant having the required knowledge will partake in the problem solving and find an extreme-value solution. This idea, termed the "parallel path effect", is deeply ingrained in the contemporary crowdsourcing literature, which has focused on identifying ways to encourage participation so as to facilitate greater parallel independent search (e.g., Jian et al. 2019, Jiang et al. 2020, Lysyakov and Viswanathan 2023).

Despite the predominance of focus on the parallel path effect in the existing crowdsourcing literature, a growing body of research has recognized that this perspective overlooks the co-creation of solutions via interactions among solvers (Majchrzak and Malhotra 2013). Formerly independent solvers may join forces to collaboratively develop solutions or share intermediate outputs for others to build upon during crowdsourcing to co-create solutions that would not have emerged from any parallel yet independent path alone (Jin et al. 2021). While prior work has shown that such co-creation can occur among relative strangers temporarily organized into teams (Majchrzak et al. 2012), the crowdsourcing literature has largely treated teams as pre-formed problem-solving units (e.g., Cao et al. 2022, Dissanayake et al. 2015), without considering the possibility that solvers can form teams on the fly to co-create innovative solutions. (Henceforth, we refer to *teaming*, the process through which a group of strangers are temporarily assembled into a multidisciplinary team for a task as per Edmundson (2012), to describe team formation during ongoing crowdsourcing contests.) Compared to pre-formed teams, teaming offers a more flexible approach to team formation by allowing greater variation in how diverse knowledge from different solvers is combined during crowdsourcing and as a result, has the potential to better address the inherent uncertainty in solving a crowdsourced problem.

Paradoxically, although teams, when properly structured and composed, are expected to generate higher quality solutions compared to individual solvers (Cao et al. 2022, Riedl and Woolley 2017, Dissanayake et al. 2015, Javadi and Hirschheim 2021), simply allowing teaming in crowdsourcing may not necessarily enhance global crowdsourcing effectiveness. With teaming, previously independent solvers work together and submit only an integrated solution to the solution seeker for final selection. The number of developed solutions submitted by the crowd at the end of crowdsourcing will thus be reduced. As such, teaming, on the one hand, adversely affects crowdsourcing effectiveness by diminishing the number of parallel paths, as fewer solutions are ultimately discovered. On the other hand, teaming may (or may not) favorably affect crowdsourcing effectiveness by increasing the quality of the discovered solutions, albeit fewer in numbers. Whether the positive effect of teaming on individual-level solution quality materializes into overall crowdsourcing effectiveness depends on how the distribution of quality of all discovered solutions shifts due to teaming and the associated reduction of parallel paths. For example, the solutions discovered by the crowd with teaming may be of lesser variety leading to marginal (or no) gains to the likelihood of discovering the best possible (or a set of very high quality) solution(s) – the extreme-value outcome that solution seekers desire. The impact of teaming on global crowdsourcing effectiveness can therefore be much more complex than their immediate impact on the quality of individual solutions. Even if the average quality or performance of individual solutions has been enhanced through teaming, overall crowdsourcing effectiveness may still be hindered if the extreme-value solutions are insufficiently explored by the crowd. Whilst prior research, likely due to limitations of archival data, has primarily focused on the average effect of forming teams on the quality or performance of individual solutions (e.g., Cao et al. 2022, Dissanayake et al. 2015), we still lack a nuanced understanding of how teaming affect global crowdsourcing effectiveness.

Moreover, unlike in organizations where managers can compel how individuals are coordinated from a global view, teaming in crowdsourcing takes place locally in a self-selection manner in which solvers are motivated to choose other solvers with whom to team up with the goal of enhancing the performance of their own solution so as to increase their chances of winning the crowdsourcing contest (Riedl and Seidel 2018, Majchrzak et al. 2021); and as such individual solvers are not preoccupied with enhancing global crowdsourcing effectiveness. For instance, the whole population of solvers tends to converge around similar ideas rather than exploring novel ones in response to the seeker's selection criteria (Park et al. 2022). In addition, because most solvers in online crowdsourcing are unfamiliar with each other, they primarily rely on publicly available information to gauge the quality of other solvers/solutions to team up with (Riedl and Seidel 2018, Park et al. 2022). Depending on how numerous solvers are guided by different quality signals in selecting targets for teaming to different extents, the impacts of teaming on global crowdsourcing effectiveness may also vary.

Considering these gaps of the extant crowdsourcing literature, there is a need for further theoretical development and analysis to deepen our understanding of how team formation during crowdsourcing among formerly independent solvers impact global crowdsourcing effectiveness. We adopt a simulation-based approach to answer our research questions. The simulation-based approach enables us to uncover second-order outcomes that emerge from first-order interactions among agents (Benbya et al. 2020, Miranda et al. 2022, Nan and Tanriverdi 2017). The simulation model we develop explicitly captures how solvers decide whether to and with whom to team up under the influence of publicly available quality signals; and how solvers search for solutions independently (before teaming) and jointly (after teaming). We calibrate the simulation parameters governing solvers' teaming decisions based on empirical data from Kaggle[1] and then conduct counterfactual simulations to compare global crowdsourcing effectiveness with and without teaming.[2] Our simulation results suggest that the impact of teaming on crowdsourcing originates from the immediate returns from identifying high quality solution to integrate with and potential returns from teamwork-based collaboration, albeit conditionally depending on problem complexity and the timing of teaming. Moreover, under certain conditions, teaming may not make up for the loss in parallel search and will hamper global crowdsourcing effectiveness.

Our findings offer a range of theoretical insights. First, to our knowledge, this is one of the first studies to investigate the impact of interactions that occur during crowdsourcing on global crowdsourcing effectiveness – the most relevant outcome metric for solution seekers that reflects the power of crowdsourcing. However, prior crowdsourcing studies have thus far investigated the quality of individual solutions which cannot fully portray global crowdsourcing effectiveness. Second, this study contributes to the crowdsourcing literature by highlighting the interplay between the parallel path effect and teaming as a potential mechanism underlying crowdsourcing effectiveness. Third, it contributes to the research stream on crowdsourcing platform design by highlighting the impact that information provided on crowdsourcing platforms may have on teaming decisions and subsequently on crowdsourcing effectiveness.

2.2 Theoretical Foundations

2.2.1 Crowdsourcing: Parallel Landscape Search

The creation of novel technologies, products and services all involve addressing complex problems that require configuring a large number of interdependent decisions (Simon 1962). The rugged landscape metaphor of the *NK* fitness landscapes model (Kauffman 1993) has been a useful theoretical apparatus to understand complex

[1] `www.kaggle.com`

[2] Even though empirical data from crowdsourcing platforms such as Kaggle are available, our research questions cannot be answered with an empirical approach. A limitation of empirical analysis in this context is that it only allows direct observations of team performance (i.e., the quality of team-based solutions) instead of global crowdsourcing effectiveness (e.g., whether or not the set of extremely good solutions was discovered).

problem solving. Prior work on problem solving as landscape search has shown that effective problem solving is stymied by the ruggedness of the landscape (i.e., the complexity of the solution space) – as the solution space becomes more complex, problem solving is likely to end up with suboptimal solutions where no improvement can be made by merely considering incremental alternatives. It has also been shown that the ability to explore many different parts of the problem's solution landscape yields the greatest chance of finding the optimal solution, especially for complex problems (Fleming and Sorenson 2003).

Crowdsourcing offers a tenable approach to finding high quality (and potentially optimal) solutions by enabling multiple parallel searches of the solution landscape on a larger scale compared to internal R&D (Afuah and Tucci 2012). It is fundamentally a process where innovative solutions are discovered through extensive parallel experimentation. The diversity in knowledge about the problem distinguishes the problem solvers' initial positions on the solution landscape and consequently influences their subsequent search trajectories, enabling the crowd of solvers to explore the landscape more extensively than a single solution seeker could through serial trial and error search (e.g., through internal R&D). At the end of crowdsourcing, the solution seeker collects all of the solutions submitted by solvers seeking to win the contest. As a result, all solutions discovered by individual solvers are made available to the solution seeker. In this way, the solution seeker is able to conduct parallel distant search on the rugged landscape and may find high-quality solutions at a lower cost compared to internal R&D. This is essentially the mechanism of the "parallel path effect" (Boudreau et al. 2011) where increasing the number of crowdsourcing participants leads to more exploration of all possible parts of the landscape and as a result, enhances crowdsourcing effectiveness in terms of the likelihood of finding extremely good solutions to the seeker's problem.

As such, research in crowdsourcing has thus far emphasized the importance of crowd participation to better leverage the parallel path effect. For example, the reward structure (Yang et al. 2009, Liu et al. 2014), project characteristics (Yang et al. 2010), problem specification (Jiang et al. 2020), feedback on progress (Jiang et al. 2021, Sanyal and Ye 2023), prize guarantees (Jian et al. 2019), the number of live crowdsourcing campaigns (Mo et al. 2021), the use of AI assistants (Lysyakov and Viswanathan 2023), and the timing of participation (Lee and Mishra

2025) have been studied in terms of their impact on solver participation. This perspective, albeit valid, neglects the fact that merely enhancing parallel search by increasing participation levels may fail to ensure a thorough exploration the problem's solution landscape. As the crowdsourced problem becomes broad in scope and requires the integration of knowledge from many different domains, there may be no individual solver who has knowledge of all aspects of the problem (Gurca et al. 2023, Majchrzak et al. 2021, Majchrzak and Malhotra 2020). Imperfect knowledge possessed by solvers can hamper the parallel path effect from two aspects. First, it limits solvers' search space on the landscape and hinders solvers from developing complete solutions beyond the scope of their individual knowledge. Only solvers who happen to be initially located in the vicinity of optimal solutions might be able to discover high quality solutions to the crowdsourced problem (Afuah and Tucci 2012). For example, Lifshitz-Assaf (2018) found that only 3 of 14 R&D problems crowdsourced by NASA have been completely resolved. Moreover, solvers are expected to build a mental representation of the problem based on its description and requirement and then develop solutions based on this mental representation (Sanyal and Ye 2023). Imperfect knowledge, however, hinders solvers from fully understanding the problem and thus leads to incomplete mental representations of the problem. Incomplete representations, in turn, will misguide how solvers explore the landscape.

In sum, although crowdsourcing enables the solution seeker to conduct parallel search on the problem's solution landscape, the effectiveness of such parallel search for complex problems is limited by the knowledge of individual solvers. This limitation underscores the need for knowledge combination across different solvers.

2.2.2 Teaming: The Joining of Parallel Search Paths

Several recent studies have begun to investigate the combination of knowledge via teams in crowdsourcing. The performance of a team is associated with the diverse yet complementary skills and knowledge possessed by the individuals in the team (Zenger and Lawrence 1989, Devine and Philips 2001) and the process through which different skills and knowledge are combined (Wegner 1987, Marks et al. 2001, Faraj and Sproull 2000). As such, the crowdsourcing literature has mainly focused on member composition in terms of skill levels and diversity (Dissanayake et al. 2019, Cao et al. 2022, Riedl and Woolley 2017) and on the collaboration process among team members (Dissanayake et al. 2015, Riedl and Woolley 2017, Javadi and Hirschheim 2021). When a team is formed with a particular composition and operates in a collaborative manner, it is expected that the team will be able to overcome the limitation of individual knowledge and outperform individual solvers in jointly searching along a single path on the problem's solution landscape. Similarly to individual solvers, attracting as many teams as possible can strengthen parallel search, especially if teams generally outperform individual solvers. With prior collaboration experience, solvers sometimes pre-form teams and participate in crowdsourcing contests as a team. A pre-formed team, on the one hand, combines diverse knowledge possessed by different members. On the other hand, it retains the same team composition which makes them relatively rigid in the face of complex crowdsourcing problems with variations in knowledge requirements. As the domain of the crowdsourcing problems changes, it is reasonable to expect that a pre-formed team is also less likely to solve them completely.

In contrast, teaming, whereby solvers initially participate in the crowdsourcing contest as independent problem-solving units and then identify potential teammates on the fly to solve the crowdsourcing problem collectively, provides a more flexible approach to team formation in crowdsourcing. Simply put, solvers initially engaged in independent search paths are essentially joining forces to jointly search the problem's solution landscape. Teaming allows crowdsourcing to flexibly leverage the whole crowd instead of being constrained to a smaller group of mutually familiar solvers and thus has greater potential to overcome the limitations of individual knowledge and enhance global crowdsourcing effectiveness by increasing the chances that some spontaneous combination of knowledge will fit the problem at hand. However, whenever previously independent solvers team up during a crowdsourcing contest, they integrate independently developed in-progress solutions and will submit only one joint solution to the solution seeker for selection at the end of crowdsourcing. Even though the quality of the solutions may be higher due to teaming, the number of solutions discovered and submitted by the crowd will inevitably be reduced. The more extensive teaming takes place among solvers, the greater this reduction will be. Moreover, forming teams on the fly during ongoing contests can be much more complicated compared to in conventional (organizational) contexts, where team formation can be conceptualized as a deliberate decision (Owens et al. 1998) in which diverse factors such as homophily (Gompers et al. 2017, Ruef et al. 2003), diversity (Reagans and Zuckerman 2001) and pre-existing relationships (Ruef et al. 2003, Reagans et al. 2004) are carefully considered by managers to strategically optimize team composition and process and thus maximize team performance (Hamman and Martínez-Carrasco 2023, Millhiser et al. 2011). In the crowdsourcing context, there is no manager who can orchestrate team formation with a global view to enhance overall crowdsourcing effectiveness, but rather, team formation occurs organically via self-selection into teams. Individual solvers select possible teammates from strangers with the expectation that together, the team will be able to solve the given problem more effectively compared to solving it independently as individuals.

The self-selection nature of teaming introduces great uncertainty because most solvers in real crowdsourcing contests are unfamiliar with each other due to limited past collaboration experiences. It is thus difficult for solvers to accurately evaluate the inherent capabilities of numerous strangers and as a result, solvers face challenges in estimating the likelihood of team success. Moreover, teaming can happen at any stage of crowdsourcing. Similar to those studied in innovative settings, solvers in crowdsourcing could form teams as nominal groups where they can start by working independently and later might decide to team up to collaborate (Girotra et al. 2010). The discretionary timing of teaming further exacerbates the uncertainty of teaming as forming teams with the same targets may induce different performance at different stages of crowdsourcing (i.e., when the individuals' search paths join). The prior literature suggests that individuals in online crowd spaces will make use of publicly available information to alleviate the uncertainty in their self-selection process (Hahn et al. 2008, Riedl and Seidel 2018). In the context of crowdsourcing, solvers can also rely on information that signal different aspects of potential teammates to narrow down the scope of consideration and gauge the potential benefits of teaming. What remains unknown, however, is how self-selected teaming affects global crowdsourcing effectiveness.

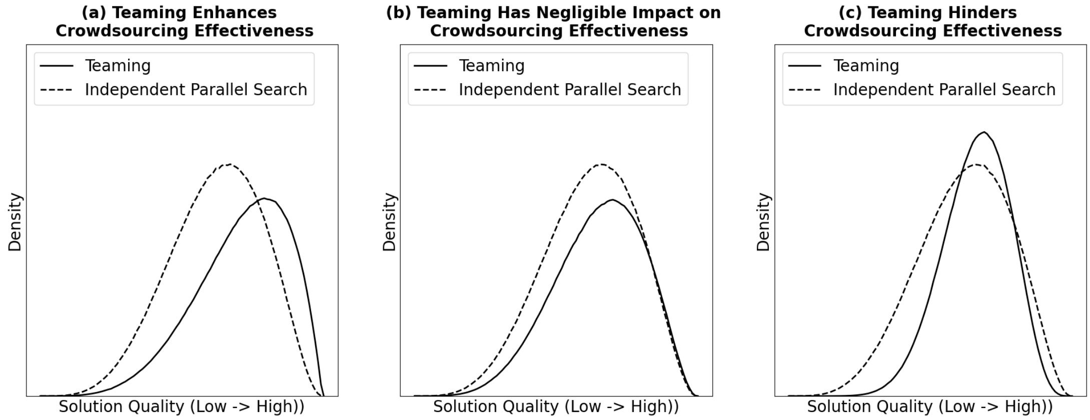

Figure 2.1: How Teaming Affects Crowdsourcing Effectiveness

2.2.3 Crowdsourcing Effectiveness: The Interplay between the Parallel Path Effect and Teaming

Although teams may generate higher quality solutions that their individual counterparts with imperfect knowledge cannot achieve, team formation via teaming inherently reduces parallel search because fewer solutions will eventually be submitted to the seeker. The prior literature on crowdsourcing, however, has mainly focused on the former at the expense of the latter. For instance, solution sharing and learning within teams and crowdsourcing communities (Jin et al. 2021, Javadi and Hirschheim 2021, Lysyakov and Viswanathan 2020, Riedl and Seidel 2018) can increase scores as well as the likelihood of winning for the submitted solutions. But a reduction in parallel search may actually render crowdsourcing less effective overall for the solution seeker.

Not considering the reduction of parallel search is critical because global crowdsourcing effectiveness depends on which solvers (and their solutions) are involved in teaming. As mentioned earlier, solvers self-select their potential teammates based on publicly available information that signal solvers' inherent qualities. Consequently, solvers who are considered superior based on these signals are more likely to be targets of teaming. However, solvers with favorable signals are also likely to be high performers themselves. Integrating solutions possessed by these solvers via teaming may yield marginal additional benefits since their solutions are already well-developed. Moreover, the prior literature on collective problem-solving has suggested an exploration-exploitation trade-off when independent solvers work together (Bernstein et al. 2018). That is, although the quality of individual solutions can be generally increased through intensive exploitation of existing knowledge via interactions among solvers (such as teaming), the best solution (or the set of very good solutions), which is fundamentally the seeker's motivation for hosting a crowdsourcing contest, may decrease in quality due to insufficient independent exploration of novel solutions (Lazer and Friedman 2007, Mason and Watts 2012). For example, when a solver with a solution that is promising yet underdeveloped forms a team with another, its efforts may be redirected to the other's inferior solutions that generate positive feedback in the short run. Depending on which solvers are involved in teaming, whether the subsequent teamwork can balance exploration and exploitation to compensate for the loss of parallel search remains uncertain.

To illustrate the interplay between the parallel path effect and teaming, Figure 2.1 presents three plausible scenarios illustrating instances where teaming enhances (2.1a), has no impact on (2.1b), or diminishes (2.1c) global crowdsourcing effectiveness. Each figure displays possible (and likely) distributions of solution quality via mere parallel search (i.e., when there is no teaming; see the dotted lines) and with teaming (see the solid lines).[3] Figure 2.1a presents a desirable case where high quality solutions are discovered by solvers through teaming (see the right area), for example, when solvers in the left tail complement imperfect knowledge via teaming and jointly develop comprehensive solutions that any individual solvers cannot configure. In contrast, Figure 2.1b shows a status-quo scenario where teaming only reduces the number of solutions and generates marginal benefits on high-quality solutions (see the right area). This could happen when only high performers with similar knowledge integrate well-developed solutions via teaming. Figure 2.1c indicates a counterproductive case where although the average quality of solutions increases compared to independent parallel search, the number of high-quality solutions decreases (see the right area). As discussed before, solvers who could have developed high-quality solutions may be deviated away from exploring their solutions after teaming and as a result, reduce the number of high quality solutions.

In sum, although teaming, by allowing collaboration, may increase the average

[3] In generating these figures, we ensured that density of the distribution for teaming is smaller than that of parallel search, indicating the reduction in the number of parallel trial and error.

quality of individual solutions (i.e., shift the mean of the distribution rightwards in Figure 2.1), its impact on global crowdsourcing effectiveness still depends on teaming can compensate for the reduction of parallel search.

2.3 Empirical Evidence

To highlight the prevalence of teaming, we show suggestive empirical evidence from crowdsourcing competitions on Kaggle. Kaggle is the world's largest crowdsourcing platform for complex machine learning problems. Kaggle allows solvers to form teams with other unfamiliar solvers on the fly during crowdsourcing competitions, making it a good fit for observing real-world teaming behaviors. One of the most notable pieces of information provided by Kaggle is a public contest leaderboard that records the contemporaneous performance of all solvers (both individuals and teams) during the contests (Lee et al. 2018). We use leaderboards on Kaggle to support our argument that solvers will likely be influenced by publicly available quality signals in their teaming decisions. We tracked how teaming takes place according to the public leaderboards on a daily basis for 63 live contests on Kaggle from May 2020 through January 2022. As discussed earlier, this study focuses on teaming – team formation within competitions and as a result, we exclude those solvers who start the contest as part of teams and only include solvers who have worked independently and then formed teams with others. Our data includes a total of 4,116 solvers who joined teams via teaming.

The solid lines in Figures 2.2a and 2.2b show the cumulative distribution of teaming probability across teammates' performance difference in rank percentile and the probability of teaming across solvers' performance in rank percentile at the timing of teaming, respectively. (We will discuss the dotted lines representing teams in our simulations later.) There are two noteworthy observations. First, we find that solvers were more likely to form teams with others with very similar performance (i.e., those with smaller differences in rank percentiles; see Figure 2.2a). Around 70% of teams are formed by individual solvers who have a performance difference smaller than 20% in rank percentile. Second, we find that solvers with higher ranks were more likely to form teams (around 0.20 density), whereas individuals with lower ranks were less likely to form teams (the density declines to 0; see Figure

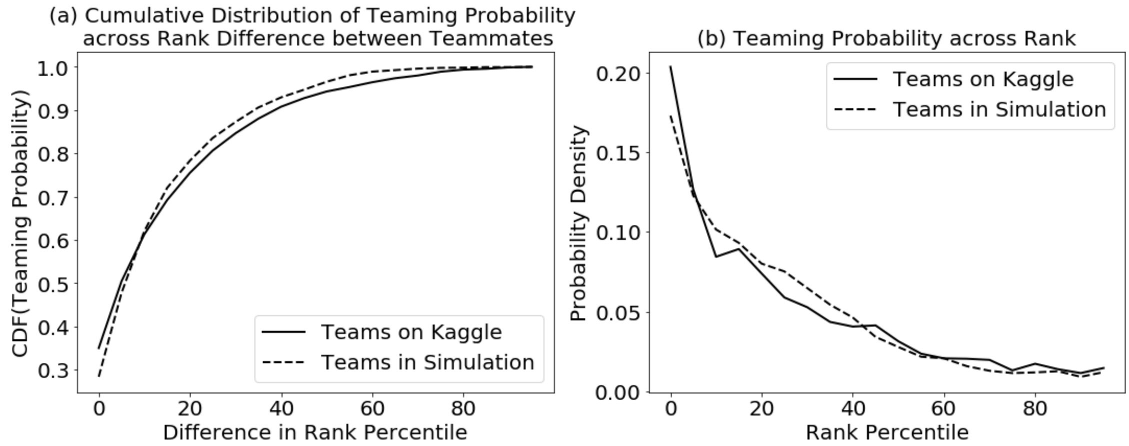

Figure 2.2: Teaming on Kaggle

2.2b). These results suggest that teaming on the fly during competitions is not random but correlated with solvers' contemporaneous performance. These results are also aligned with expectations – when publicly available information, such as contemporaneous performance on the crowdsourced problem as displayed on the contest leaderboard, is used as quality signals in teaming decisions, solvers with favorable signals (e.g., high ranking on the leaderboard) are more likely to form teams. Additionally, considering the mutual selection among solvers, teams are typically formed by solvers with comparable qualities (e.g., similar performance on the leaderboard).

Our analysis of the empirical observations from Kaggle contests shows evidence, albeit model-free, that teaming decisions are not entirely random but seem to be driven by strategic considerations of information that is publicly available on crowdsourcing platforms. This reliance on quality signals can determine whether and with whom solvers will form teams during the competition and eventually overall crowdsourcing effectiveness (as illustrated in Figure 2.1). However, empirical data fails to capture the counterfactual loss in parallel search and the gains to global crowdsourcing effectiveness. Crowdsourcing effectiveness may not necessarily be enhanced even if teams outperform individual solvers after teaming because those solvers who form teams could have found (a greater number of) high-quality solutions independently; and the population of teamed-up solvers may fail to discover novel solutions. Moreover, the empirical evidence only reflects teaming behaviors within a specific context. Crowdsourcing solvers may exhibit different teaming patterns on other platforms. Although the leaderboard is one type of information featured on crowdsourcing platforms that could potentially guide solvers' teaming decisions, solvers may rely on other signals about potential teammates. There are also possible boundary conditions that shift the tension between teaming and parallel search on affecting crowdsourcing effectiveness. Therefore, we adopt a simulation-based approach that enables us to conduct counterfactual experiments with and without teaming across various contingencies.

2.4 Simulation Model

As discussed earlier, crowdsourcing can be conceptualized as the parallel search on a problem's solution landscape (Afuah and Tucci 2012). Therefore, the *NK* fitness landscapes model (Kauffman 1993, Levinthal 1997), which models problem solving as adaption on rugged landscapes, is a suitable analytical apparatus for our study (Davis et al. 2007).

The classical *NK* fitness landscapes model creates a multidimensional space (i.e., the fitness landscape) consisting of N elements and K interdependencies among these elements. Each configuration of N elements $\mathbf{d} = \langle d_1, d_2, \ldots, d_N \rangle$ is associated with a fitness value $F(\mathbf{d})$. Each element d_i contributes to the overall fitness value by c_i, the value of which depends not only on the specification of the focal element d_i but also on K other interdependent elements (i.e., $c_i = c_i(d_i|k \text{ other } d_j\text{'s})$). Formally, c_i is randomly drawn from a uniform distribution $U(0, 1)$. The overall fitness value $F(\mathbf{d})$ of a solution $\mathbf{d}$ can be calculated as the average of the fitness contributions c_i across every element d_i (i.e., $F(\mathbf{d}) = \frac{1}{N} \sum_i c_i$; $0 \leq F(\mathbf{d}) \leq 1$). By adjusting the parameter K for a given N, the *NK* fitness landscapes model enables the researcher to tune the complexity of the problem (i.e., the ruggedness of the landscape). When there are few interdependencies among elements (i.e., low K), the generated landscapes are smoother, whereas if elements are tightly coupled with each other (i.e., high K), the landscape becomes more rugged and the search for the global optimum becomes more difficult because of the existence of multiple peaks and valleys on the rugged landscape (see Appendix A.1.1 for technical details concerning the generation of rugged landscapes.

We extend the classical *NK* fitness landscapes model to better reflect the crowdsourcing context as follows. First, the crowdsourced problem is modeled as a complex solution-performance space where many solvers search for solutions in parallel. However, individual solvers' search and evaluation of solutions are constrained by their diverse yet incomplete knowledge. During search, solvers can team up with others to collaboratively search for a joint solution.

2.4.1 Crowdsourced Problem

We model a crowdsourced problem as consisting of N elements $d_1, d_2, \ldots, d_N$ that need to be configured by solvers, where each element d_i can take the value of 0 or 1. The d_i's represent knowledge-related decisions. For instance, in algorithm contests, they can include the choice of the analytical method (e.g., deep learning-based model vs. causality model) and the choice of computational architecture (e.g., local vs. cloud clusters); and in logo design contests, they can include the usage of color palettes (e.g., warm vs. cool colors) and different design approaches (e.g., minimalist vs. maximalist), etc. Formally, a crowdsourcing problem is represented by a space consisting of 2^N possible solutions, where each solution is an N-bit configuration: $\mathbf{d} = \langle d_1, d_2, \ldots, d_N \rangle$. Each location in the space (i.e., one particular solution configuration) is associated with a fitness value representing the quality/performance of the configured solution. The elements can be interdependent such that the fitness contribution of a particular configuration choice for an element could also depend on the configuration of K other elements. For example, in algorithm design contests, using cloud-based computation may be more beneficial when a deep learning model is adopted because of the scalability of cloud computing. Similarly, warm colors may generate better performance if a minimalism style is chosen in the logo design context.

2.4.2 Solvers

2.4.2.1 Decision Set

Solvers are heterogeneous in the knowledge they possess. In algorithm design contests, some solvers are proficient in technical domains, whereas some are more skilled at other contextual aspects such as business, clinical medicine, or computational biology. In logo design contests, some solvers are experts in visual communication, whereas some are more skillful in written communications such as copywriting of slogans. To incorporate knowledge heterogeneity, we model solvers as agents who are capable of configuring only a subset of the elements of $\mathbf{d}$. Formally, each agent holds a decision set D, where $D \subset \{1, 2, \ldots, N\}$ and $|D| = N_D$. Elements in D index the d_i elements in the knowledge domain of each agent and thus those that the agent can configure. Therefore, an agent can only search among 2^{N_D} possible solutions on the landscape that correspond with its knowledge. The indices for each agent are selected at random in order to reflect the diversity of knowledge across heterogeneous solvers. More specifically, at the beginning of the simulation (i.e., each simulated crowdsourcing contest), we randomly assign each agent N_D decisions and initialize them as 0 or 1 with equal probability. For decisions outside of the agent's knowledge domain (i.e., the $N - N_D$ decisions), global default values are set. These default values are identical for all agents who lack knowledge on the respective elements. Formally, agents' solutions can be represented as $\mathbf{d} = \langle \mathbf{d}_D, \mathbf{d}_{D^C} \rangle$, where subscript D signifies that d_i is from decision set D and can be configured during search, whereas subscript D^C signifies that d_i is outside D (i.e., C here represents that d_i is in the complement of set D, or $N \backslash D$) and cannot be modified by the agent.

2.4.2.2 Independent Search

We model crowdsourcing solvers' search behaviors as an incremental and iterative adaptation process. Before teaming, all agents search the landscape independently. During each time step, agents randomly select one decision from their decision set D and change their values from 0 to 1 or vice versa to consider an alternative configuration $\mathbf{d}'$.[4] However, due to limited knowledge (i.e., $|\boldsymbol{D}| < N$), agents cannot perfectly assess the value of decision configurations but are only able to evaluate them locally given other unknown knowledge domains. In other words, elements outside of an agent's knowledge (i.e., D^C) form a black box that cannot be directly observed. In algorithm contests, a technical expert can perceive the difference (e.g., efficiency and accuracy) among alternative algorithms but may fail to identify the contextual factors that impact the feasibility and suitability of the chosen algorithm. Similarly,

[4]We later relax this assumption and allow agents to select multiple decisions simultaneously.

in logo design contests, a visual communication expert can make choices on color and graphics but may not be able to evaluate the effectiveness of different slogan copies. Prior research has modeled such bounded rationality as cognitive abstraction (Gavetti and Levinthal 2000, Hahn and Lee 2021) where the cognitive landscape is usually a smoother version of the actual landscape and fitness values generally represent how well corresponding configurations work across unknowns. For all 2^{N_D} possible solutions, the perceived fitness values $\tilde{F}(\mathbf{d})$ can thus be calculated as the average fitness values across all possible configurations of the elements outside of the agents' decision spaces. Formally, $\tilde{F}(\mathbf{d}) = \mathbf{E}[F(\mathbf{d})|\mathbf{d}_{D^C}]$, where $\mathbf{E}[\cdot]$ represents the expected fitness across unknown decisions.[5] The perceived fitness values of different solutions may be higher or lower than the actual fitness values due to imperfect knowledge. Yet, how the agent perceives all solution configurations on the landscape as a whole would be largely correlated with the actual fitness landscape. For the current configuration $\mathbf{d}$ and proposed alternative $\mathbf{d}'$, where the Hamming distance between $\mathbf{d}$ and $\mathbf{d}'$ is 1, agents would choose the new configuration $\mathbf{d}'$ if $\tilde{F}(\mathbf{d}') > \tilde{F}(\mathbf{d})$. Conversely, if the above condition is not satisfied, agents would keep the current configuration and conduct another round of search.

2.4.2.3 Teaming

The teaming process is modeled as follows. At the time of teaming, each agent has a probability (p_{teamup}) to initiate a team-up invitation. If an agent u decides to initiate a team-up invitation, it sorts other agents according to the publicly available information on which it relies as signals and considers possible teammates one by one. The invitation process stops if an agent on the ordered list accepts the invitation or if no agents accept the invitation. At the same time, the recipient v of the team-up invitation also sorts the inviting agent u according to the signal information and have a probability (p_{accept}) to accept the team-up invitation. We model p_{teamup} as a constant such that each agent has the same tendency to initiate

[5] For example, suppose $N = 6$ and $N_D = 4$, if an agent is only knowledgeable about the first four knowledge elements (i.e., $D = \{1, 2, 3, 4\}$), then one of its solutions could be represented as $\mathbf{d} = \langle 1, 0, 1, 0, *, * \rangle$, where $*$ represents unknown elements (i.e., $D^C = \{5, 6\}$). The perceived fitness value $\tilde{F}(\mathbf{d})$ of solution $\mathbf{d}$ can be calculated as the average fitness value across unknowns. Namely, the average fitness value of $\mathbf{d} = \langle 1, 0, 1, 0, \mathit{0}, \mathit{0} \rangle$, $\mathbf{d} = \langle 1, 0, 1, 0, \mathit{0}, \mathit{1} \rangle$, $\mathbf{d} = \langle 1, 0, 1, 0, \mathit{1}, \mathit{0} \rangle$, and $\mathbf{d} = \langle 1, 0, 1, 0, \mathit{1}, \mathit{1} \rangle$. Also see Appendix A.1.2 for how this computational approach models the limitation of agents' imperfect knowledge.

teaming invitations and p_{accept} as a decreasing function of the relative order of the sender agent so those with favorable signals (e.g., good performance on the public leaderboard) are more likely to be accepted by peers. Whenever two agents team up, they are removed from subsequent matches, ensuring that each agent can only be part of one team. [6] Our modeling choice reflects how teaming takes place in real crowdsourcing contests. On some crowdsourcing platforms (e.g., Kaggle), solvers can indicate their willingness to team up in contests and as a result, may receive direct messages from other solvers to form a team. Other platforms (e.g., TopCoder and Crowdspring), despite not having explicitly implemented features to indicate teaming willingness, also provide communication channels that make the initiation and acceptance of team-up invitations possible. As noted, solvers on these platforms do not have to form teams prior to joining a contest but can team up with other solvers at any time during the contest. In our simulation experiments, we calibrate p_{accept} and p_{teamup} according to the empirical evidence from Kaggle in §2.3 (i.e., Figure 2.2).

2.4.2.4 Joint Search

Once a team is formed, agents in the team stick together for the remainder of the contest. One final necessary component of our model is how agents jointly search for a solution after teaming. Team members will gradually become aware of their teammates' knowledge and construct a shared understanding of the problem (Faraj and Sproull 2000, Wegner 1987, Robert et al. 2008, 2018). For example, in algorithm contests, a technical expert is likely to be aware of the contextual factors that determine the feasibility of algorithms if it teams up with a domain specialist. Formally, at each time step, an agent u has a probability of p_{learn} to augment its evaluation of alternatives $\mathbf{d}$ by incorporating one knowledge element from its teammate v's knowledge such that $\tilde{F}_u(\mathbf{d}) = \mathbf{E}[F(\mathbf{d})|\mathbf{d}_{D_u^C - \{d_j\}}]$, where $j \in D_v$.[7] Agents' evaluation of solution configurations gradually becomes more coherent and

[6] Whereas we assume teams are formed by two agents in the main model, we later relax this assumption in our sensitivity analyses and allow agents to form teams of varying sizes.

[7] For example, suppose $N = 6$ and $N_D = 4$, if an agent u is only knowledgeable about the first four knowledge elements (i.e., $D_u = \{1, 2, 3, 4\}$), now suppose this agent learns the fifth element from its teammate v with $D_v = \{3, 4, 5, 6\}$ and will be aware of the configuration of the fifth element thereafter. If the agent v's current specification of the fifth element is 1, then the solution becomes $\mathbf{d} = \langle 1, 0, 1, 0, 1, * \rangle$ for agent u and correspondingly, the perceived fitness value can be calculated as the average fitness value of $\mathbf{d} = \langle 1, 0, 1, 0, 1, \mathit{0} \rangle$ and $\mathbf{d} = \langle 1, 0, 1, 0, 1, \mathit{1} \rangle$.

as a result, their search behaviors become more aligned. Whenever an alternative is proposed by a team member according the search procedure described in §2.4.2.2, the proposal will be evaluated by the team collectively before they are accepted (Knudsen and Levinthal 2007, Csaszar and Eggers 2013, Christensen and Knudsen 2010). For example, in algorithm contests, a proposal to change the data structure from a technician needs to be evaluated by the domain expert in the team in terms of the appropriateness of the data structure in the real-life problem context. Unlike groups in traditional organizations where hierarchies play an important role in collaboration, self-organized teams in crowdsourcing are often small and relatively more decentralized. As such, following prior *NK* models on collaboration (Rivkin and Siggelkow 2003, Siggelkow and Rivkin 2005), a proposed alternative $\mathbf{d}'$ to $\mathbf{d}$ would only be accepted if $\tilde{F}_v(\mathbf{d}') \geq \tilde{F}_v(\mathbf{d})$ and $\tilde{F}_u(\mathbf{d}') \geq \tilde{F}_u(\mathbf{d})$ simultaneously – i.e., if both team members u and v perceive the alternative to be at least as good as the status quo.[8]

2.5 Simulation Results

We use the model described above to run a series of simulations that examine the impact of teaming on crowdsourcing effectiveness. For the purpose of model docking, we first validate our simulation model by comparing baseline results with prior findings in classical *NK* models and the crowdsourcing literature. The details are presented in Appendix A.1.5. We then explore the impact of teaming according to rank information on the public leaderboard (cf., our empirical observation from Kaggle) in a baseline model. We identify the underlying mechanisms that explain the findings observed in the baseline experiment. Finally, we apply the identified mechanisms to examine the impact of teaming on crowdsourcing effectiveness in broader crowdsourcing environments. Specifically, we explore the impact of two additional types of information that signal solvers' inherent qualities from the prior literature – i.e., knowledge diversity and problem-solving capability as well as different teaming tendencies.

For each simulation experiment, we adopt the following strategy. We fix $N = 12$

[8] See Appendix A.1.4 for a step by step illustration of how agents jointly search for a solution after teaming.

and vary K from 0 to 11 (i.e., the full range of complexity settings for $N = 12$). Our simulations converge after around 30 simulated time periods, so we set the earliest and latest timings of teaming from 0 to 30, in steps of 3. We keep the number of crowd participants at 400, the size of the agents' decision set N_D at 6, the learning rate among teammates p_{learn} at 0.10.[9] To ensure that the results of the simulation reflect the underlying structure of the model rather than a particular realization of a stochastic process, our results are based on 10,000 replications on independently generated random fitness landscapes. All results reported here are statistically significant with a 99% confidence level.

2.5.1 Baseline Model

We begin our analysis by examining whether teaming according to rank information (i.e., contemporaneous solution quality signal), as on Kaggle, can indeed enhance crowdsourcing effectiveness. In accordance with prior studies that use matching analysis to validate the results generated by simulation models (Butler et al. 2014, Johnson et al. 2014), we match p_{teamup} and p_{accept} to real-world observations such that the overall pattern of teaming generated by our simulations is consistent with empirical observations. We set p_{teamup} to be 0.20 and p_{accept} to be $0.5 \cdot (e^{-0.1x} + 0.001)$, where x is the performance rank percentile of the sender agent.[10] See dashed lines in Figures 2.2a and 2.2b for how our choice of p_{teamup} and p_{accept} approximate empirical observations from Kaggle.

2.5.1.1 Impact of Teaming according to Rank

To illustrate how teaming according to rank information affects global crowdsourcing effectiveness, Figure 2.3a shows the likelihood of finding high quality solution configurations on the landscape via crowdsourcing with and without teaming after 30 simulated time periods when problem complexity $K = 6$ and teaming timing $t = 15$, sorted by the absolute fitness value of each configuration in descending order. The x-axis in the figures capture the absolute rank of configurations on the landscapes. Here, rank refers to the rank of the solution configuration among all

[9] We relax these assumptions later in our sensitivity analysis.

[10] More specifically, we set p_{teamup} to be equal to the proportion of team observed on Kaggle and specify p_{accept} as an exponential function for which we conduct a grid search for hyperparameters. The details are discussed in Appendix A.1.3.

possible configurations on the fitness landscape, even those not submitted by the solvers, and not the rank of the solutions discovered by solvers. For a landscape with $N = 12$, there is a total of 4,096 ($= 2^{12}$) possible configurations. So if the winning solution (i.e., ranked 1^{st} among submitted solutions) has the 10^{th} highest fitness value among all 4,096 possible configurations, then its rank (on the x-axis of Figure 2.3a) would be 10. Although this solution has the best relative performance among all submitted solutions, it does not have the highest absolute quality among all possible configurations. And so to the solution seeker, this would not be ideal. The y-axis capture the likelihood that some solvers have submitted the x^{th} ranked configuration at the end of crowdsourcing. From Figure 2.3a, we find that teaming according to rank on the public leaderboard could increase the likelihood of finding high-quality solutions (see the solid line with cross markers). The likelihood of finding the best solution on landscapes increases around 5% compared to the case when all individual solvers search independently (when *Absolute Rank* $\rightarrow 1$).

Figures 2.3b, c, and d show the likelihood of finding the best solution (i.e., the leftmost point in Figure 2.3a) across timing of teaming for problems with low, medium, and high complexity, respectively. We find that teaming according to rank increases the likelihood of finding the best solution compared to merely individual search when problem complexity is low (see the solid line in Figure 2.3b). Interestingly, the relative advantage of teaming compared to individual search shows an inverted-U shape across timing of teaming. The peak performance is associated with timing around the 10_{th} period. As the timing approaches the end of the contest, the likelihood of finding the best solution decreases.

The above findings are moderated by problem complexity. First, as the level of problem complexity increases, the optimal timing of teaming shifts leftward (i.e., to earlier teaming timing). In problems with high complexity, the earlier the timing of teaming, the better the crowdsourcing effectiveness is (see Figure 2.3d). Moreover, teaming could be detrimental to crowdsourcing effectiveness compared to individual parallel search as problem complexity increases. In problems with medium complexity, teaming at the end of simulations would yield a comparable likelihood of finding the best solution as individual parallel search (see late timings in Figure 2.3c); in problems with high complexity, only teaming at early timings can result in higher likelihood of finding the best solution compared to independent

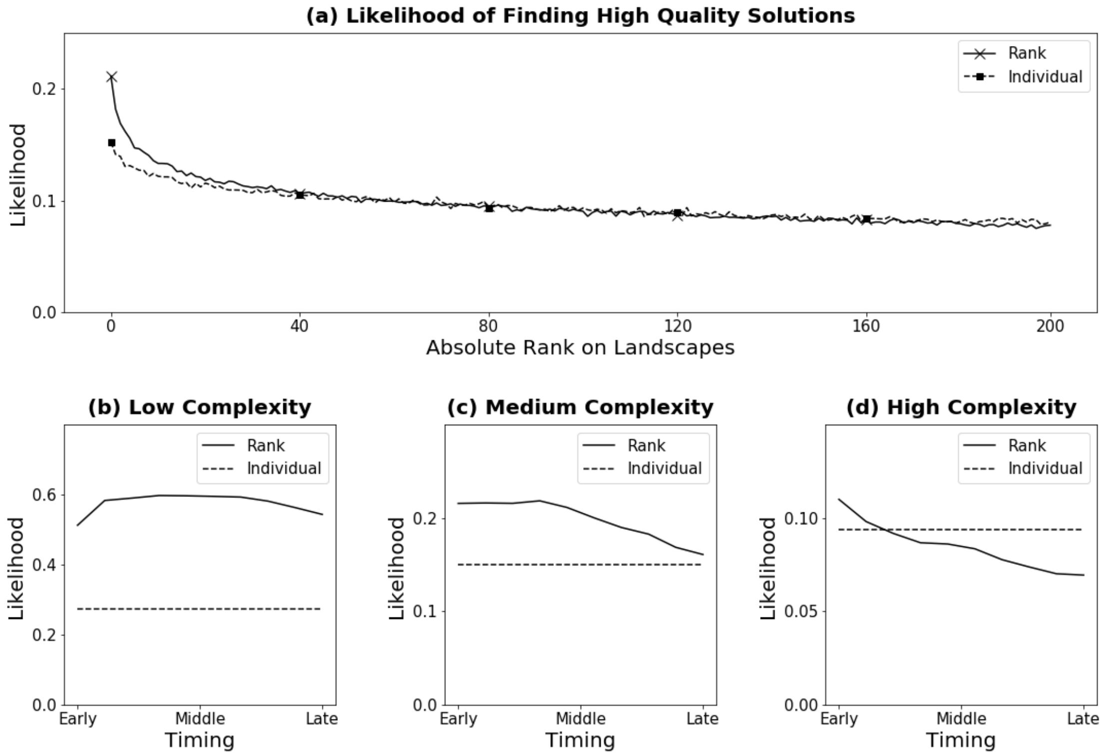

Figure 2.3: How Teaming According to Rank Affects Crowdsourcing Effectiveness

Notes: Problem complexity $K = 6$, teaming timing $t = 15$ in (a); for brevity, (b)-(d) only show the likelihood of finding the best solution for combinations of problem complexity (K) at low, medium and high levels (across columns of figures) and timing at early, middle and late stages (across the x-axis) – i.e., K=1, 6, 11 and T=3, 15, 27.

parallel search (see early timings in Figure 2.3d).

Taken together, the results show that the impact of teaming on crowdsourcing depends on the timing of teaming and problem complexity. If not well balanced, teaming can hamper global crowdsourcing effectiveness.

2.5.1.2 How the Impact of Teaming Unfolds

To better understand the patterns of the results presented above, we employ a two-step approach to uncover the underlying mechanisms that explain the impacts of teaming according to rank on crowdsourcing effectiveness. We first investigate the (dis)advantages of teaming with a counterfactual setup when solvers randomly team up with each other to rule out the impact of strategical usage of quality signals in the self-selected teaming process. We then analyze to what extent strategic teaming according to rank magnifies the (dis)advantages. When solvers form teams during crowdsourcing, parallel search is inherently weakened because fewer solutions will be submitted to the solution seeker. However, teaming may at the same time generate high-quality solutions that cannot be developed merely via individual search and make up for the loss in parallel search. The quality of solutions generated through teaming, on the one hand, depends on which solvers (and their respective solutions) are teamed up. Teaming up with solvers who possess the required knowledge about the crowdsourced problem or who have already developed more advanced solutions is likely to improve the outcomes of teaming. However, the crowdsourced problem usually entails great uncertainty regarding the knowledge and solution approaches required, making the identification of such solvers and solutions challenging in the early stages of crowdsourcing when solvers have yet to develop good solutions. Conversely, as the possible solution space is explored through large-scale parallel search and effective solution approaches gradually emerge over time, it becomes more likely for solvers to identify and select suitable teammates. On the other hand, the quality of solutions generated through teaming is determined by how previously independent solvers join forces after teaming. Solvers' independently developed solutions are integrated into a combined solution upon teaming, and the solvers then further develop the integrated solution collaboratively. The integration of separate solutions and co-development of the integrated solutions is influenced by the solvers (and their respective solutions) identified for teaming. The integration of two well-developed (underdeveloped) solutions is more (less) likely to enhance the quality of the integrated solution. Likewise, collaboration with solvers who possess the required knowledge tends to facilitate the co-development process. While the integration of separate solutions can occur immediately upon teaming, the co-development of the joint solution takes time. The advantages of teams over individual solvers, such as broader search scope and more comprehensive evaluation of solution configurations, only unfold over time. There is therefore an inherent trade-off between the mechanisms through which teaming generates high quality solutions – the identification of suitable teammates and the integration of their solutions would be more effective when teams are formed late, while the benefits of collaboration are more likely to manifest when teaming occurs early.

Figure 2.4a illustrates the possible mechanisms through which teaming generates high-quality solutions. The first mechanism involves identifying possible teammates

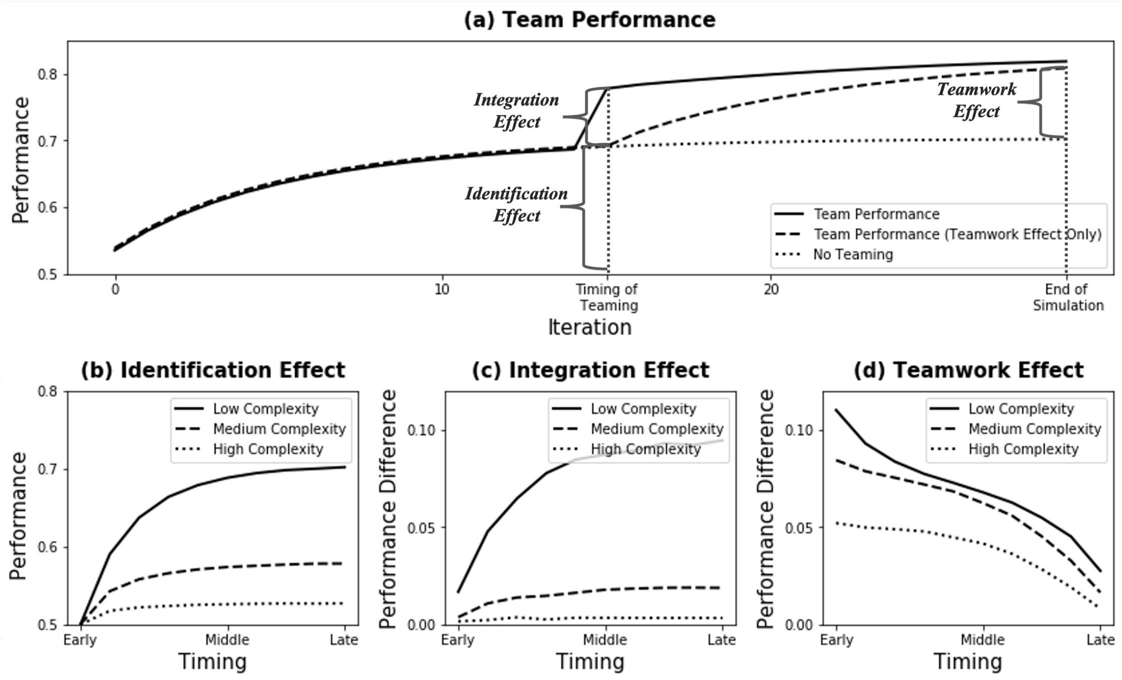

Figure 2.4: Performance Improvement from Teaming

to integrate solutions with at the time of teaming (hereafter *identification effect*). We capture it by assessing the extent to which high quality solvers (and their solutions) are selected for teaming. It is operationalized as the average performance of solvers who form teams. The second mechanism concerns the extent to which separate solutions are integrated into high quality ones right after teaming (hereafter *integration effect*). It is quantified as the performance difference between teams and individual solvers right after teaming. The last mechanism pertains to leveraging the benefit of teams after teaming (hereafter *teamwork effect*). It is captured by the performance difference between teams and individual solvers at the end of crowdsourcing.[11] Taken together, the *identification*, *integration*, and *teamwork effects* capture the extent to which teaming improves the quality of solutions developed by teams, albeit at the cost of reducing the number of solutions submitted (i.e., a reduction in the parallel-path effect).

[11]It is important to note that strong *identification* and *integration effects* would leave little room for the *teamwork effect* to manifest. In other words, if a team has already obtained a high-quality solution via identification and integration of high quality solutions, its potential to develop high-quality solutions via teamwork might be overlooked. As such, we consider a hypothetical scenario where solvers jointly develop a randomly given solution after teaming to capture the size of the *teamwork effect*.

Figures 2.4b, 2.4c and 2.4d show the size of the *identification*, *integration* and *teamwork effects* across different timings and different levels of complexity, respectively. Consistent with our conjecture, the *identification* and *integration effects* become stronger as the timing approaches the end of simulations, whilst the *teamwork effect* declines as teaming occurs later. Their sizes are moderated by problem complexity. As shown by the difference between the solid, dashed and dotted lines representing low, medium and high levels of complexity in Figures 2.4b and 2.4c, the *identification* and *integration effects*, albeit remain positive, decline significantly as problem complexity increases. When problem complexity is low, there are few interdependencies between design elements. As a result, solvers are more likely to configure the design elements within their respective decision sets to a high level of quality. The integration of two well-configured solutions will also likely produce a high-quality integrated solution. With more complex problems, however, individual solvers are less likely to independently develop high quality solutions. The integration of separate solutions also entails significant uncertainty as greater extents of interdependencies between separate solutions largely determine the quality of the integrated solution. It becomes more difficult to generate better solutions by simply integrating two disparate solutions because detailed tuning and integrated refinement are required to resolve the underlying interdependencies in the problem, which are very likely to be present across solvers' respective decision sets. Conversely, as shown by the difference between the solid and dotted lines in Figure 2.4d, although the *teamwork effect* also decreases as problem complexity increases, its reduction is less stark compared to the *identification* and *integration effects*. When problem complexity is high, conflicts arising from interdependencies across solvers' respective decision sets can be (partially) resolved through collaborative search. A previously overlooked design option may be triggered as a result of interacting with teammates' respective solution parts, and a previously misjudged solution can be more accurately evaluated by incorporating other solvers' knowledge. These changes can stimulate solvers to explore a wider range of alternative solutions and help solvers to escape local optima and reach configurations that cannot be attained without teamwork after teaming.

2.5.1.3 Regression Analyses

Next, we estimate how the proposed effects (i.e., identification, integration and teamwork effects), along with the parallel path effect, jointly affect crowdsourcing effectiveness. Specifically, we retrieve the *identification*, *integration* and *teamwork effects* using the previously discussed procedure from a dataset of 1,500,000 records (i.e., 5 different numbers of solvers (100, 200, 300, 400 and 500) across 10 different timings (T=0, 3, ..., 27) and 3 different levels of complexity (K=1, 6, 11) for 10,000 replications), each representing a randomly generated crowd of solvers. We capture the parallel path as the number of solutions submitted by the crowd at the end of crowdsourcing.

Table 2.1 includes the models used to examine global crowdsourcing effectiveness (i.e., whether the crowd of solvers in each replication has discovered the best solution on the landscape). In addition to the variables of interest, we also control the average performance of individual solvers (*IndividualPerformance*). Models A.1, B.1, and C.1 are baseline models that examine whether the crowd could have discovered the best solution without teaming for problems with different levels of complexity.[12] In line with expectations, parallel search is positively associated with crowdsourcing effectiveness (Model A.1, $\textit{ParallelPath} = 0.2147$, $p < 0.01$). As problem complexity increases, the parallel path effects become even more critical for improving global crowdsourcing effectiveness (Model C.1, $\textit{ParallelPath} = 0.2705$, $p < 0.01$). Models A.2-A.4, B.2-B.4, and C.2-C.4 examine whether teaming helps the crowd discover the best solution. These models differ from A.1, B.1, and C.1 by allowing teaming, which in turn leverages teaming effects but results in a lower number of solutions submitted. The dependent variable is a binary indicator that takes the value 1 if the crowd is unable to discovered the best solution via merely independent search (i.e., the dependent variable of Models A.1, B.1, and C.1 takes the value 0) but succeed in discovering it when teaming is allowed, and 0 otherwise.

Models A.2, B.2, and C.2 are baseline models that include only control variables and the parallel path effect. We find that the coefficients for *ParallelPath* remain positive, suggesting that with a greater number of solutions, teaming is more likely

[12]To retrieve the dependent variable, we ran a counterfactual experiment in which solvers started from the same initial positions on the same landscapes as in the teaming case but searched for the global peak in an independent manner.

Table 2.1: Regression Results

Panel A: *Low Complexity*	**Model A.1**	**Model A.2**	**Model A.3**	**Model A.4**
	DV: Whether the Best Solution Has Been Found via Independent Search	*DV: Whether Teaming Helps the Crowd Find the Best Solution*		
Constant	-10.6923^{***}	0.2352^{***}	-4.4880^{***}	-5.6081^{***}
	(0.03705)	(0.02758)	(0.05858)	(0.06174)
ParallelPath	0.2147^{***}	0.2400^{***}		0.2578^{***}
	(0.00252)	(0.00257)		(0.00263)
IdentificationEffect			-0.6207^{***}	-0.5388^{***}
			(0.08168)	(0.08344)
IntegrationEffect			12.4137^{***}	13.3400^{***}
			(0.11323)	(0.11740)
TeamworkEffect			10.9624^{***}	12.0053^{***}
			(0.15990)	(0.16770)
Timing		-0.0175^{***}	-0.0092^{***}	-0.0090^{***}
		(0.00038)	(0.00067)	(0.00068)
IndividualPerformance	12.5424^{***}	-2.5031^{***}	3.6204^{***}	3.9255^{***}
	(0.04786)	(0.03749)	(0.10653)	(0.10870)
Obs	500,000	500,000	500,000	500,000
Log Likelihood	-244792.30	-277229.20	-275171.92	-270242.00

Panel B: *Medium Complexity*	**Model B.1**	**Model B.2**	**Model B.3**	**Model B.4**
	DV: Whether the Best Solution Has Been Found via Independent Search	*DV: Whether Teaming Helps the Crowd Find the Best Solution*		
Constant	-6.9259^{***}	-2.3378^{***}	-3.9553^{***}	-5.0164^{***}
	(0.06519)	(0.08234)	(0.09382)	(0.09784)
ParallelPath	0.2460^{***}	0.2376^{***}		0.2562^{***}
	(0.00312)	(0.00437)		(0.00442)
IdentificationEffect			-0.7462^{***}	-0.8988^{***}
			(0.18893)	(0.20116)
IntegrationEffect			7.4958^{***}	8.5451^{***}
			(0.18396)	(0.19732)
TeamworkEffect			8.8960^{***}	10.2876^{***}
			(0.17336)	(0.18760)
Timing		-0.0211^{***}	-0.0009	0.0022^{**}
		(0.00064)	(0.00091)	(0.00094)
IndividualPerformance	7.2243^{***}	-1.1242^{***}	1.9676^{***}	2.4026^{***}
	(0.10890)	(0.13979)	(0.23611)	(0.24816)
Obs	500,000	500,000	500,000	500,000
Log Likelihood	-184466.49	-126387.67	-126468.44	-124737.02

Panel C: *High Complexity*	**Model C.1**	**Model C.2**	**Model C.3**	**Model C.4**
	DV: Whether the Best Solution Has Been Found via Independent Search	*DV: Whether Teaming Helps the Crowd Find the Best Solution*		
Constant	-4.6799^{***}	-4.1426^{***}	-4.3349^{***}	-5.2257^{***}
	(0.09921)	(0.14028)	(0.14294)	(0.14657)
ParallelPath	0.2705^{***}	0.2417^{***}		0.2567^{***}
	(0.00392)	(0.00609)		(0.00614)
IdentificationEffect			-1.9360^{***}	-2.3708^{***}
			(0.28478)	(0.31063)
IntegrationEffect			6.3134^{***}	7.4715^{***}
			(0.26750)	(0.29257)
TeamworkEffect			6.7798^{***}	8.0499^{***}
			(0.22027)	(0.24149)
Timing		-0.0115^{***}	-0.00002	0.0021^{**}
		(0.00089)	(0.00099)	(0.00101)
IndividualPerformance	2.4899^{***}	0.5356^{**}	3.2740^{***}	3.7964^{***}
	(0.18541)	(0.26217)	(0.36131)	(0.38343)
Obs	500,000	500,000	500,000	500,000
Log Likelihood	-132376.19	-75662.49	-75974.64	-75073.72

Notes: (1) We have varied the range of solutions of interest for the solution seeker from the 1^{st} to the top 10 and top 100 solutions, the results remained qualitatively the same. (2) *ParallelPath* is measured as the number of solutions submitted (in hundreds). (3) A logit model is used, the results remain qualitatively unchanged for linear probability and probit models.
Significance Levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

to help the crowd find the best solution. This is because, given the same teaming tendency, an increase in the number of solvers would increase the number of solutions submitted and at the same time, amplify the impact of teaming.

Models A.3, B.3, and C.3 examine the impact of teaming on crowdsourcing effectiveness. When problem complexity is low, the *integration* and *teamwork ef-*

fects are positive and significant (Model A.3, *IntegrationEffect* $= 13.3400$, $p < 0.01$; *TeamworkEffect* $= 12.0053$, $p < 0.01$), while the *identification effect* is negative (Model A.3, *IdentificationEffect* $= -0.5388$, $p < 0.01$). These results suggest that teaming can effectively compensate for the loss of parallel search by leveraging the *integration* and *teamwork effects*. However, identifying and including high quality solvers and their respective solutions in teaming to leverage the *integration* and *teamwork effects* may hamper crowdsourcing effectiveness. As complexity increases, the *identification effect* becomes increasingly negative (Model C.3, *IdentificationEffect* $= -2.3708$, $p < 0.01$), whilst the *integration* and *teamwork effects* becomes less effective in helping the crowd discover the best solution (Model C.3, *IntegrationEffect* $= 7.4715$, $p < 0.01$; *TeamworkEffect* $= 8.0499$, $p < 0.01$). We also find that as complexity level increases, the coefficient of the *teamwork effect* becomes relatively larger than that of the *integration effect*. We observe similar patterns when we consider the parallel path effect and the impact of teaming simultaneously (i.e., Models A.4, B.4, and C.4 in Table 2.1). Taken together, these results suggest that as problem complexity increases, teaming is less likely to compensate for the loss of parallel search.

From a global perspective, the *identification effect* involves the identification of effective approaches to high quality solutions. However, identifying such approaches alone is not sufficient to ensure crowdsourcing effectiveness, especially when problem complexity is high. On the one hand, there is greater uncertainty in which approaches will ultimately yield high-quality solutions. Focusing only on approaches that are developed from solvers' limited knowledge and appear promising at the moment may hinder the exploration of alternative paths. On the other hand, even if the identified approaches are indeed effective, their potential can be undermined if the interdependencies among design elements are not properly resolved in the integration and subsequent co-development of solutions. Conversely, the *integration* and *teamwork effects* enable the crowd to reach solutions that are otherwise inaccessible through independent search. By combining separate solutions, the crowd is able to generate solutions that cannot be developed by any single solver and therefore reach uncharted areas on the rugged landscape. In subsequent teamwork, solvers are more likely to explore alternative configurations of their respective solution parts in response to the changes made by their teammate(s), with each configuration more comprehensively

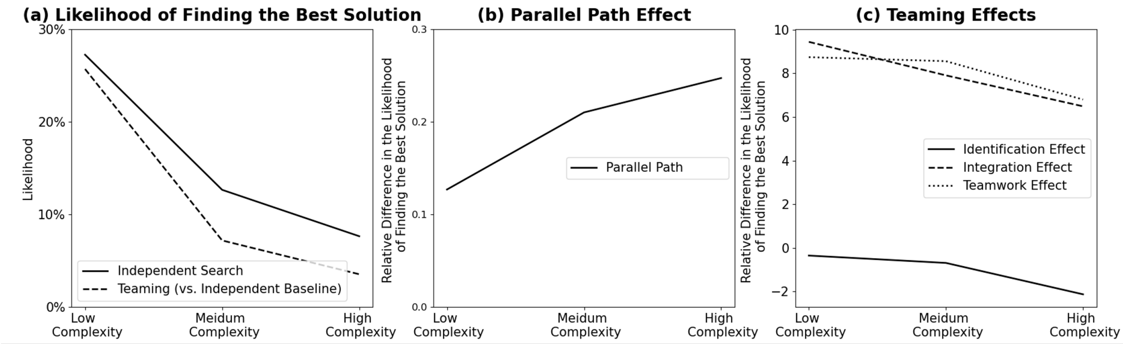

Figure 2.5: Marginal Effects of the Parallel Path, Identification, Integration, and Teamwork Effects

evaluated using both their own knowledge and that of their teammates. However, as the problem becomes more complex, unresolved interdependencies among design elements within solvers' decision sets increase. The extent to which high quality solutions can be developed from the newly explored positions (i.e., the integrated solutions) therefore depends increasingly on the subsequent teamwork to resolve these interdependencies.

Figure 2.5 offers a graphical depiction of these results. The solid and dashed lines in Figure 2.5a show the average likelihood of finding the best solution via merely independent search (i.e., when there is no teaming) and the likelihood that teaming improves the discovery of the best solution, respectively. Figure 2.5b shows the marginal effect of parallel paths. It is rescaled according to the baseline likelihood of finding the best solution (i.e., the solid line in Figure 2.5a). We find that the marginal effect of parallel paths increases with problem complexity. When problem complexity is low, an one unit increase in the number of parallel trials and errors (i.e., 100 more solutions) can increase the likelihood of finding the best solution by approximately 15% (i.e., from 28% to 28% $\times$ (1+15%)). In contrast, when problem complexity is high, it can increase the likelihood by approximately 25%. For problems with low complexity, the landscape is smooth with few peaks, making solvers' parallel search more likely to end up in the same peaks. As problem complexity increases, individual solvers are more likely to get stuck with suboptimal solutions as the landscape exhibits great ruggedness. As a result, more parallel search is needed in problems with high complexity such that some solvers might manage to discover a path toward the globally optimal solution.

Figure 2.5c shows the marginal effects of the *identification*, *integration*, and *teamwork effects*. When problem complexity is low, a 0.05-unit increase in the *identification effect* (e.g., from 0.5 to 0.55, as observed in the solid line in Figure 2.4b) reduces the likelihood that teaming compensates for the loss of parallel search (i.e., the dashed line in Figure 2.5a) by around 1.7%. The negative impact, however, can be offset by the benefits of the *integration* and *teamwork effects*. A 0.05-unit increase in the *integration* and *teamwork effects* (e.g., from 0.05 to 0.10, as observed in the solid lines in Figures 2.4c and 2.4d) can raise the likelihood by 47.2% and 43.7%, respectively. As problem complexity increases, however, the negative impact of the identification effect becomes more pronounced. A 0.05-unit increase in the *identification effect* reduces the likelihood that teaming compensates for the loss of parallel search by 10.6%, whilst a 0.05-unit increase in the *integration* and *teamwork effects* can raise the likelihood by 32.4% and 33.9%. Moreover, as shown in Figure 2.4, although the crowd can still achieve an *identification effect* above 0.5, attaining a 0.05-unit increase in the *integration* and *teamwork effects* becomes substantially more difficult in high complexity problems. In other words, it becomes less likely to leverage the benefits of the *integration* and *teamwork effects* to offset the negative impact of the *identification effect* and subsequently, compensate for the loss of parallel search and enhance crowdsourcing effectiveness. Taken together, these results suggest that as the loss of parallel search becomes more critical with increasing problem complexity, teaming can still counteract the loss through the *integration* and *teamwork effects*. However, it becomes increasingly difficult to offset the negative impact of the *identification effect*.

2.5.1.4 Revisiting the Baseline Results via the Identified Mechanisms

The impact of teaming according to rank can be explained by integrating how it emphasizes the *identification*, *integration* and *teamwork effects* and how different effects impact crowdsourcing effectiveness (i.e., Figure 2.5c). Figure 2.6 offers a graphical depiction of how teaming according to rank emphasizes different effects compared to randomly teaming up (i.e., Figures 2.4b-2.4d). We find that teaming according to rank (the gray bars in Figure 2.6) accentuates all three effects more significantly than the random baseline, especially when teams are formed late. On the one hand, rank information enables solvers to identify those with high quality

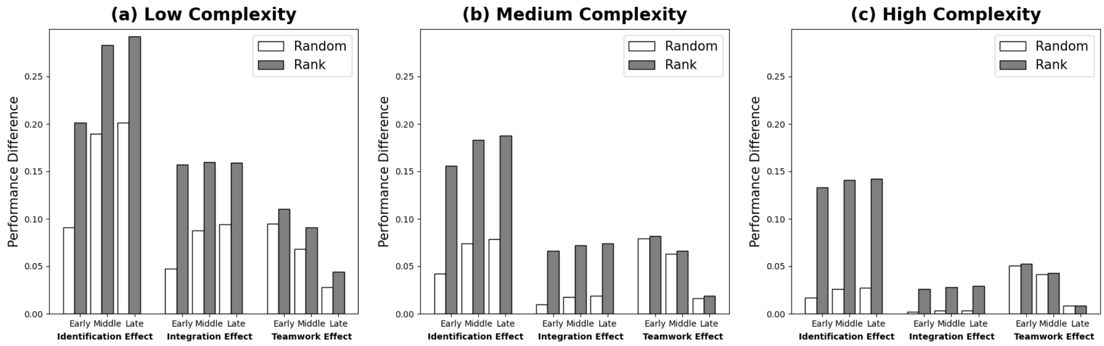

Figure 2.6: How Teaming According to Rank Emphasizes Different Effects

Notes: For visualization purposes, we subtract 0.5 (the baseline performance) from the *identification effect*.

solutions to integrate solution with. On the other hand, those identified as high-ranking solvers are also more likely to possess the requisite knowledge for the focal problem. As a result, teaming according to rank utilizes the *identification*, *integration* and *teamwork effects* simultaneously at the expense of weakening parallel search. When problem complexity is low, the benefits of the enhancing *integration* and *teamwork effects* outweigh the negative impacts of the *identification effect* (as shown in Figure 2.5c). A combination of three effects can therefore make up for the loss of parallel search, leading to an increase in the global crowdsourcing effectiveness (see Figure 2.3b). The optimal timing of teaming is thus around the middle of the crowdsourcing contest to leverage both the *integration* and *teamwork effects.* As the level of problem complexity increases, the *teamwork effect* becomes more critical and at the same time, is relatively easier to attain compared to the *integration effect.* The optimal timing of teaming advances (to earlier) to fully utilize the *teamwork effect* (see Figures 2.3c and d). Teaming according to rank at late timings, however, leverages the *teamwork effect* the least and relies mostly on the *identification* and *integration effects.* The *integration effect* along (approximately 0.03 units), however, is insufficient to offset the negative impacts of the *identification effect* (around 0.65 units). As a result, teaming may fail to compensate for the loss in parallel search (see Figure 2.3d). We synthesize the baseline results into a coherent theoretical framework (see Figure 2.7).

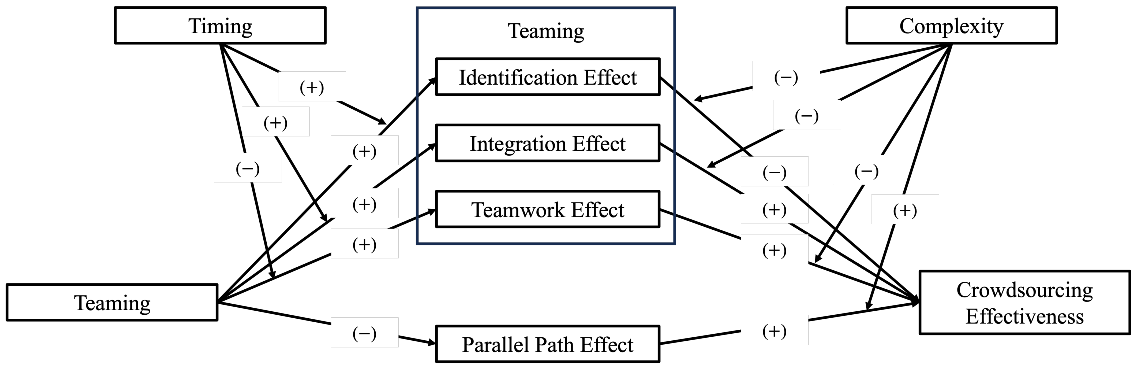

Figure 2.7: An Integrated Theoretical Framework for the Effects of Teaming on Crowdsourcing Effectiveness

2.5.2 Alternative Signals for Teaming

Next, we explore the impact of teaming on crowdsourcing effectiveness in generalized crowdsourcing environments through the lens of the identified mechanisms. Rank information (e.g., from the contest leaderboards on Kaggle) is only one type of publicly available information that may be utilized by solvers as signals in their teaming decisions. Other quality signals can also readily be afforded by crowdsourcing platforms by showcasing different types of information in the forms of, for example, detailed profile pages of solvers that highlight their expertise areas and historical records of solution submissions. To generalize our findings beyond teaming according to rank, we run a series of counterfactual experiments when alternative information is used in teaming decisions. Specifically, we identify and categorize the different types of information cues that can serve as quality signals on crowdsourcing platforms from the prior literature into three distinctive yet parsimonious categories. The first category of quality signals, which we have already investigated above, relates to solvers' contemporaneous performance, such as the ranking on the public leaderboard (Lee et al. 2018), the public rating given by the solution seeker (Mihm and Schlapp 2019, Wooten and Ulrich 2016), and the quality of the contemporaneous solutions made public (Jin et al. 2021). The second category of signals relates to solvers' knowledge background, such as academic qualifications or personal interests listed on their profile pages, the domains of past contests in which they participated (Jiang et al. 2021), and the content of past solutions (Majchrzak et al. 2021, Piezunka and Dahlander 2015). The third category of signals refers to solvers' general capability in problem solving beyond a specific task, such as performance on past contests as reflected on their profile page, their reputation on the platform (Zhang et al. 2019, Boudreau et al. 2016) and the number of submissions made in past competitions (Lysyakov and Viswanathan 2023, Jian et al. 2019).

2.5.2.1 Knowledge Background

This experiment examines the impacts of knowledge background as the basis of teaming decisions on crowdsourcing effectiveness via its impact on the parallel path, identification, integration and teamwork effects. Specifically, we extend our model to allow each solver's decision set D to be visible to other solvers. In considering teaming decisions, solvers will select the targets of teaming according to the extent of knowledge overlap between them and potential teammates (i.e., $|\boldsymbol{D}_u \cap \boldsymbol{D}_v|$ for solver agents u and v). Solvers could seek potential teammates with greater common ground (i.e., maximal knowledge overlap) to reduce the occurrence of knowledge incompatibilities and conflict. Alternatively, solvers can conduct teaming with solvers to maximize new knowledge being brought to the problem (i.e., minimal knowledge overlap). Finally, solvers could seek potential teammates by balancing common ground and new knowledge (i.e., moderate knowledge overlap). To capture different preferences, we allow solvers to select potential teammates according to minimal, moderate, and maximal extents of knowledge overlap. We set p_{teamup} as identical to the baseline experiment and p_{accept} as $0.14 \cdot (0.5 \cdot (e^{-0.1x} + 0.001))$, where x represents the relative order of a potential teammate's knowledge overlap with a focal agent over all possible teammates, to maintain a similar proportion of teams among the crowd of solvers as in the baseline (also see Appendix A.1.3 for details).

Figure 2.8 show the likelihood of finding the best solution when solvers seek to team up with other solvers with minimal (see solid lines), moderate (see dotted lines), and maximal knowledge overlap (see dash-dot lines) across the combinations of the timing of teaming and of problem complexity. We also display the results of individuals' independent parallel search (see dashed lines) and the rank-based teaming (see loosely dotted lines) from prior results as benchmarks. There are two noteworthy findings. First, in general, teaming according to minimal knowledge overlap increases the global crowdsourcing effectiveness the most, followed by moderate overlap and then maximal overlap. The only exception is when problem complexity

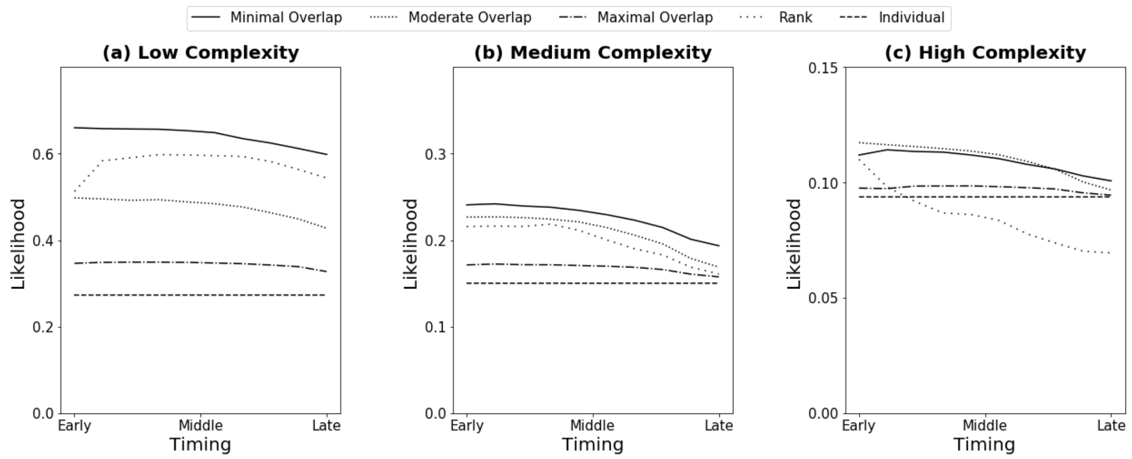

Figure 2.8: How Teaming According to Knowledge Overlap Affects Crowdsourcing Effectiveness

Notes: For brevity, (a)-(c) only show results for combinations of problem complexity (K) at low, medium and high levels (across columns of figures) and timing at early, middle and late stages (across the x-axis) – i.e., K=1, 6, 11 and T=3, 15, 27.

is high – in this particular case, teaming according to moderate knowledge overlap tends to outperform teaming according to both minimal and maximal knowledge overlap when the timing of teaming is early (but teaming according to minimal overlap still outperforms when the timing of teaming is late). Compared to the baseline results, teaming according to minimal knowledge overlap tends to outperform teaming according to rank, and the relative improvement is more pronounced when problem complexity is high (see solid, dotted, and loosely dotted lines for minimal overlap, moderate overlap and rank, respectively, in Figure 2.8). Second, the benefits brought by teaming according to knowledge overlap are moderated by both the timing of teaming and problem complexity. The likelihood of finding the best solution decreases as the timing of teaming becomes later, especially for complex problems. In the extreme case, teaming according to maximal knowledge overlap is only marginally beneficial to overall crowdsourcing effectiveness (see Figure 2.8c).

The above findings can also be explained by the interplay between the teaming and parallel path effects. Figure 2.9 summarizes how teaming up with other solvers with minimal (see gray bars), moderate (see white bars with slashes), and maximal knowledge overlap (see black bars with horizontal hatching) emphasize different effects (i.e., across rows) for different levels of problem complexity (i.e., across

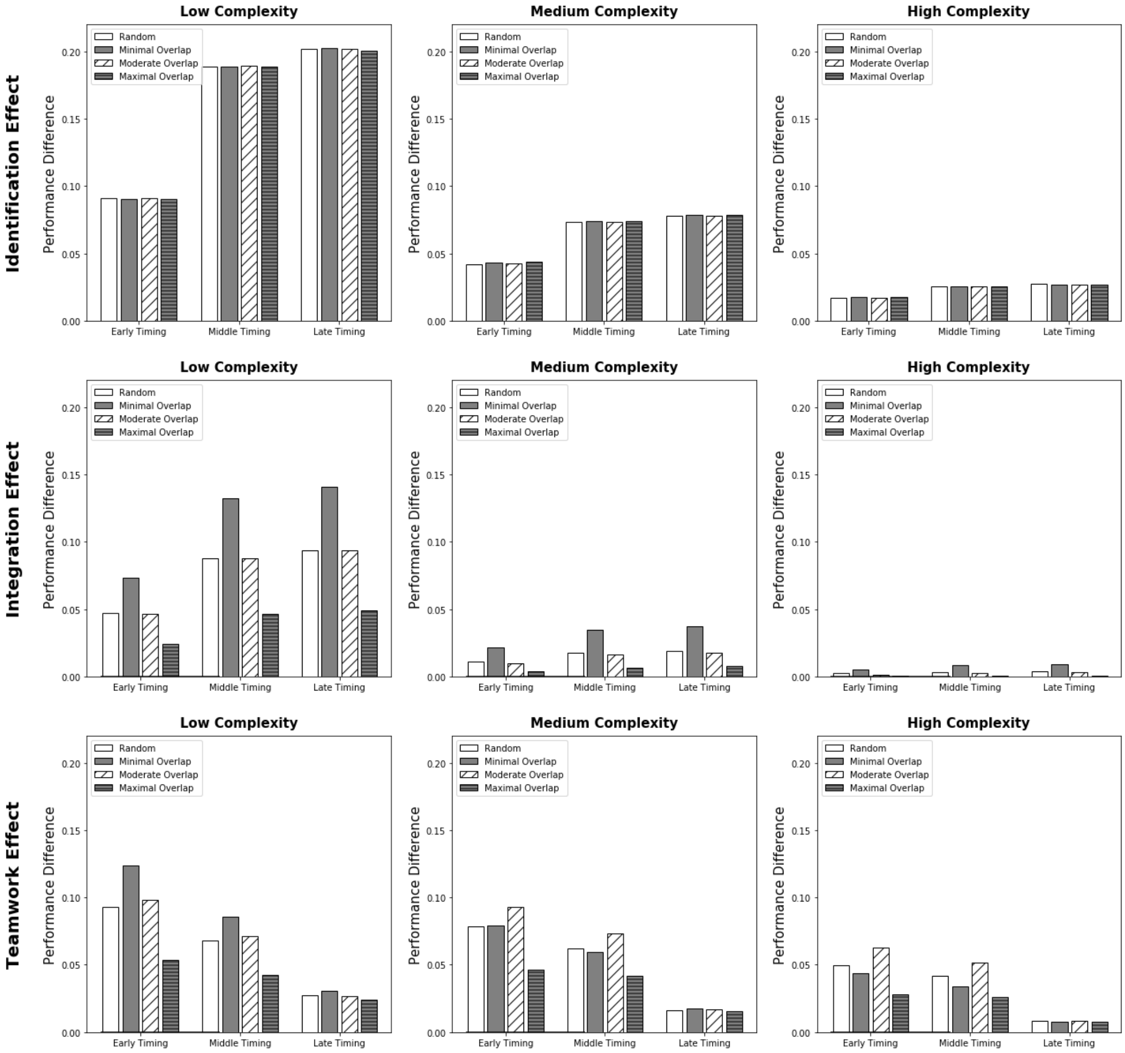

Figure 2.9: How Teaming According to Knowledge Overlap Emphasizes Different Effects

columns). From the first row of Figure 2.9, we observe that teaming according to different extents of knowledge overlap leverages the *identification effect* equivalently. As a result, the variations in the likelihood of finding the best solution are driven by their different levels of emphases on the *integration* and *teamwork effects*. Teaming with solvers with a minimal extent of knowledge overlap enables solvers to integrate their solution configurations of different sets of design elements into a high quality solution (i.e., *integration effect*) and to work on separate solution parts afterwards (i.e., *teamwork effect*). Considering both the *integration* and the *teamwork effects* are positively associated with crowdsourcing effectiveness (as in Figure 2.5c), teaming according to minimal knowledge overlap thus generates the highest overall crowdsourcing effectiveness when problem complexity is low because it leverages these two effects the most (see gray bars in the first column of Figure 2.9). As problem complexity increases, however, teaming according to minimal knowledge overlap may fail to leverage the *teamwork effect* due to a lack of common ground among teammates that facilitate solution co-development.

In contrast, teaming according to moderate extents of knowledge overlap does not leverage the *integration* and *teamwork effects* as much as teaming according to minimal knowledge overlap when problem complexity is low because it may fail to separate solvers to work on different sets of design elements (see white bars with slashes in Figure 2.9). Yet, it enables solvers to team up with those with whom they have common knowledge, which is essential to resolving conflicts arising from interdependent design elements through teamwork. As such, teaming according to moderate extents of knowledge overlap best leverages the *teamwork effect* (see white bars with slashes at the bottom right corner of Figure 2.9). Given that the *teamwork effect* plays a more crucial role than the *integration effect* in enhancing overall crowdsourcing effectiveness when problem complexity is high (as depicted in Figure 2.5c), teaming according to moderate extents of knowledge overlap emerges as most beneficial for overall crowdsourcing effectiveness (see Figure 2.8c). With late timings, however, it emphasizes the *teamwork effect* to a similar extent as teaming according to minimal knowledge overlap, as there remains insufficient time for the *teamwork effect* to fully manifest. Consequently, it tends to yield a lower likelihood of finding the best solution because it leverages the *integration effect* less than teaming according to minimal knowledge overlap (see gray bars and white bars with slashes in the second row of Figure 2.9).

As for teaming according to maximal knowledge overlap, it enables solvers to work on the same set of design elements but at the same time brings little new knowledge into the team. As such, it is unable to sufficiently leverage the *integration* and *teamwork effects* (see black bars with horizontal hatching in Figure 2.9). Therefore, teaming according to maximal knowledge overlap is always at best marginally beneficial compared to mere parallel search.

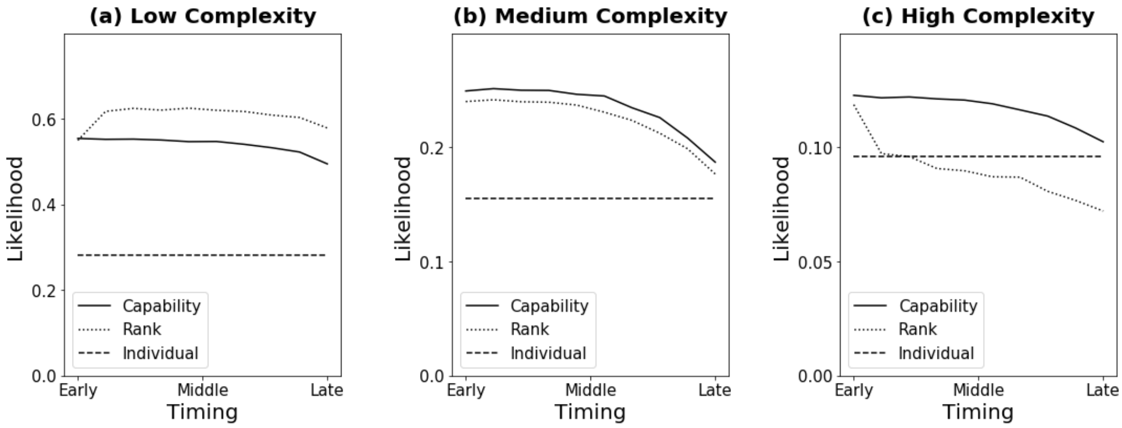

Figure 2.10: How Teaming According to Capability Affects Crowdsourcing Effectiveness

Notes: For brevity, (a)-(c) only show results for combinations of problem complexity (K) at low, medium and high levels (across columns of figures) and timing at early, middle and late stages (across the x-axis) – i.e., K=1, 6, 11 and T=3, 15, 27.

2.5.2.2 Problem-Solving Capability

In this experiment, we vary solvers' problem-solving capability and allow solvers to select the targets of teaming according to other solvers' search capabilities. Problem-solving capability is parameterized by the number of alternatives an agent can consider in each period. Solvers with high problem-solving capability can consider more alternatives and access the ramification of changing more than one design element at once within their decision set D. In other words, they are able to conduct distant search and are less likely to get stuck at local optima on the problem's solution landscape. Specifically, we vary the number of alternatives an agent can consider from 1 to $2 \times N_D$. Solvers who can consider less than or equal to N_D alternatives can only search locally with a Hamming distance of one. Solvers with higher problem-solving capability can consider N_D local alternatives, and additionally, distant alternative(s) with a Hamming distance of two. We set p_{teamup} and p_{accept} as identical to the baseline experiment and sort solvers according to their search capability instead of contemporaneous performance.

Figure 2.10 shows the likelihood of finding the best solution when solvers conduct teaming according to problem-solving capability. These results show that teaming according to problem-solving capability improves crowdsourcing effectiveness generally. Similarly, the benefits brought by teaming gradually diminish as the timing

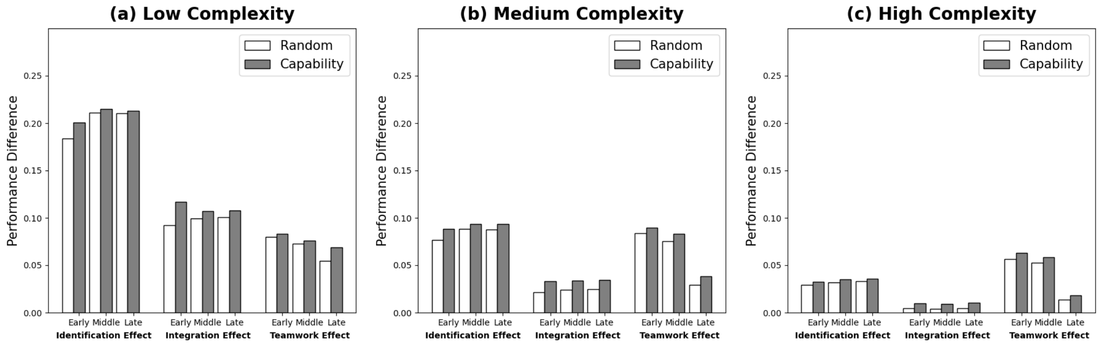

Figure 2.11: How Teaming According to Capability Emphasizes Different Effects

of teaming becomes later, and as problem complexity increases. Compared to our baseline results when teaming decisions are based on rank, teaming according to capability is less (more) beneficial when problem complexity is low (high) (see solid and loosely dotted lines for teaming based on capability and rank, respectively, in Figure 2.10).[13]

Figure 2.11 shows how teaming according to problem-solving capability emphasizes the various effects. It magnifies the *teamwork effect* by allowing for more intensive and distant search after teaming. As shown by the gray bars regarding the *teamwork effect*, teaming according to problem-solving capability always leverages the *teamwork effect* marginally more compared to randomly teaming up. Meanwhile, problem-solving capability only roughly correlates with the quality of separate solutions as high problem-solving capability usually leads to better individual solutions at the time of teaming. As shown by the gray bars regarding the *identification* and *integration effects*, teaming according to problem-solving capability utilizes these two effects marginally more than the random baseline. These results suggest that when problem complexity is low, teaming based on problem-solving capability is less effective than teaming based on rank in enhancing overall crowdsourcing effectiveness due to its limited use of the *integration* and *teamwork effects.* However, it is also less susceptible to the adverse impact of the *identification effect.* Consequently, it can maintain or even increase crowdsourcing effectiveness under high problem complexity (see Figure 2.10).

[13]Here, we also vary the number of alternatives agents can consider when they team up according to rank.

2.5.3 Alternative Teaming Tendencies

The simulation parameters governing solvers' teaming decisions (i.e., p_{teamup} and p_{accept}) were calibrated using empirical data from Kaggle in the baseline model. However, crowdsourcing solvers may exhibit different teaming tendencies across platforms and contexts. For instance, Kaggle is well known for its well-established culture of collaboration, where solvers frequently form teams and share knowledge to improve performance (e.g., Jin et al. 2021, Cao et al. 2022). Solvers on other platforms may be more competition-oriented and solve the crowdsourced problems under different incentive structures, leading to variation in the likelihood of forming teams on the fly and in the probability of accepting team-up invitations from solvers with either less or more favorable signals. This experiment focuses on two representative cases to generalize our findings to broader crowdsourcing environments. First, solvers are more willing to work with others and therefore have a higher (lower) likelihood to initiate team-up invitations in crowdsourcing environments that advocate collaboration (competition). The first set of analyses therefore focuses on varying the probability of forming teams on the fly (i.e., p_{teamup}). While our main analyses examine the effect of teaming in a binary setup (i.e., with and without teaming; Table 2.1), this set of analyses also allows us to investigate whether teaming can compensate for the loss of parallel search and enhance global crowdsourcing effectiveness in a continuous manner. Second, solvers' teaming behaviors are likely influenced by the competitive dynamics within crowdsourcing contests. The second set of analyses examines global crowdsourcing effectiveness when solvers' decisions to send and accept team-up invitations are influenced by contest specific competition.

We first vary solvers' tendency to initiate a team-up invitation (p_{teamup}) in the baseline model (i.e., teaming according to rank; $p_{teamup} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$) to explore how different levels of teaming probability on the crowdsourcing platform influence the baseline results. Figure 2.12 shows the likelihood of finding the best solution across solvers' tendency to initiate a team-up invitation (p_{teamup}) for problems with low, medium, and high complexity. In general, as more teams are formed (i.e., as p_{teamup} increases), the impact of teaming on global crowdsourcing effectiveness is amplified. When teaming according to rank is beneficial (e.g., under low problem complexity as shown in Figure 2.3b), a higher teaming

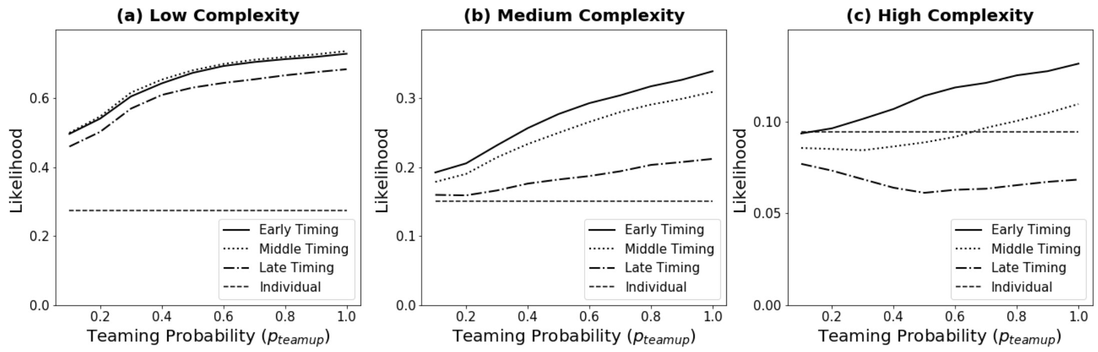

Figure 2.12: The Impact of Teaming According to Rank Across Teaming Probability

Notes: When $p_{teamup} = 1$, a solver may still fail to form a team if none of its team-up invitations are accepted by other solvers.

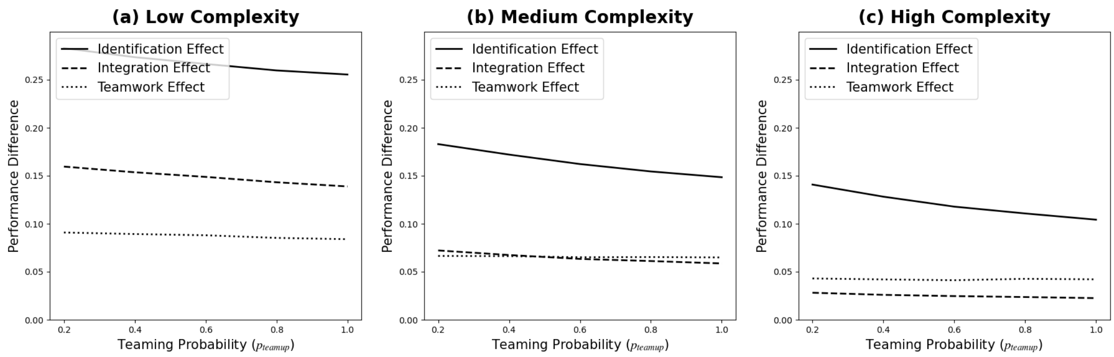

Figure 2.13: How Teaming According to Rank Emphasizes Different Effects Across Teaming Probability

Notes: For brevity, we only report results for middle timing (i.e., $T = 15$). Similar patterns were observed across other timing choices.

probability increases the likelihood that the crowd would discover the best solution; if teaming is detrimental (e.g., when problem complexity is high and teaming occurs at the late stages as shown in Figure 2.3d), a higher teaming probability exacerbates the negative impact when $p_{teamup} \leq 0.4$ (as shown by the dash-dot line in Figure 2.12c). There are several noteworthy findings. First, when problem complexity is high, teaming at the middle stages shifts from being detrimental to beneficial as the teaming tendency increases (see dotted line in Figure 2.12c). Second, teaming at the late stages also becomes less detrimental if teaming probability further increases (i.e., $p_{teamup} > 0.4$).

These discrepancies are consistent with the theoretical mechanisms identified earlier regarding how rank emphasizes the different teaming effects. As shown in Figure 2.13, as teaming probability increases, the *identification effect* declines significantly. In the extreme case, when all agents engage in teaming, the *identification effect* becomes equivalent to random selection of teammates. Similarly, the *integration* and *teamwork effects* also decline. The decrease, especially for the *teamwork effect*, is less pronounced. It thus becomes more likely for the *integration* and *teamwork effects* to alleviate the negative impact of the *identification effect* and enhance global crowdsourcing effectiveness as teaming probability increases (as shown in Figure 2.12). There is therefore an inverted-U shape relationship between teaming probability and crowdsourcing effectiveness, in which the amplification of the negative impact and the reduction of the *identification effect* reach a balance.

Second, to capture how solvers' decisions to send and accept team-up invitations are influenced by the competitive nature of crowdsourcing, we model a prize structure in which top performers (according to the leaderboard) are disincentivized from teaming, as doing so would require sharing the prize. Conversely, low-performing solvers would engage in teaming in order to outperform those prize contenders. Figure 2.14 shows the results when the top 10 performers are designated as prize contenders.[14] Compared to the baseline results, we find that teaming according to rank in competitive settings is less likely to enhance crowdsourcing effectiveness when complexity is low (Figure 2.14a), but more likely to do so when complexity is high (Figure 2.14c). Figure 2.15 shows how teaming emphasizes different effects of teaming. As indicated by the gray bars, the *identification effect* declines significantly compared to the baseline results, When top performers are reluctant to team up, it becomes unlikely to engage high quality solvers and their respective solutions in teaming. Consequently, both the integration of separate solutions (*integration effect*), and the co-development of the integrated solution (*teamwork effect*) would be less effective. When the problem complexity is low, the *identification effect* is only marginal detrimental, the reduction of the *integration* and *teamwork effects* thus explains the performance decline compared to the baseline. When the problem is more complex, however the overemphasis on the *identification effect* is avoided in competitive settings without substantially reducing the *integration* and *teamwork effects*. As a result, teaming according to rank has a higher likelihood to enhance

[14]The results remain qualitatively unchanged if we vary the group of prize contenders to include the top 1 or top 100 performers.

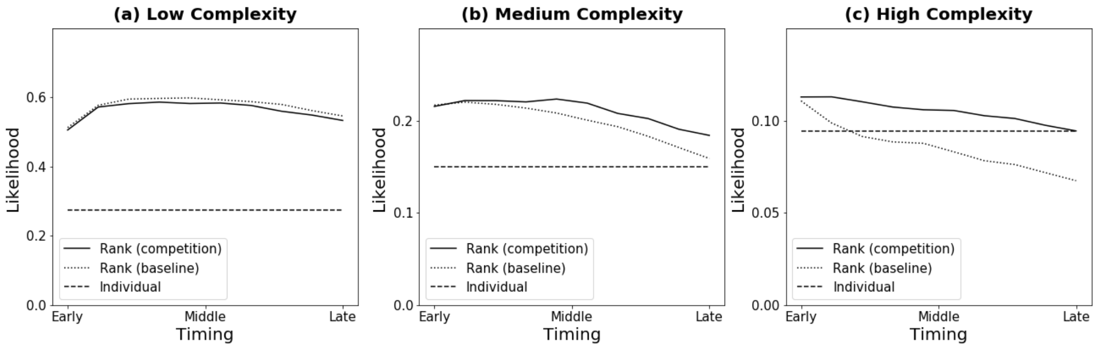

Figure 2.14: The Impact of Teaming According to Rank under Competition

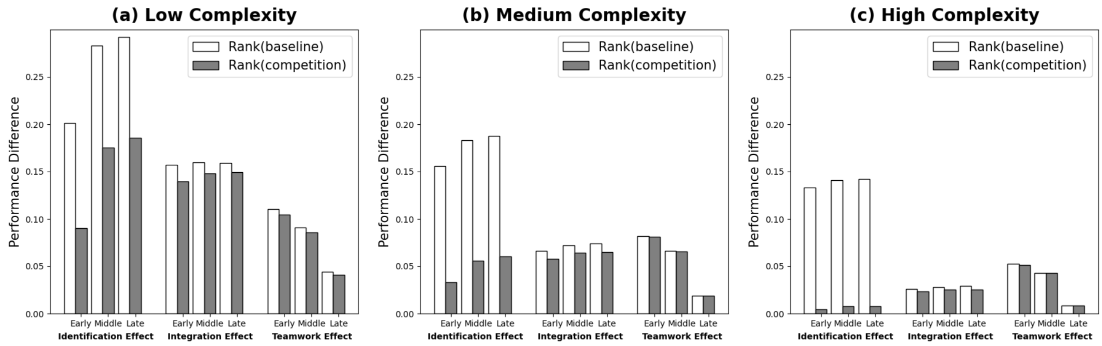

Figure 2.15: How Teaming According to Rank Emphasizes Different Effects (under Competition)

global crowdsourcing effectiveness in environments that are more competitive than our empirical context.

2.5.4 Sensitivity Analyses

We conducted a number of sensitivity analyses to ensure that our results are robust to our choice of model invariants. First, we varied the number of solvers $(40, 100, 200, 300, 400, 500)$ and found no substantive change in our results. Second, we varied the number of elements $N = 16$ (i.e., with a solution space with $2^{16} = 65,536$ unique configurations) and found consistent results. Third, we changed the learning rate among solvers p_{learn} from $(0, 1]$. Our results were robust across variations in p_{learn}. Fourth, in the main analyses, we only report results for the likelihood of finding the best solution. We also varied the range of best solution(s) from the 1st to the first 10 and first 100 solutions and found no qualitative difference. Fifth, we varied the size of solvers' decision sets (i.e., $N_D = 4, 6, 8$). The size of solvers' decision sets N_D affects the number of design elements each solver can configure and the average knowledge overlap among solvers. As a result, it alters the strengths of the *identification*, *integration*, and *teamwork effects*. Yet, despite the change in the effect sizes, the qualitative conclusions regarding how teaming interacts with the parallel path effect on impacting global crowdsourcing effectiveness remains robust. It is also possible that certain design elements being more commonly represented in the population of solvers and others being less so. We also conducted experiments in which elements are randomly selected into agents' decision sets according to a nearly normal distribution (i.e., $Beta(2, 2)$) instead of a uniform distribution and found no qualitative difference. Lastly, it is possible that more than two formerly independent solvers team up. We allowed agents to flexibly form teams ranging from two to five team members and found our conclusions remain qualitatively unchanged.

2.6 Discussion and Conclusion

In this paper, we offer a novel perspective on how to better leverage the power of the crowd in crowdsourcing via teaming – self-selected team formation during crowdsourcing contests. Teaming offers a flexible approach through which previously independent solvers interact with each other to collectively solve problems in crowdsourcing. Solvers can start crowdsourcing independently and join forces to jointly develop solutions. Teaming, however, shifts the distribution of quality of all discovered solutions via diminishing the parallel path effect, thereby influencing the likelihood of discovering the best possible solutions (i.e., crowdsourcing effectiveness). Moreover, the teaming process is driven by solvers' self-selection. Self-selection shapes how the whole population of the crowd searches the problem's solution landscape (independently and jointly) and ultimately impacts global crowdsourcing effectiveness. We first find suggestive empirical evidence from Kaggle that solvers are likely to rely on quality signals from publicly available information in teaming on crowdsourcing platforms in the teaming process. Based on the empirical observations, we extend the *NK* fitness landscape model and conduct counterfactual experiments to unlock the impact of self-selected teaming on global crowdsourcing effectiveness. Our results suggest that the impact of teaming on crowdsourcing effectiveness originates from the immediate returns from identifying other solvers (or their solutions) to integrate with at the time of teaming and potential (but time-consuming) returns from joint teamwork after teaming. The *identification*, *integration*, and *teamwork effects* are moderated by problem complexity and the timing of teaming and may make teaming detrimental to overall crowdsourcing effectiveness under certain conditions (see Figure 2.7). The proposed mechanisms are used to explain the results of counterfactual experiments on teaming according to different quality signals and under different tendencies. These findings offer theoretical and practical implications.

2.6.1 Contribution to Theory

Our study contributes to the crowdsourcing literature by highlighting the tension between the interactions among crowdsourcing solvers (such as through teaming) and the parallel path effect. The conventional wisdom regarding crowdsourcing has assumed crowdsourcing effectiveness as a function of solver participation levels (i.e., the parallel path effect) (Boudreau et al. 2011, Afuah and Tucci 2012). Recent studies have begun to question this stream of research by examining the effectiveness of knowledge combination among solvers (Jin et al. 2021, Riedl and Seidel 2018, Cao et al. 2022). This study bridges these two streams by showing that when interactions take place on the fly, they are interdependent with the parallel path effect. When solvers team up on the fly in crowdsourcing, the number of submitted solutions at the end of crowdsourcing is inherently reduced thus weakening the extent to which parallel path effect can be utilized, but teaming also impacts crowdsourcing effectiveness through varying combinations of identification, integration and teamwork effects. Crowdsourcing research should continue to bridge these two perspectives.

More specifically, we contribute to the crowdsourcing literature by highlighting that the parallel path effect and teaming jointly influence global crowdsourcing effectiveness. The prior literature on crowdsourcing has mainly used the quality or performance of individual solutions to represent crowdsourcing effectiveness (Jin et al. 2021, Riedl and Seidel 2018, Cao et al. 2022). This perspective, albeit valid, does not necessarily apply to teaming when the parallel paths join force and change how the whole population of the crowd searches the problem's solution landscape. Our simulation results suggest that although identification, integration and teamwork effects can all help teams to improve solution quality compared to their individual counterparts, their impacts on crowdsourcing effectiveness (i.e., the probability to generate extreme-value outcomes) are moderated by problem complexity. When the *identification effect* dominates, teaming may only reposition solvers' solutions within already explored areas on the landscape and infuse little new knowledge into solutions for complex problems through the integration of separate solutions and particularly through subsequent teamwork to reach previously unexplored configurations. As a result, teaming may fail to compensate for the reduction in parallel search and worsen overall crowdsourcing effectiveness. Scholars can re-examine previous findings and resolve the ambiguity between solution quality and crowdsourcing effectiveness. Moreover, teaming also uncovers a novel and interesting aspect that has yet to receive much attention. In the crowdsourcing context, participants conduct teaming in a self-selection manner. Most solvers are strangers to one another and can hardly verify others' underlying qualities from repeated collaboration. However, they may make use of a variety of publicly available information that crowdsourcing platforms can (and do) provide as quality signals and team up to varying extents. We identify a number of types of quality signals and tendencies to form teams. We find that they would influence how the identification, integration, and teamwork effects are leveraged and, consequently, differentially affect crowdsourcing effectiveness. The impact of teaming and the potential misalignment between solution quality and crowdsourcing effectiveness may thus become very contextual, depending on what types of information are made salient by crowdsourcing platforms.

We also contribute to the crowdsourcing literature by providing a new perspective to inform crowdsourcing platform design. Various design elements have been examined under the guidance of the parallel path effect to encourage solver participation and facilitate independent search (Lee et al. 2018, Jiang et al. 2021, Jian et al. 2019, Mihm and Schlapp 2019, Boudreau et al. 2016). However, these design elements might also affect crowdsourcing effectiveness through the mechanism of interactions among solvers (e.g., via self-selected teaming). Our results suggest that crowdsourcing effectiveness is determined by the interplay of the parallel path effect and teaming (i.e., *identification*, *integration* and *teamwork effects*). If a platform design element strengthens the parallel path effect to the neglect of teaming, the overall effectiveness of crowdsourcing might be at risk when no individual participant has the breadth of knowledge to be able to derive a high-quality solution for the problem. Conversely, some design elements might sacrifice the parallel path effect to some extent in exchange for greater likelihood of teaming (e.g., the teaming willingness indicator on Kaggle). The benefits from these features might be questionable if little new knowledge is infused to affect the solution search process. Even if many participating solvers are highly capable, the collective crowd might end up with redundant exploitation of what has already been discovered (i.e., weak integration and teamwork effects). The present study provides future research a framework to systematically examine the impacts of different design elements on crowdsourcing effectiveness from the perspective of both the parallel path effect and the effects associated with solver interactions.

Lastly, our extensions to the *NK* fitness landscapes model and its usage with simulation experiments to theorize crowdsourcing also contributes to the crowdsourcing literature. Even though crowdsourcing was aptly conceptualized as landscape search more than a decade ago (Afuah and Tucci 2012), there have only been a few studies that actually investigate different aspects of crowdsourcing through the theoretical and methodological lens of landscape search. The *NK* fitness landscapes model offers a natural fit for investigating crowdsourcing with its advantage in seemlessly capturing massive parallel search for problems with different levels of complexity and patterns of interdependencies (Baumann et al. 2018). It gives scholars a new way to theorize and reflect on empirical findings, test the boundary conditions of conventional wisdom, and explore multilevel outcomes (e.g., not only the impact on the quality of individual solutions but also on the distribution of quality of all solutions). A range of additional parameters and rules can be incorporated into our baseline model. Here, we suggest several directions for future research. First, to simplify our analysis, we fixed the search pattern for all agents in our simulation experiments. However, as pointed out by recent empirical observations (Jiang et al. 2021, Sanyal and Ye 2023), solvers' search distance and intensity might vary based on predetermined goals and in-progress feedback. Future research can extend our model by allowing for heterogeneous search patterns. Second, we generally set the K parameter at different levels to study the impact of overall problem complexity. Our model can be extended to study the impact of different problem structures such as nearly-decomposable (Levinthal and Workiewicz 2018) and non-decomposable problems (Chan et al. 2021, Rivkin and Siggelkow 2007) by adjusting the pattern of interdependencies. Third, we model the limited knowledge possessed by boundedly rational solution solvers but assume the solution seeker has a full range of knowledge to effectively evaluate solution solvers' submissions. A handful of recent studies have started to examined whether the seeker can accurately evaluate all submitted solutions (Piezunka and Dahlander 2015). Relaxing this assumption may allow a closer look at how a solution seeker might adopt and integrate the winning solution and provide further insights to why many crowdsourcing campaigns succeed in attracting a large number of participants but still fail to produce significant innovations for the solution seeker.

2.6.2 Implications for Practice

Our study has important implications for different stakeholders. First, the solution seeker should better understand the population of solvers to whom they are going to crowdsource their problems. The outcomes of crowdsourcing campaigns depend not only on the knowledge portfolio of solvers but also on the process through which the problem solutions are discovered. If solvers on a crowdsourcing platform are known to favor collaborating, the solution seeker should be aware of the upsides and downsides of teaming. The solution seeker can strategically design crowdsourcing campaigns to avoid the side effect of teaming on innovative solutions. For example, our study suggests that both problem complexity and team-up timing affect overall crowdsourcing effectiveness. The solution seeker could split a complex problem into several relatively less complex sub-problem or artificially set an earlier deadline for teaming up for complex problems to alleviate the negative side effects of teaming. In a similar vein, the solution seeker could actively call attention to the upside of teaming. For example, the solution seekers could also strategically design the reward structure to better incentivize solvers to conduct teaming in a desirable way. For complex problems where crowdsourcing effectiveness is most likely to be hampered by teaming, the solution seeker could change the reward structure (e.g., the best solution from collaboration across distant domains) to motivate solvers to form teams in a manner that induces proper levels of diversity in solution search.

This study also offers implications for crowdsourcing platforms. First, crowdsourcing platforms should carefully consider the impacts of teaming in organizing crowdsourcing contests. For example, the prior literature on team recommendation systems on platforms finds that users tend to form teams with low diversity (e.g., Gomez-Zara et al. 2019, 2020), our results suggest this tendency will have ambivalent effects on crowdsourcing effectiveness. Crowdsourcing platforms should consider other contextual factors to better design team recommendation systems for solvers. Crowdsourcing platforms could embed solvers' experience and knowledge background in the real-time leaderboards to make capability and knowledge background information also more salient at different stages of crowdsourcing campaigns to nudge the teaming process. In a similar vein, while some platforms have already implemented features relevant to teaming, we encourage crowdsourcing platforms to carefully re-examine these design practices and check whether they are counterproductive to crowdsourcing effectiveness above and beyond merely increasing solvers' average performance (which may or may not lead to overall crowdsourcing effectiveness). Furthermore, our results suggest that crowdsourcing effectiveness is not only affected by those design elements directly related to crowdsourcing per se but also by those elements that serve as quality signals for teaming. For instance, developing an effective reputation system that captures solvers' problem-solving capability might induce teaming to occur in a more desirable way.

2.6.3 Limitations and Directions for Future Research

Despite the novel theoretical and practical insights generated, this study is not without limitations. We treat crowdsourcing campaigns as separate and independent. Real-world campaigns may have spillover effects on subsequent contests within a crowdsourcing platform (Zhang et al. 2019, Mo et al. 2021, Park et al. 2022). For instance, solvers may decide to collaborate with knowledgeable others or do more explorative search to maximize learning opportunities. This, of course, reduces solvers' current performance but intuitively facilitates search in future crowdsourcing campaigns. Nonetheless, our study, while being at the campaign-level, offers novel insights for the solution seeker. Given that there are many solution seekers on crowdsourcing platforms, we believe that solution seekers care more about the outcome of their campaigns given a particular pool of the crowd rather than the long-term development of the crowd. As a result, the outcomes of learning that happens across campaigns can be abstracted as different initializations of the number of modifiable decisions (N_D) in our model. Future research can extend our model to capture the interdependencies across different campaigns and solvers' longitudinal performance to investigate the long-term impact of (re)combination via teaming for crowdsourcing platforms.

Another limitation is that our extended *NK* model, similar to other simulation models, is still a simplified abstraction of the reality. First, we assume that solvers are relatively homogeneous except for the knowledge background and problem-solving capability they possess. Team performance in real crowdsourcing campaigns might also be affected by diversity in demographics (van Knippenberg et al. 2004, Gompers et al. 2017, Dissanayake et al. 2019). Future work could explore the impact of demographics on teaming decisions and subsequent crowdsourcing effectiveness. Second, we only capture the basic properties of teams with several simplifications. For instance, solvers might also use various quality signals in a hybrid manner. Future studies can apply our simulation model and experimental design to examine the interaction effect of different information. It is worth noting that most teams in crowdsourcing are virtual and have to work across geographical boundaries. However, these kinds of internal behavioral interactions among team members are unbeknownst to scholars. Most intra-team coordination and teamwork would not happen on the crowdsourcing platform but privately through third-party communications and collaborative development platforms. It would be cost-prohibitive to track all solvers' interactions in an empirical experimental study of sufficient scale to reenact the parallel path effect, especially since the objective is to examine the performance (i.e., crowdsourcing effectiveness) of the population of a large and anonymous crowd. The simulation methodology equips us with a theoretical framework to abstract these unobserved factors and captures their nuance under different contingencies and anticipate possible outcomes for future empirical inquiries. In addition to the possible extensions discussed earlier, we invite other researchers to further develop and extend the *NK* fitness landscapes model to study other aspects of crowdsourcing.

Chapter 3 The Cave of Crowdsourcing: A Systems Investigation of Crowdsourcing Effectiveness

"The truth would be literally nothing but the shadows of the images."
– Plato, *The Republic*

3.1 Introduction

Crowdsourcing, defined as attracting external solvers with rewards to solve challenging internal problems, is fast emerging as a mainstream tool for firms' innovation (Boudreau and Lakhani 2013, Liu et al. 2014, Terwiesch and Xu 2008). By involving a crowd of solvers with diverse knowledge, crowdsourcing enables firms to harness knowledge beyond what they already know and as a result, addresses the uncertainty in the knowledge required to solve complex problems (Boudreau et al. 2011). This mechanism, termed "the parallel path effect," theoretically accounts for the potential of crowdsourcing.

However, despite the tremendous potential in theory, firms as solution seeker are still failing to effectively leverage crowdsourcing (Dahlander and Piezunka 2014, 2020). For example, the Netflix Prize contest, one of the most famous crowdsourcing campaigns, managed to attract more than forty thousand teams to work on the problem but eventually awarded the prize to a solution which could not be implemented due to unexpected engineering costs that were unforeseen at the time of initial problem specification (Amatrian and Basilico 2012). In a similar vein, recent studies find that firms tend to myopically favor solutions they are familiar with in spite of the numerous and varied ideas and solutions that solvers identify (Piezunka and Dahlander 2015). These observations are consistent with absorptive capacity theory (Cohen and Levinthal 1990), which suggests that a firm's ability to identify and exploit external knowledge (such as solutions developed by the crowd) is primarily determined by its existing knowledge. The interplay between the parallel path effect and absorptive capacity points to a theoretical tension that has yet to be carefully explored. On the one hand, solvers' diverse knowledge is the source of crowdsourcing effectiveness (i.e., the solicitation of high quality solutions from the crowd) in addressing the firm's knowledge uncertainty, especially when the firm's problem involves a broad set of knowledge domains (Boudreau et al. 2011, Jeppesen and Lakhani 2010). On the other hand, the more diverse solvers' knowledge is, the more difficult it will be for a seeker firm to effectively absorb the knowledge with its existing knowledge.

The tension between the parallel path effect and absorptive capacity suggests the seeker firm and solvers with diverse knowledge inevitably constitute a suboptimal system constrained by the seeker firm's knowledge.(Hereafter, we use the crowdsourcing system to refer to the union of the seeker firm, solvers, and their interactions.) In the mode of crowdsourcing, the solvers' parallel search needs to be evaluated and selected by the seeker firm and then applied to the crowdsourced problem. As the crowdsourced problem becomes increasingly complex and requires knowledge from multiple domains, there may be no individual actor (not only among solvers and also including the seeker firm) that possesses the full range of knowledge (Gurca et al. 2023, Majchrzak and Malhotra 2020). Due to a limited understanding of domains beyond its existing knowledge, the seeker firm may fail to evaluate solvers' solutions and steer their parallel search correctly. As such, solvers' endeavors to increase the likelihood of being selected by the seeker firm as winners may deviate from effective problem solving to merely fitting the firm's knowledge about the problem (Park et al. 2022). The crowdsourcing system thus produces satisfactory solutions for the firm but (potentially) ineffective solutions for the crowdsourced problem. For instance, when the seeker firm uses crowdsourcing contests to recruit solvers to build machine learning algorithms and test proposed algorithms on a holdout data set (e.g., Jin et al. 2021, Lee et al. 2018), how the holdout data set along with how various features were generated from the firm's daily operations and what metrics should be used to assess algorithm performance depends on the firm's initial understanding of the problem. The firm might heed less to solvers' trial and error outside its knowledge space. Hosting a crowdsourcing contest might thus help fit the given data but not necessarily solve the business problem that motivated the firm to opt for crowdsourcing rather than internal R&D (Amatrian and Basilico 2012).

However, the existing literature on crowdsourcing has yet to differentiate satisfying the firm within the system (of the firm and solvers) from enhancing the overall system performance in discovering high quality solutions. Interactions between the seeker firm and solvers within the system, such as the acquisition of solutions via parallel search (Liu et al. 2014, Jian et al. 2019) and ongoing feedback from the seeker firm to solvers (Jiang et al. 2021, Mihm and Schlapp 2019, Sanyal and Ye 2023, Wooten and Ulrich 2016), have been investigated to help the seeker firm get better solutions from crowdsourcing. Yet, what constitutes "better" in the system is defined by the firm's knowledge about the solution space. Satisfying the seeker firm is only part of the crowdsourcing system. The selected solution is always with the highest quality as *perceived* by the seeker firm but not necessarily be the one with the highest *absolute* quality. When the firm's existing knowledge is limited, solutions adopted by the firm might be attractive for the firm at first glance but fail to ultimately induce genuine innovation. The neglect of a systems view of crowdsourcing thus hinders crowdsourcing from being utilized effectively. As such, an important question remains unanswered: How to enhance crowdsourcing effectiveness from a systems view?

This study aims to generate theoretical insights from a systems view to stimulate future research on crowdsourcing. Specifically, we study crowdsourcing by considering the seeker firm and a crowd of solvers as a system and investigate the performance implications of the interactions between the seeker firm and solvers on the ability of the crowdsourcing system to obtain high quality solutions. Given that in most crowdsourcing practices, the seeker firm evaluates the solvers solutions, the evaluation should be according to its knowledge. Therefore it is difficult, if not impossible, to observe the ground truth – i.e., how the system consisting of the firm and solvers performs. Therefore, this study takes a simulation approach to answer the above questions. Simulation is particularly useful in developing theory when the focal phenomenon involves multiple interacting processes and when empirical data is challenging to obtain (Davis et al. 2007). The created simulation model explicitly captures different levels of the seeker firm's existing knowledge and enables us to examine the interactions between the seeker firm and external solvers from a systems view. The simulation results reveal two paradoxes related to the solution flow from the crowd of solvers to the seeker firm and related to the knowledge flow from the seeker firm to the crowd. First, when crowdsourcing is most needed to develop solutions beyond the seeker firm's existing knowledge, the crowdsourcing system is also most likely to overlook high quality solutions developed by the crowd. Second, the seeker firm's feedback to solvers may make the seeker firm more satisfied with crowdsourcing at the cost of worsening crowdsourcing effectiveness, leading to a system collapse. To resolve these paradoxes, we investigate whether the seeker firm could enhance the crowdsourcing system's ability to obtain high quality solutions by proactively learning from the crowd about what constitute a good solution. We find this knowledge flow from the crowd to the seeker firm can enhance crowdsourcing effectiveness, especially when the seeker firm has a low level of knowledge and the problem is not complex. Notably, we also find a potential synergy between the seeker firm's knowledge and the wisdom of the crowd, where modest learning from the crowd can lead to higher quality solutions compared to relying solely on the firm's own knowledge or the wisdom of the crowd. This knowledge flow also interacts with other flows in affecting the overall crowdsourcing effectiveness. The crowdsourcing system is able to obtain higher quality solutions by first learning from the crowd and providing feedback only in the latter half of crowdsourcing.

Our study makes several contributions. First, to our knowledge, this is one of the first studies to investigate crowdsourcing from a systems view. It highlights the potential misalignment between satisfying the seeker firm in the system and enhancing the overall system performance in discovering high quality solutions. Second, this study contributes to research on the parallel path effect and the seeker firm's feedback in affecting crowdsourcing effectiveness by pointing to the critical role of the seeker firm's knowledge in the crowdsourcing system. Third, this study also contributes to the crowdsourcing literature by highlighting the possibility and boundary conditions of using the wisdom of the crowd to further enhance the overall system performance. Each of these contributions points to exciting future research inquiry.

3.2 Theoretical Foundation

3.2.1 Crowdsourcing as Landscape Search

The landscape metaphor has become a useful theoretical framework to study complex problem solving (Kauffman 1993). It uncovers the performance implication of problem solving as adaptive search in different problem environments. The ruggedness of the landscape (i.e., the complexity of the solution space) hinders landscape search and as a result, creates challenges to effective problem solving – when the solution space is complex, problem solving is likely to end up with suboptimal solutions (i.e., local optima on the landscape) where no improvement can be attained by merely considering incremental alternatives. As such, the ability to escape local optima and explore many different parts of the solution landscape (i.e., solution space) generates the greatest chance of finding the optimal solution, especially for complex problems (Fleming and Sorenson 2003).

When relying on internal R&D, the seeker firm can only search the solution landscape according to its existing knowledge and may fail to develop solutions beyond the scope of its knowledge base. In contrast, crowdsourcing offers the seeker firm a cost-efficient approach to finding high quality solutions through large scale parallel search of the solution landscape (Afuah and Tucci 2012). Through crowdsourcing, a seeker firm can attract solvers with diverse knowledge about the problem. The diversity in knowledge distinguishes the solvers' initial positions on the solution landscape and consequently influences their subsequent search trajectories, enabling the crowd of solvers to explore the landscape more extensively than the seeker firm could through internal R&D. At the end of crowdsourcing, all solutions discovered by the crowd of solvers are submitted to the seeker firm for reward and as a result, become visible to the seeker firm. In this way, the seeker firm is able to conduct distant search on the rugged solution landscape to increase the likelihood of obtaining high quality solutions. This is essentially the mechanism of the "parallel path effect" (Boudreau et al. 2011) where increasing the number of crowdsourcing participants is associated with more extensive exploration of different parts of the landscape and as a result, increases the likelihood that seeker firm gains access to high quality solutions.

This idea is deeply ingrained in the contemporary crowdsourcing literature. A variety of elements, including the reward structure (Yang et al. 2009, Liu et al. 2014, Jian et al. 2019), problem specification (Jiang et al. 2020), feedback on progress (Jiang et al. 2021, Sanyal and Ye 2023), the number of live crowdsourcing campaigns (Mo et al. 2021), and availability of AI assistants (Lysyakov and Viswanathan 2023), have been investigated in terms of their impact on solver participation rates and the amount of effort solvers exert. A key aim of this line of research is to enable the seeker firm to obtain high quality solutions by enhancing the parallel path effect. What requires careful consideration, however, is whether and to what extent the seeker firm can fully leverage the parallel path effect. When the problem is broad and complex, the seeker firm may not fully understand all aspects of solutions due to its limited knowledge regarding the domains relevant to the crowdsourced problem. Its limited knowledge, on the one hand, is what actually motivates the seeker firm to adopt crowdsourcing to conduct distant search (Afuah and Tucci 2012). On the other hand, the limitations of the seeker firm's knowledge also implies that the seeker firm may lack the capability to identify and utilize external solvers' knowledge (Cohen and Levinthal 1990). The knowledge base of the firm has been identified as the key determinant of its ability to exploit knowledge from a variety of external sources, for instance, online communities (Huang et al. 2022) and open platforms (Niculescu et al. 2018). In the crowdsourcing context, the seeker firm's (limited) knowledge regarding the domains relevant to the crowdsourced problem determines the mental representation of the solution space (Sanyal and Ye 2023). In turn, the firm's inaccurate mental representation may lead them to misevaluate solvers' parallel search and eventually affect how well the seeker firm leverages the parallel path effect.

In sum, the seeker firm conducts large scale parallel distant search on the solution landscape via crowdsourcing. Although the seeker firm could enhance solvers' parallel search by strategically designing crowdsourcing, it may still fail to leverage the parallel path effect as expected due to its limited knowledge. A re-evaluation of crowdsourcing that incorporates the seeker firm's limited knowledge thus becomes important.

3.2.2 Firm's Knowledge in Crowdsourcing

The prior literature on crowdsourcing has hinted at how the seeker firm's knowledge may affect solvers' parallel search via interactions between the seeker firm and solvers. First, when launching the crowdsourcing contest, the seeker firm describes the crowdsourced problem by articulating what it wants the solution solvers to achieve. The information provided by the seeker firm determines what constitutes a "good" solution and will inevitably affect solvers' search (Jiang et al. 2020, Zaggl et al. 2023). Second, throughout the contest, the seeker firm can provide ongoing feedback to incentivize solvers as well as to reduce the uncertainty of solvers' in-progress search (Jiang et al. 2021, Sanyal and Ye 2023, Mihm and Schlapp 2019, Jian et al. 2019, Wooten and Ulrich 2016). On-going feedback, similar to the initial specification of the crowdsourced problem, does not objectively reflect the performance of solvers' contemporaneous solutions but is noisy and subjective to the seeker firm's private knowledge and preferences (Jiang et al. 2021, Mihm and Schlapp 2019). When the seeker firm opts for crowdsourcing for knowledge beyond its existing knowledge, feedback may inadvertently transmit the bias sourced from the firm's limited knowledge to the solvers. Last, at the completion of the crowdsourcing contest, the seeker firm evaluates all submitted solutions to select the best one from the pool. This process, however, is influenced not only by the absolute solution quality but also by how the seeker firm perceives and evaluates the submitted solutions. The seeker firm tends to filter out solutions that capture distant knowledge beyond its local base (Piezunka and Dahlander 2015), despite the potential of such solutions to be of high quality (Jeppesen and Lakhani 2010). The seeker firm may evaluate solutions according to different (and potentially imperfect) criteria due to a lack of knowledge.

Crowdsourcing can thus be conceptualized as a system influenced by the seeker firm's (limited) knowledge in all aspects. The system's ability to obtain high quality solution is also constrained by the seeker firm's (limited) knowledge. For example, concrete execution guidelines in the problem specification have been shown to be more useful than high-level conceptual objectives for crowdsourcing outcomes (Jiang et al. 2020). However, the seeker firm might lack knowledge about the detailed problem-solving process and can only provide general objectives about possible solutions in the problem specification at the launch of crowdsourcing. In a similar vein, the seeker firm might not have sufficient knowledge about the interdependencies among different design elements and fail to effectively break down the problem into separable modules that are easier to tackle (Afuah and Tucci 2012, Hu and Wang 2020). The firm's in-progress feedback and final selection can thus become misleading for solvers as it conditions on some hidden contingencies which are unbeknown to the solvers. Moreover, the impact of the seeker firm's (limited) knowledge on interactions between the seeker firm and solvers, albeit in isolation, paints a systems view of crowdsourcing. The impact of the seeker firm's (limited) knowledge on various interactions may influence each other, leading to unexpected system level outcomes. Feedback, for instance, may navigate solvers from searching distant domains to those more familiar to the seeker firm (Park et al. 2022, Sanyal and Ye 2023). Yet, considering the seeker firm's tendency to select solutions that are more familiar to them (Piezunka and Dahlander 2015), providing feedback to solvers may hinder solvers' exploration of the solution landscape and at the same time, increases the likelihood that the seeker firm will be satisfied with the familiar solutions. Put differently, the crowdsourcing system may end up with a scenario where the seeker firm constantly acquires solutions that it deems satisfactory but are of low quality. It is thus crucial to understand the performance of the crowdsourcing system and more importantly, how to overcome the potential suboptimality caused by the seeker firm's limited knowledge about the solution space.

3.2.3 Learning from the Crowd

As discussed above, problems arising from the outflow of knowledge from the seeker firm to the crowd of solvers has been hinted at from different aspects. In contrast, issues related to the inflow of knowledge from solvers to the seeker firm remains relatively underexplored in the extant crowdsourcing literature which has predominantly taken the selection of the winning solution as the primary approach through which the seeker firm intakes solvers' knowledge (e.g., Mo et al. 2021, Mihm and Schlapp 2019). Yet, as the scope and knowledge requirements of the crowdsourced problem expand, it becomes unlikely for any single solver to possess the full range of knowledge required to effectively solve the problem (Gurca et al. 2023, Majchrzak and Malhotra 2020). In turn, solvers' limited knowledge impedes the quality of their submitted solutions, even if one is chosen to be the winner among the submissions. As such, acquiring knowledge via adopting the winning solution could be ineffective for the seeker firm. Whilst a body of research has attempted to address this issue by encouraging more solvers to participate in crowdsourcing (Boudreau et al. 2011, Liu et al. 2014, Jian et al. 2019), scant attention has been paid to acquiring knowledge from all submitted solutions rather than from only the winning solution.

The crowd, as a collective of solvers with diverse knowledge, exhibits its own wisdom (Surowieck 2004). The crowd's wisdom hinges on each solver making independent decisions as per its own knowledge. When individual decisions from different aspects are aggregated, it is expected that the errors entailed in individual decisions would cancel each other out, leading to more accurate crowd-level decisions (Muller-Trede et al. 2017, Da and Huang 2020). It has been shown to reach comparable or even superior performance compared to experts in various tasks such as funding decisions (Mollick and Nanda 2016), earning forecasts (Da and Huang 2020), and hedonic judgments (Muller-Trede et al. 2017). Applying the same logic, although the seeker firm may not comprehend all aspects of its problem, it could learn about what solution elements should be taken and how these elements should be organized in these aspects to constitute a "good" solution from all solvers' submissions. For example, in the Netflix Prize contest, Netflix gained insights into algorithms with high performance by examining the source code submitted by solvers (Amatrian and Basilico 2012). In this way, the seeker firm can upweight (downweight) the quality of a solution if it has incorporated (missed) the solution elements considered by other solvers, even if the seeker firm may lack a nuanced understanding of them. The prior literature on crowdsourcing has implied such opportunities for the seeker firm to learn from the crowd of solvers and update its understanding of the solution space. When all solvers submit their solutions to the seeker firm for in-progress feedback (Jian et al. 2019, Jiang et al. 2021) and final selection (Piezunka and Dahlander 2015, Park et al. 2022), the solution seeker is able to observe all solutions along with how various solution elements are included and considered by solvers and aggregate them into an updated mental model. In some crowdsourcing practices, such as two-stage contests (Hou and Zhang 2021, Liu and Kim 2023), the seeker firm can review all submitted solutions and select eligible

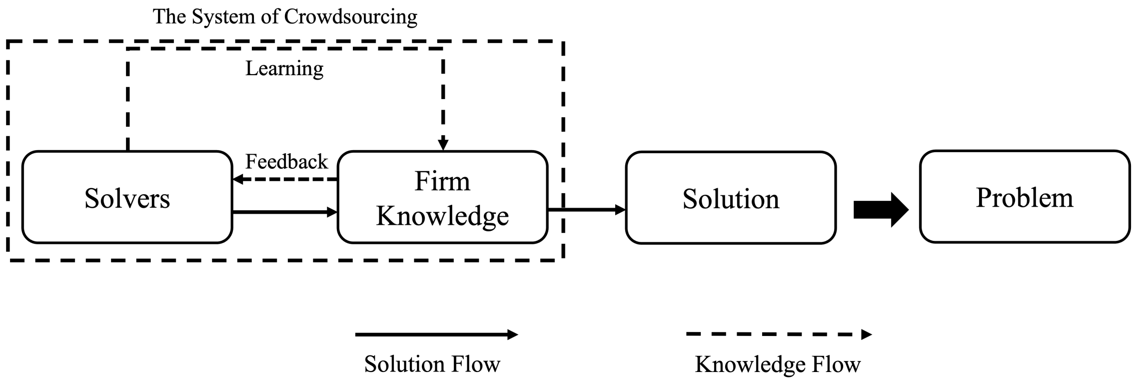

Figure 3.1: A Systems View of Crowdsourcing

finalists for the second round contest. The seeker firm can also refine its evaluation and selection criteria accordingly during the initial first round selection.

Learning from the crowd, however, does not occur in isolation either. Whether and to what extent it could overcome the potential suboptimality caused by the seeker firm's limited knowledge depends on other flows in the crowdsourcing system. The seeker firm's intermittent feedback, for instance, has been shown to be associated with solution diversity (Sanyal and Ye 2023). In turn, solution diversity determines how a solution element considered by the crowd of solvers has been validated from diverse perspectives, affecting the effectiveness of the seeker firm's learning.

We summarize the aforementioned interactions between the seeker firm and solvers that have been constrained by the seeker firm's knowledge as well as learning from the crowd of solvers in Figure 3.1. As indicated by the solid arrows, all solvers' solutions will be evaluated and selected by the seeker firm according to its knowledge and then applied to solve the crowdsourced problem. The dashed arrows show how knowledge flows within the crowdsourcing system. On the one hand, the seeker firm's knowledge determines what to be crowdsourced and how it would provide feedback to solvers. On the other hand, the seeker firm can also learn from the wisdom of the crowd about what constitute a "good" solution. As discussed, satisfying the seeker firm is merely a subcomponent of the crowdsourcing system and does not guarantee producing high quality solutions from the system. To the best of our knowledge, this study is among the first to distinguish crowdsourcing effectiveness from seeker firm satisfaction and investigate crowdsourcing effectiveness from a systems view.

3.3 Methodology

In this study, we use a simulation to investigate crowdsourcing for two reasons. First, as shown in Figure 3.1, the focus of this study is to examine crowdsourcing effectiveness from a systems view. However, the ground truth is difficult to establish because most observational data only captures whether the seeker firm is satisfied with crowdsourced solutions within the crowdsourcing system. Simulation allows us to investigate crowdsourcing when observational data is accessible (Miranda et al. 2022). Second, we attempt to examine the complex interactions among the seeker firm and solvers and how these interactions impact the overall system performance. Simulation enables us to observe the emergence of system-level patterns from micro-level interactions (Davis et al. 2007). We follow the simulation-based theory development process as per Davis et al. (2007). We draw from the theoretical foundations discussed above to choose the simulation model and create the computational representations for each theoretical constructs in the crowdsourcing system. The developed model is verified against prior research findings and with robustness checks. A series of simulation experiments are conducted to generate theoretical insights from a systems view.

Specifically, we model a crowdsourcing contest consisting of three entities – a problem with a certain level of complexity, a number of solution solvers who independently search for solutions, and a seeker firm who selects the winning solution at the end of the contest. During the contest, the seeker firm can also guide solvers' search process by providing feedback on solvers' contemporaneous solutions as well as learn from the crowd. Each entity will be discussed in detail below.

3.3.1 The *NK* Fitness Landscape Model

As we discussed above, crowdsourcing is essentially a process through which the seeker firm conduct large scale parallel distant search on the solution landscape. Therefore, we choose the *NK* fitness landscapes model (Kauffman 1993, Levinthal 1997, Rivkin and Siggelkow 2003) as the analytical apparatus for our study.

The *NK* fitness landscapes model creates a multidimensional space (i.e., the fitness landscape) that includes N elements and K interdependencies among these elements. Each configuration of N decisions $\mathbf{d} =< d_1, d_2, \ldots, d_N >$ is associated with a fitness value $F(\mathbf{d})$ (i.e., the altitude on the landscape), which is calculated as the average of the fitness contribution c_i of every element d_i, where c_i depends not only on the specification of the focal element d_i but also on K other interdependent elements – i.e., $F(\mathbf{d}) = \Sigma_{i=1}^{N} c_i(d_i|K \text{ other } d_j\text{'s})/N$, where c_i is randomly drawn from a uniform distribution $U(0,1)$.

By adjusting the parameter K for a given N, the NK fitness landscapes model enables the researcher to tune the ruggedness of the landscape (i.e., the complexity of the problem). When there are few interdependencies among elements (i.e., low K), the generated landscapes are smoother, whereas the landscape becomes more rugged if elements are tightly coupled with each other (i.e., high K).

3.3.2 Crowdsourcing Problem

Following convention, we model a crowdsourcing problem consisting of N elements $d_1, d_2, \ldots, d_N$ that need to be configured by solvers, where each element d_i can take the value of 0 or 1. The d_i's represent decisions on a particular knowledge domain i. For example, when the seeker firm crowdsources its internal problem and data to external solvers for predictive algorithms (e.g., Lee et al. 2018, Jin et al. 2021), the elements (d_i) of the crowdsourcing problem can include the choice of the analytical method (e.g., deep learning-based model vs. causality model), choice of computing power (e.g., local vs. cloud clusters), and the adoption of available exemplar (e.g., yes vs. no), etc. The elements can be interdependent such that the fitness contribution of a particular choice for an element could also depend on the choice of other decisions. For instance, to develop an effective algorithm, using cloud-based computation is more beneficial when a deep learning model is adopted given the scalability of cloud computing. Formally, a crowdsourcing problem is represented by a space consisting of 2^N possible solutions, where each solution is an N bit configuration: $\mathbf{d} =< d_1, d_2, \ldots, d_N >$. Each location in the space (i.e., one particular solution configuration) is associated with a fitness value representing the configured solution's performance. The elements can be interdependent such that the fitness contribution of a particular choice for an element could also depend on the choice of other K decisions.

3.3.3 Solvers

3.3.3.1 Knowledge Set

Solvers differ in the knowledge they possess. For instance, NASA used to post its internal challenges online to collaborate with domain experts from a variety of technological fields, including material science, clinical medicine, mechanical engineering, and microbiology (Lifshitz-Assaf 2018). Therefore, solvers are modeled as agents who are capable of determining the values of a subset of the configuration space ($\mathbf{d}$). Formally, each agent in the crowd holds a knowledge set D_c, where $D_c \subset \{1, 2, \ldots, N\}$ and $|D_c| = N_c$. Elements in D_c index the elements in the knowledge domain of an agent and as a result, the agent can configure.

Because of limited knowledge outside their domains, solvers' search for solutions is imperfect and can only reflect those aspects within their knowledge domains. We model it as search on simplified cognitive representations of the actual landscape (Gavetti and Levinthal 2000, Hahn and Lee 2021). Specifically, a solver would perceive a solution $\mathbf{d}$ as $\mathbf{d}_{D_c}$, where the subscript D_c represents those d_i are within this solver's knowledge set D_c. The solver can only search and evaluate $\boldsymbol{d}_{D_c}$ given unknown elements outside D_c. Simply put, solvers abstract the actual landscape into a smoother version where some peaks and valleys are ignored due to a lack of nuanced understanding of that area (i.e., simplified cognitive representation). On the abstracted landscape, fitness values generally represent how well configurations work across unknowns. Formally, a solver would perceive a solution $\boldsymbol{d} =< \boldsymbol{d}_{D_c}, \boldsymbol{d}_{D_c^C} >$ as with fitness value $\tilde{F}_c(\boldsymbol{d}) = \mathbf{E}[F(\boldsymbol{d})|\boldsymbol{d}_{D_c^C}]$, where $\tilde{F}_c(\boldsymbol{d})$ represents the quality of a solution $\boldsymbol{d}$ perceived by a solver, $\boldsymbol{d}_{D_c{}^C}$ represents those d_i that are outside of D_c (i.e., in the complement of D_c), and $\mathbf{E}[\,]$ represents the expected fitness across unknown decisions. For example, suppose $N = 6$ and $N_c = 4$, if a solver is only knowledgeable about the first four design elements (i.e., $D_c = \{1, 2, 3, 4\}$), then one of its solutions could be represented as $\boldsymbol{d} =< 1, 0, 1, 0, *, * >$, where $*$ represents unknown elements. The solver's perceived value towards this solution can be calculated as the average fitness value across unknowns. Namely, the average fitness value of $\boldsymbol{d} =< 1, 0, 1, 0, \mathit{0}, \mathit{0} >$, $< 1, 0, 1, 0, \mathit{0}, \mathit{1} >$, $< 1, 0, 1, 0, \mathit{1}, \mathit{0} >$, and $< 1, 0, 1, 0, \mathit{1}, \mathit{1} >$.[1]

[1]This modeling approach does not mean that a solver actually knows the fitness values of

At the beginning of the simulation, *CrowdSize* solvers are randomly seeded on the generated landscape. Each solver starts search from a configuration $\boldsymbol{d} =< d_1, d_2, \ldots, d_N >$ with all elements in $\boldsymbol{d}_{D_c}$ being initialized as 0 or 1 with equal probability for a randomly generated knowledge set D_c according to N_c. Solvers' solutions to configuring $\boldsymbol{d}_{D_c}$ will be evaluated by the seeker firm during crowdsourcing.

3.3.3.2 Search with Performance Feedback

During each iteration, solvers randomly select one element d_i from their knowledge set D_c and change its value from 0 to 1 or vice versa to propose an alternative configuration $\boldsymbol{d'}$. For the current configuration $\boldsymbol{d}$ and the proposed alternative $\boldsymbol{d'}$, where the Hamming distance between $\boldsymbol{d}$ and $\boldsymbol{d'}$ in D_c is 1, solvers would adopt $\boldsymbol{d'}$ if $\tilde{F}_c(\boldsymbol{d'}) > \tilde{F}_c(\boldsymbol{d})$ or remain with status quo $\boldsymbol{d}$, otherwise.

Solvers' search is sensitive to the seeker firm's feedback. For instance, if submitted solutions are rated as of high quality by the seeker firm, solvers would refrain from redesigning the whole solution but would rather make minor modifications and improvements to their current solutions (Jiang et al. 2021). Our simulation model captures this process as follows. When solvers receive feedback F_* from the firm, they would choose an alternative $\boldsymbol{d'}$ if $F_*(\boldsymbol{d'}) \geq F_*(\boldsymbol{d})$ or remain with the status quo $\boldsymbol{d}$, otherwise. After receiving feedback, solvers would update their understanding of the tested solution $\boldsymbol{d}$ such that $\tilde{F}_c(\boldsymbol{d})$ is replaced by $F_*(\boldsymbol{d})$. In this way, when solvers receive positive (negative) feedback from the firm, they are less (more) likely to change the current configuration.

3.3.4 Seeker Firm

3.3.4.1 Knowledge Space

As discussed ealier, the seeker firm may possess limited knowledge regarding the domains outside its existing knowledge. Therefore, we model the seeker firm also as an agent with imperfect knowledge about the solution space. Similar to solvers,

all solutions and is able to calculate the average fitness values across configurations of unknown elements. Instead, the average value is simply an unbiased expected value that a solver would perceive a solution configuration given its imperfect knowledge. The perceived fitness values of different configurations may be higher or lower than the actual fitness values due to a lack of knowledge. Yet, how the agent perceives all configurations on the landscape as a whole would be largely correlated with the actual fitness landscape.

the seeker firm is knowledgeable about the value of a subset of a configuration $\boldsymbol{d} =< d_1, d_2, \ldots, d_N >$. Formally, the seeker firm holds a knowledge set D_f, where $D_f \subset \{1, 2, \ldots, N\}$ and $|D_f| = N_f$. Elements in D_f index the elements in the knowledge domain of the seeker firm and as a result, it can assess. The seeker firm would perceive a solution as $\boldsymbol{d} =< \boldsymbol{d}_{D_f}, \boldsymbol{d}_{D_f^C} >$, where the subscript D_f represents d_i is from the firm's knowledge set D_f. The seeker firm would perceive a solution $\boldsymbol{d}$ as with quality $\tilde{F}_f(\boldsymbol{d}) = \mathbf{E}[F(\boldsymbol{d})|\boldsymbol{d}_{D_f^C}]$, where D_f^C represents d_i is outside D_f, and $\mathbf{E}[\,]$ represents the expected fitness across unknown decisions.

3.3.4.2 Solution Evaluation

The seeker firm has a randomly generated initial configuration $\boldsymbol{d}$. If a solver submits a solution $\boldsymbol{d}_{D_c}$ to the seeker firm, the seeker firm would hypothetically adopt the elements that have been configured in $\boldsymbol{d}_{D_c}$ and evaluate the integrated configuration $\boldsymbol{d_{int}}$ according to its knowledge set D_f. For example, suppose the seeker firm's initial configuration $\boldsymbol{d} =< 1, 0, 1, 0, 1, 0 >$, if a solver is knowledgeable about the first four design elements (i.e., $D_c = \{1, 2, 3, 4\}$) and submits a solution $\boldsymbol{d_{D_c}} =< 1, 1, 1, 1, *, * >$ to configure the first four elements, the seeker firm would evaluate $\boldsymbol{d_{D_c}}$ according to its perceived value towards $\boldsymbol{d_{int}} =< 1, 1, 1, 1, 1, 0 >$.

3.3.4.3 Feedback

During each iteration, the seeker firm has a probability of $p_{feedback}$ to give feedback to solvers during their parallel search. We model the firm's feedback towards a submitted solution $\boldsymbol{d}_{D_c}$, F_*, as the firm's perceived value $\tilde{F}_f(\boldsymbol{d_{int}})$, where $\boldsymbol{d_{int}}$ is the integrated solution based on the seeker firm's initial configuration $\boldsymbol{d}$ and the submitted solution $\boldsymbol{d}_{D_c}$.

3.3.4.4 Learning from the Crowd

The seeker firm can learn from solvers throughout crowdsourcing by enumerating solvers' solutions and analyzing what constitutes a "good" solution from the solvers' perspectives. If most submitted solutions share the same configuration of certain design elements, the seeker firm could view the configuration as ground truth and devalue a solution that has not taken the same configuration, even if the seeker firm may not possess the knowledge to assess the configuration accurately. To incorporate the seeker firm's learning from the crowd, we allow the seeker firm to update its

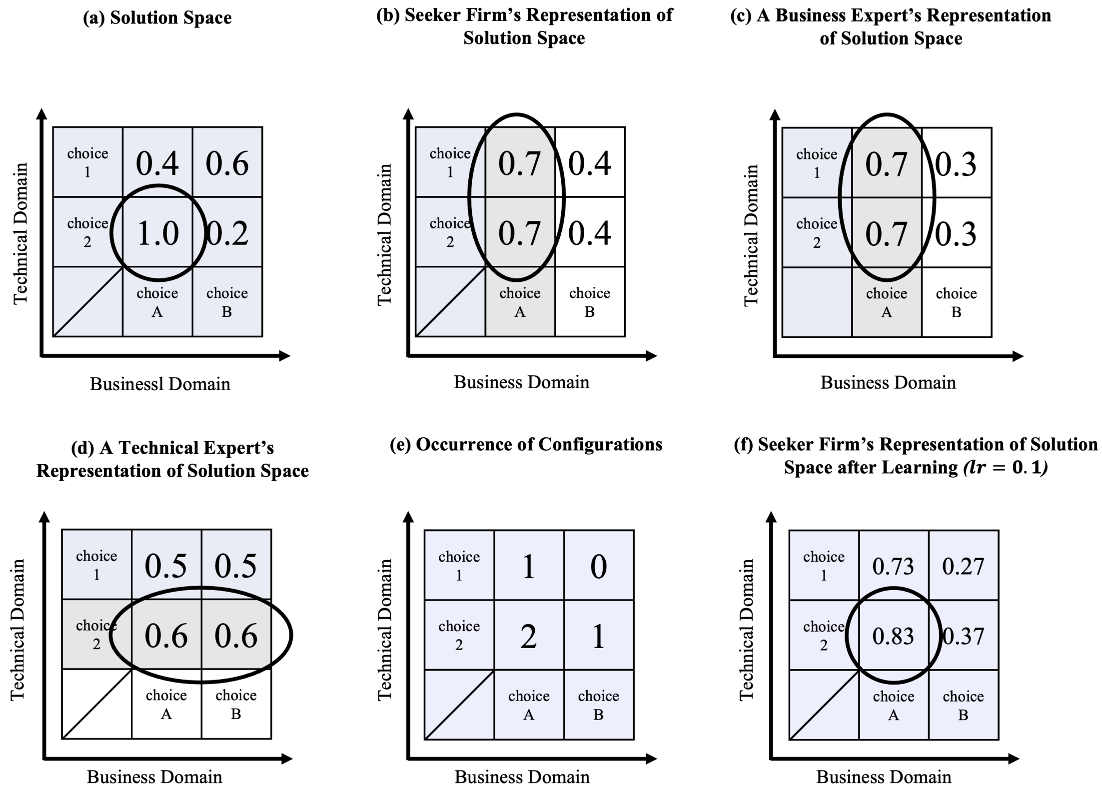

Figure 3.2: Illustration of the Updating of the Seeker Firm's Perceived Values

Notes: Figure 3.2a represents absolute fitness values of different configurations; Figure 3.2b-d represent how the seeker firm and solvers perceive the fitness landscape according to their knowledge bases; the perceived values are calculated as the average of absolute fitness values across unknown domains; Figure 3.2e shows the occurrence of different configurations; Figure 3.2f displays how the seeker firm updates its perceived values towards different configurations according to their occurrences.

evaluation of solvers' solutions. Specifically, the seeker firm will count the occurrence of solvers' solutions $\boldsymbol{d}_{D_c}$ to configuring certain elements in all submitted solutions. If a configuration $\boldsymbol{d}_{D_c}$ has been used for *freq* times in solvers' submissions, then for a solution $\boldsymbol{d}$ that contains the configuration $\boldsymbol{d}_{D_c}$, its perceived value is updated according to a learning rate lr. Formally,

$$\tilde{F}_f(\boldsymbol{d})_{afterLearning} = (1 - lr) \times \tilde{F}_f(\boldsymbol{d})_{beforeLearning} + lr \times freq$$

Learning from the crowd is easiest seen with an example (see Figure 3.2).[2]

[2]According to the *NK* landscapes model, the real landscape is a multidimensional space and thus cannot be visualized. These figures are simplifications of $N + 1$ dimensional spaces (i.e., N design elements and one additional dimension for fitness values) for illustration purposes. Values

Suppose a firm takes the crowdsourcing approach to develop a trading algorithm and there are, for simplicity, two domains that need to be configured to develop an effective solution: the first domain pertains to the business context and determines what evaluation metrics should be used to align with the firm's specific environment, for instance, the maximal drawdown and the frequency of trade (i.e., choices A and B in Figure 3.2); the second domain is more technical and deals with what algorithms should be selected to build the trading solution. The choice of machine learning, ensemble learning, and deep learning falls into this category (i.e., choices 1 and 2 in Figure 3.2). As highlighted by the circle in Figure 3.2a, there is a best combination of evaluation metrics and algorithms (i.e., a global optimum; the center cell associated with a fitness value = 1) that contributes to high quality solutions. However, the seeker firm only possesses knowledge about the business domain (as in Figure 3.2b) and as a result, can only perceive solutions according to the business domain, regardless of how the technical domain is configured. As shown by Figure 3.2b, the seeker firm perceives no difference between two choices made in the technical domain. Instead, it evaluates choices in the business domain based on their performance in general across unknowns (e.g., the perceived value of choice A $= \frac{1}{2} \times (0.4 + 1.0) = 0.7$). Consequently, as highlighted by the circle in Figure 3.2b, the seeker firm would prefer solutions that take "choice A" (i.e., using the maximal drawdown to evaluate the trading solution) in the business domain. Suppose one solver from the business domain and one solver from the technical domain join the crowdsourcing and configure their solutions independently according to their specific knowledge domains. As highlighted by the circle in Figures 3.2c and d, the business expert will submit a solution with the business domain configured as "choice A" because it has the highest value perceived by the expert. Similarly, the technical expert will submit a solution with the technical domain configured as "choice 1" (i.e., using deep learning methods). The seeker firm could count the frequency of different configurations being used by solvers. As shown in Figure 3.2d, configurations with "choice A" and "choice 1" are used by both solvers. A solution with both "choice A" and "choice 1" thus will have the highest frequency $freq = 2$. The seeker firm could update its perceived values towards this solution according to a learning rate

in cells represent fitness values (solution quality).

$lr = 0.1$ (as in Figure 3.2e). For the above solution with both "choice A" and "choice 1", its perceived value will be replaced with $0.83(= 0.7 \times 0.9 + 2 \times 0.1)$. In this sense, the seeker firm is able to utilize both its internal knowledge base and the wisdom of the crowd to find the global optimum.

3.3.4.5 Solution Adoption

At the completion of crowdsourcing, all solvers will complete their search and submit their solutions $\boldsymbol{d}_{D_c}$ to the seeker firm. The seeker firm will hypothetically integrate $\boldsymbol{d}_{D_c}$ into its configuration to evaluate all solutions and adopt the one with the highest perceived value $\tilde{F}_f(\boldsymbol{d_{int}})$.

Table 3.1 summarizes the parameters involved in our simulation model and the performance metrics we use to evaluate crowdsourcing effectiveness.

Table 3.1: Overview of Simulation

Crowdsourcing Problem	
N	The number of solution elements
K	The number of interdependent elements for each focal element
Solvers	
$CrowdSize$	The number of solvers
N_c	The number of solution elements solvers can configure
Seeker Firm	
N_f	The number of solution elements the seeker firm is knowledgeable about
$p_{feedback}$	The likelihood of give feedback to solvers during each iteration
lr	The extent to which the seeker firm updates its perceived values towards various configurations according to the wisdom of the crowd
Performance Measure	
$\tilde{F}_f(\mathbf{d})$	The seeker firm's perceived quality of solutions
$F(\mathbf{d})$	The absolute quality of solutions

3.4 Simulation Results

We use the model described above to run four sets of computational experiments that each examine different aspects of crowdsourcing from a systems perspective. For each simulation experiment, we adopt the following strategy. We fix $N = 14$ and vary K from 0 to 13 (i.e., the full range of complexity values when $N = 14$). We set the size of the solvers' knowledge set N_c as 7. In other words, each solver

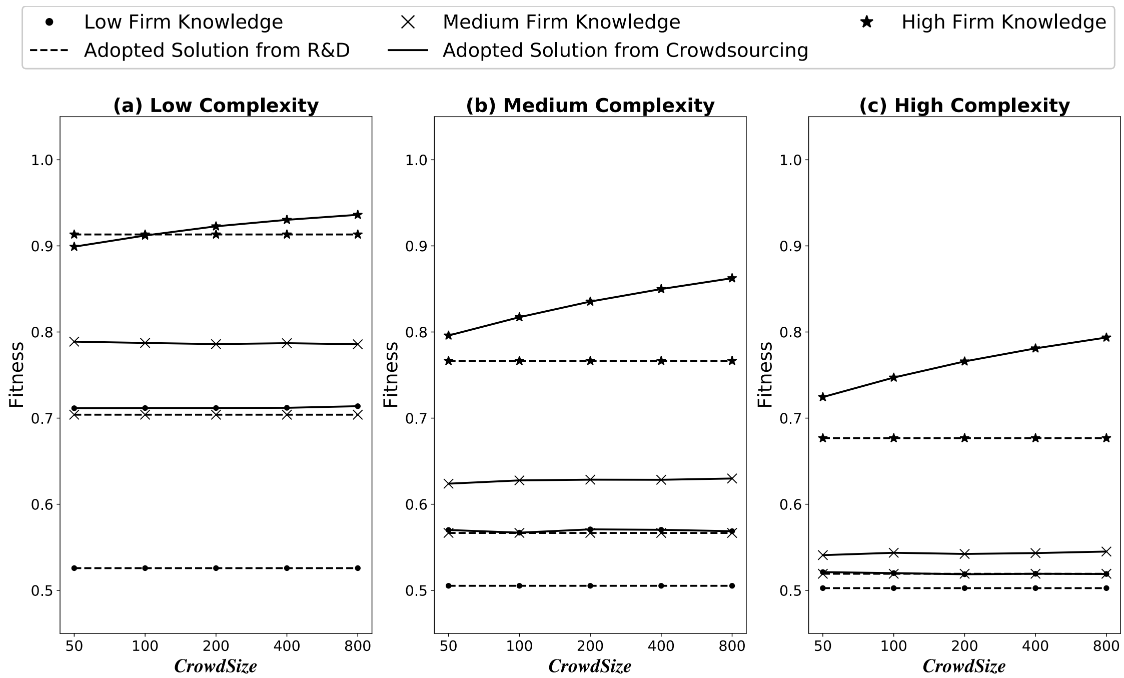

Figure 3.3: Crowdsourcing vs. R&D

Notes: For brevity, Figure 3.3 only includes the results for combinations of problem complexity (K) at low, medium and high levels (across columns of figures) and the seeker firm's knowledge (N_f) at low (circle markers), medium (cross markers), and high (star markers) levels – i.e., $K = \{1, 7, 13\}$ and $N_f = \{1, 7, 13\}$.

is knowledgeable about and therefore can change values for half of the possible elements. We run the simulation for 100 periods (when it reaches stable state). All experiments are seeded with 20,000 replications, each facing a unique landscape randomly generated according to the given N and K. The results reported below are statistically significant with a 99% confidence level.

3.4.1 Experiment 1 – Solution Flow: Crowdsourcing vs. R&D

We begin our analyses by examining the effectiveness of the solution flow of crowdsourcing versus the seeker firm's internal R&D.[3] We disallow the seeker firm to provide feedback or learn from the crowd for this baseline experiment.

Figure 3.3 shows the absolute quality of solutions (i.e., $F(\boldsymbol{d})$) developed from

[3]The internal R&D process is modeled as the seeker firm's iterative and incremental search. The seeker firm modifies one design element during each iteration and adopts the proposed configuration if it is perceived as better than the status quo, so on and so forth.

internal R&D (dashed lines) and crowdsourcing (solid lines). Our analyses are based on $980 = (14 \times 5 \times 14)$ experimental conditions under which the seeker firm may possesses 14 levels of knowledge (i.e., $N_f = \{1, \ldots, 14\}$; as indicated by the circle, cross, and star markers) and run crowdsourcing with 5 levels of number of solvers (i.e., $CrowdSize = \{50, 100, 200, 400, 800\}$; across the x-axis of each subplot) at 14 levels of problem complexity (i.e., $K = \{0, \ldots, 13\}$; across the columns in Figure 3.3). Not surprisingly, as the classical *NK* model would predict, both internal R&D and crowdsourcing suffer a performance decline as problem complexity increases. The absolute quality of the solution from internal R&D when the seeker firm has high level of knowledge (dashed lines with star markers), for instance, decline from 0.91 in low complexity problems to 0.68 in high complexity problems. Meanwhile, we also find evidence for the parallel path effect – i.e., involving as many solvers as possible addresses the uncertainty in complex problem solving and thus enhances crowdsourcing effectiveness. As shown by the solid lines with star markers, the absolute quality of the adopted solution from crowdsourcing increases as the number of solvers increases. Comparing the slopes of the solid lines with star markers in problems with different complexity, we find that the parallel path effect is most impactful for complex problems. The net benefit of increasing the number of solvers from 50 to 800 is around 2% (10%) in low (high) complexity problems. These findings are consistent with prior literature on crowdsourcing (e.g. Boudreau et al. 2011) and lends support for validity of our modeling approach.

However, to what extent the seeker firm can utilize the parallel path effect depends on its knowledge. The seeker firm is able to utilize the parallel path effect when it has a high level of knowledge (see the positive slope of solid lines with star markers across solver number). However, as the seeker firm's knowledge level decreases, merely involving more solvers does not guarantee crowdsourcing effectiveness. As shown by the solid lines with circle and cross markers, when the seeker firm does not possess sufficient knowledge, the absolute quality of the solutions from crowdsourcing adopted by the seeker firm does not increase as the number of solvers increases.

Figure 3.4 shows the quality of the best solution from crowdsourcing (which the seeker firm may fail to adopt) in the dashed lines. As suggested by the parallel path effect, the crowd of solvers manages to find solutions of higher quality as the

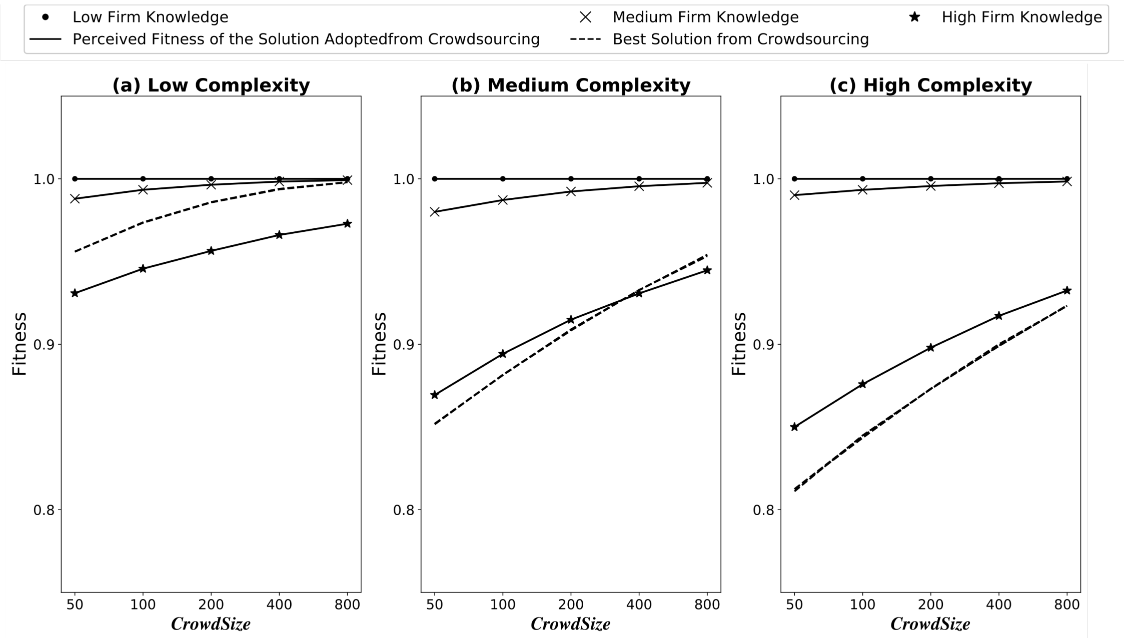

Figure 3.4: How Firm's Knowledge Impacts the Parallel Path Effect

Notes: The perceived quality was normalized to the range $[0, 1]$ within each level of complexity. It therefore captures the extent to which the best solution in the firm's mental representation of the solution space has been discovered.

number of solvers increases. Figure 3.4 also shows the perceived quality of the adopted solution from crowdsourcing (i.e., $\tilde{F}_f(\boldsymbol{d})$) in the solid lines. We find that the perceived quality is negatively associated with the seeker firm's knowledge. The slope of the solid lines across the number of solvers gradually decreases as the seeker firm's knowledge level declines. When the seeker firm has insufficient knowledge (as in the solid lines with cross and circle markers), it can only appreciate and evaluate limited aspects of the submitted solutions and as a result, becomes more likely to be satisfied with the solvers' solutions. This, in turn, hinders the seeker firm benefits from involving as many solvers as possible and makes the parallel path effect under-leveraged.

These results suggest a paradox that has yet to be carefully addressed in the crowdsourcing literature. On the one hand, crowdsourcing has the greatest potential compared to internal R&D for the seeker firm when the seeker firm does not possess sufficient knowledge about the solution space. The quality difference between the best solution from crowdsourcing (see dashed lines in Figure 3.4) and the solution from internal R&D (see dashed lines in Figure 3.3) is greater when the seeker firm possesses less knowledge. On the other hand, when crowdsourcing has the greatest potential (i.e., when the seeker firm does not possess sufficient knowledge about the solution space), the seeker firm fails to fully leverage this potential specifically due to the lack of knowledge. The quality difference between the adopted solution from crowdsourcing (see solid lines in Figure 3.3) and the best solution from crowdsourcing (see dashed lines in Figure 3.4) is also greater when the seeker firm possesses less knowledge. These results suggest that when the seeker firm turns to crowdsourcing for solutions that require knowledge beyond its existing base, it also lacks the knowledge to identify and utilize the diverse knowledge possessed by the crowd to its full potential. The conventional wisdom of crowdsourcing – i.e., the parallel path effect, however, fails to resolve this paradox as discussed above. We summarize these findings with the following paradox:

Paradox 3.1. When the seeker firm has insufficient knowledge about the solution space, crowdsourcing is most needed but also least utilized.

3.4.2 Experiment 2 – Knowledge Flow: Feedback in Crowdsourcing

Experiment 2 studies the knowledge flow from the seeker firm to the solvers, namely how the seeker firm's feedback affects the absolute quality of the adopted solution. We fix the number of solvers at 100 and allow the seeker firm to provide feedback to solvers at certain probabilities during crowdsourcing. $980 (= 14 \times 14 \times 5)$ experimental conditions were analyzed for all possible combinations of firm knowledge levels (i.e., $N_f = \{1, \ldots, 14\}$), problem complexity (i.e., $K = \{0, \ldots, 13\}$), and feedback probabilities (i.e., $p_{feedback} = \{0, 0.25, 0.50, 0.75, 1\}$).[4]

Figure 3.5 presents the absolute quality of the solutions adopted by the seeker firm (i.e., $F(\boldsymbol{d})$) across feedback probabilities. As shown by the solid lines with star markers, when the seeker firm has a high level of knowledge, providing feedback to solvers tends to increase the absolute quality of the adopted solution (i.e.,

[4] We also experimented with different numbers of solvers and found no qualitative differences. In addition to the extent to which the seeker firm provides feedback to solvers, prior studies have also suggested the moderating effect of the timing of feedback (e.g. Jiang et al. 2021, Sanyal and Ye 2023) on the impact of feedback. Therefore, we also examined the timing of feedback – i.e., early and late feedback (providing feedback only in the first and second half of crowdsourcing). The qualitative results remain unchanged.

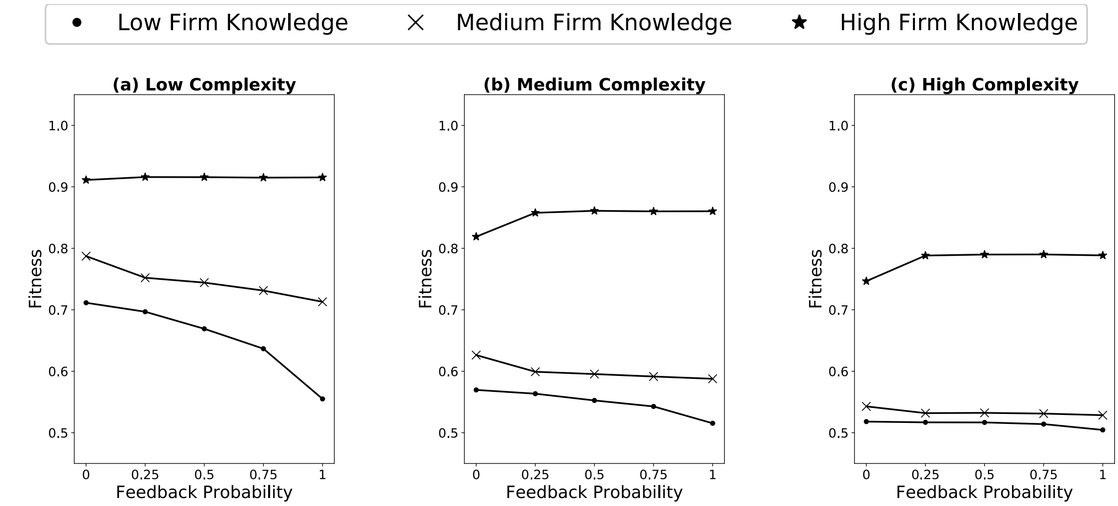

Figure 3.5: Absolute Quality of the Adopted Solution with Feedback

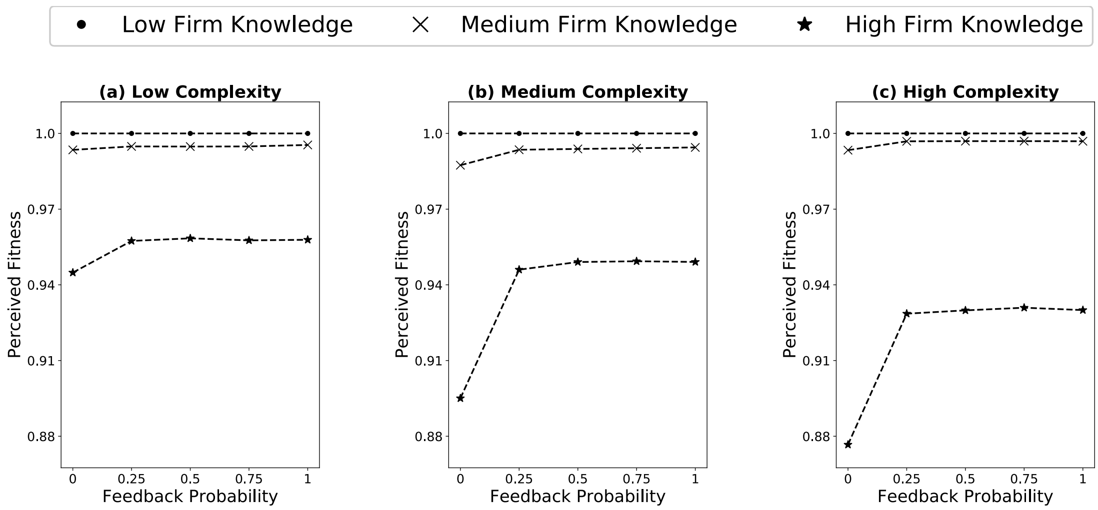

Figure 3.6: Perceived Quality of the Adopted Solution with Feedback

crowdsourcing is more effective), especially for highly complex problems. The slope of the solid line with star markers is steeper in problems with higher complexity. However, as the seeker firm's knowledge level declines, providing feedback can hinder crowdsourcing effectiveness, especially when problem complexity is low. As shown by the solid lines with circle and cross markers, providing feedback decreases the absolute quality of the adopted solutions in these cases. Taken together, these results point to the seeker firm's knowledge as a critical moderator of the effectiveness of feedback.

Next we examine how the seeker firm perceives the adopted solution from crowdsourcing. Figure 3.6 shows the perceived quality of the solutions adopted by the seeker firm (i.e., $\tilde{F}_f(\boldsymbol{d})$) across feedback probabilities. Interestingly, despite the heterogeneous impact on the absolute quality, providing feedback generally increases the quality of the adopted solution as perceived by the seeker (see the increasing trends of dashed lines with cross and star markers). In other words, the seeker firm is more satisfied with the adopted solution when it provides feedback. The only exception is when the seeker firm only has limited knowledge and as a result, can only evaluate limited aspects of the submitted solutions and is already satisfied with the adopted solution from the crowd (see dashed lines with circle markers). When combining the impact of feedback on the absolute and perceived quality of the adopted solution, we find a potential lock-in caused by the seeker firm's insufficient knowledge – the seeker firm's attempts to increase crowdsourcing effectiveness via feedback might increase the perceived quality to the detriment of the absolute quality of the adopted solution. In other words, although the seeker firm may perceive that crowdsourcing effectiveness has improved when it provides feedback, crowdsourcing effectiveness has actually worsened. These results point to a second paradox of crowdsourcing:

Paradox 3.2. When the seeker firm has insufficient knowledge about the solution space (i.e., when crowdsourcing is actually most needed), the seeker firm's feedback to solvers worsens crowdsourcing effectiveness even though firm may be more satisfied with the solutions developed via crowdsourcing.

3.4.3 Experiment 3 – Knowledge Flow: Learning from the Crowd

The previous two experiments highlight that the system of crowdsourcing is limited by the seeker firm's knowledge. More importantly, the crowdsourcing system may collapse when it produces lower quality solutions yet makes the seeker firm more satisfied. Experiment 3 aims to examine whether the other knowledge flow in the crowdsourcing system, namely the seeker firm's learning from the crowd, might help to resolve the above paradoxes. We fixed the number of solvers at 100, disallowed feedback from the seeker firm and manipulated the extent to which the

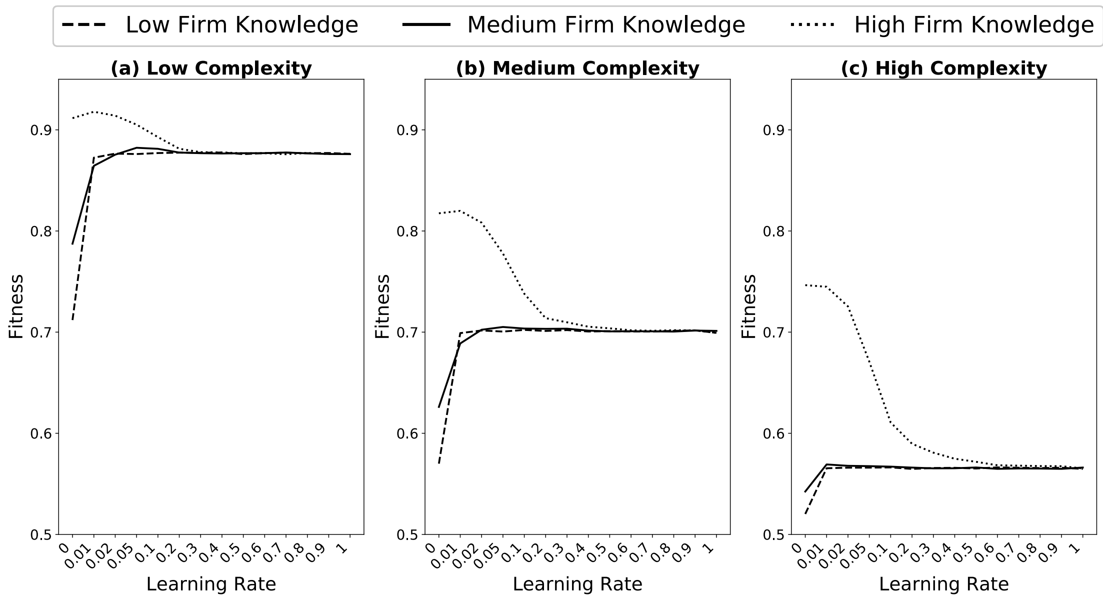

Figure 3.7: Absolute Quality of the Adopted Solution with Learning

Notes: The x-axis was rescaled to better illustrate the existence of the tipping point.

seeker firm learns from the crowd (i.e., lr). Our analyses are based on testing $19,796 = (14 \times 14 \times 101)$ experimental conditions – 14 levels of firm knowledge (i.e., $N_f = \{1, \ldots, 14\}$), 14 levels of problem complexity (i.e., $K = \{0, \ldots, 13\}$), and 101 levels of learning rate (i.e., $lr = [0, 1]$ with a step size of 0.01). Figure 3.7 shows the absolute quality of the adopted solution from crowdsourcing when the seeker firm learns from the crowd across different levels of problem complexity (across columns of figures), different levels of firm knowledge (solid, dashed, and dotted lines), and learning rates (x-axis).

We find that learning from the crowd could potentially increase the absolute quality of the adopted solution, albeit depending on the level of the seeker firm's knowledge and problem complexity. When the seeker firm has limited knowledge regarding the domains relevant to the crowdsourced problem, its existing knowledge base is less capable of evaluating solutions developed by the crowd. As a result, the seeker firm should mostly rely on the judgments made by the solvers about what constitutes a "good" solution and evaluate solutions accordingly. As shown by the dashed lines, the seeker firm is able to find solutions of higher quality as long as it learns from the crowd (i.e., when $lr > 0$). As the seeker firm's knowledge level increases, the benefit of learning from the crowd diminishes. Only a modest extent of learning (see the dotted line when $lr < 0.05$) can help the seeker firm obtain higher quality solutions. We also find that the optimal extent of learning from the crowd (i.e., lr) decreases as the seeker firm's knowledge level increases. For instance, the solid and dotted lines in Figure 3.7a show that the optimal extent of learning decreases from 0.05 to 0.01 when the seeker firm's knowledge declines from a high to a medium level. These results suggest that as the seeker firm's knowledge level increases, the seeker firm could rely less on the solvers' knowledge. Moreover, when the seeker firm's knowledge level is relatively high, the inverted-U shape functions across the extent of learning from the crowd (i.e., lr) also point to the synergy between the seeker firm's knowledge and the wisdom of the crowd (see the solid and dotted lines). By incorporating solvers' judgments about the mainstream design choices into its existing knowledge base, the seeker firm is able to obtain solutions of higher quality compared to solely relying on its own knowledge or on the wisdom of the crowd.

The above findings are moderated by problem complexity. First, as the level of problem complexity increases, the benefits of learning gradually diminish. For instance, when the seeker firm's knowledge level is low, learning from the crowd could enhance the absolute quality of the adopted solution by around 0.18 in problems with low complexity (as shown by the dashed line in Figure 3.7a) but only by 0.04 in problems with high complexity (as shown by the dashed line in Figure 3.7c). Second, as problem complexity increases, the optimal extent of learning from the crowd (i.e., lr) declines. The optimal extent of learning from the crowd drops from 0.01 to 0 when the seeker firm has a high level of knowledge (as shown by the dotted lines in Figures 3.7a and c) and from 0.05 to 0.01 when the seeker firm has a medium level of knowledge (as shown by the solid lines in Figures 3.7a and c). The only exception is when the seeker firm has limited knowledge and relies heavily on the wisdom of the crowd (as shown by the dashed lines). These results suggest that incorporating solvers' knowledge becomes less beneficial in complex problems – when the landscape is smooth(er), there is a stronger correlation among nearby solutions. Solvers' choice for a certain design element is less likely to be impacted by other design elements and is also less likely to impact the fitness contribution of those elements. As a result, the independent choices made by solvers are more accurate and tend to build a consensus about what constitutes a "good" solution. Incorporating solvers' knowledge thus tends to enhance the seeker firm's understanding of the solution space. Conversely, when the solution space is (more) rugged, solvers' choice for a certain design element depends on the configuration of other relevant design elements. Solvers are unlikely to reach an accurate consensus as each of them may have specific configurations, making learning from the crowd less beneficial and the synergy between the seeker firm's knowledge and the wisdom of the crowd less attainable.

Taken together, we propose the following:

Proposition 3.1a. Learning from the crowd enhances the absolute quality of the solution adopted from crowdsourcing when the seeker firm has a low level of knowledge and the problem is not complex.

Proposition 3.1b. As the seeker firm's knowledge level increases, synergy with the wisdom of the crowd becomes more attainable.

Proposition 3.1c. As the problem complexity increases, the benefits of learning from the crowd diminish; synergy between the seeker firm's knowledge and the wisdom of the crowd becomes less achievable.

3.4.4 Experiment 4 – A Systems View of Crowdsourcing

The above results indicate that the knowledge flow from the crowd to the seeker firm (i.e., learning from the crowd) could potentially resolve the paradoxes in crowdsourcing. This experiment examines the impact of learning from the crowd, along with other flows, on crowdsourcing effectiveness from a systems view. We fixed the number of solvers at 100 and allowed the seeker firm to provide feedback to ($p_{feedback} = 1$) as well as learn from the crowd. Our analyses are based on testing $19,796 = (14 \times 14 \times 101)$ experimental conditions – 14 levels of firm knowledge, 14 levels of problem complexity, and 101 levels of learning rate $lr = [0, 1]$ with a step size of 0.01. Figure 3.8 shows the absolute quality of the adopted solution from crowdsourcing (as shown by the dashed lines) when the seeker firm always provides feedback to and learns from the crowd across different levels of problem complexity (across columns of figures), different levels of firm knowledge (across rows of figures), and learning rates (x-axis). We also included results from Experiments 2 (as shown by the leftmost points on the dashed lines when $lr = 0$) and 3 (as shown by the

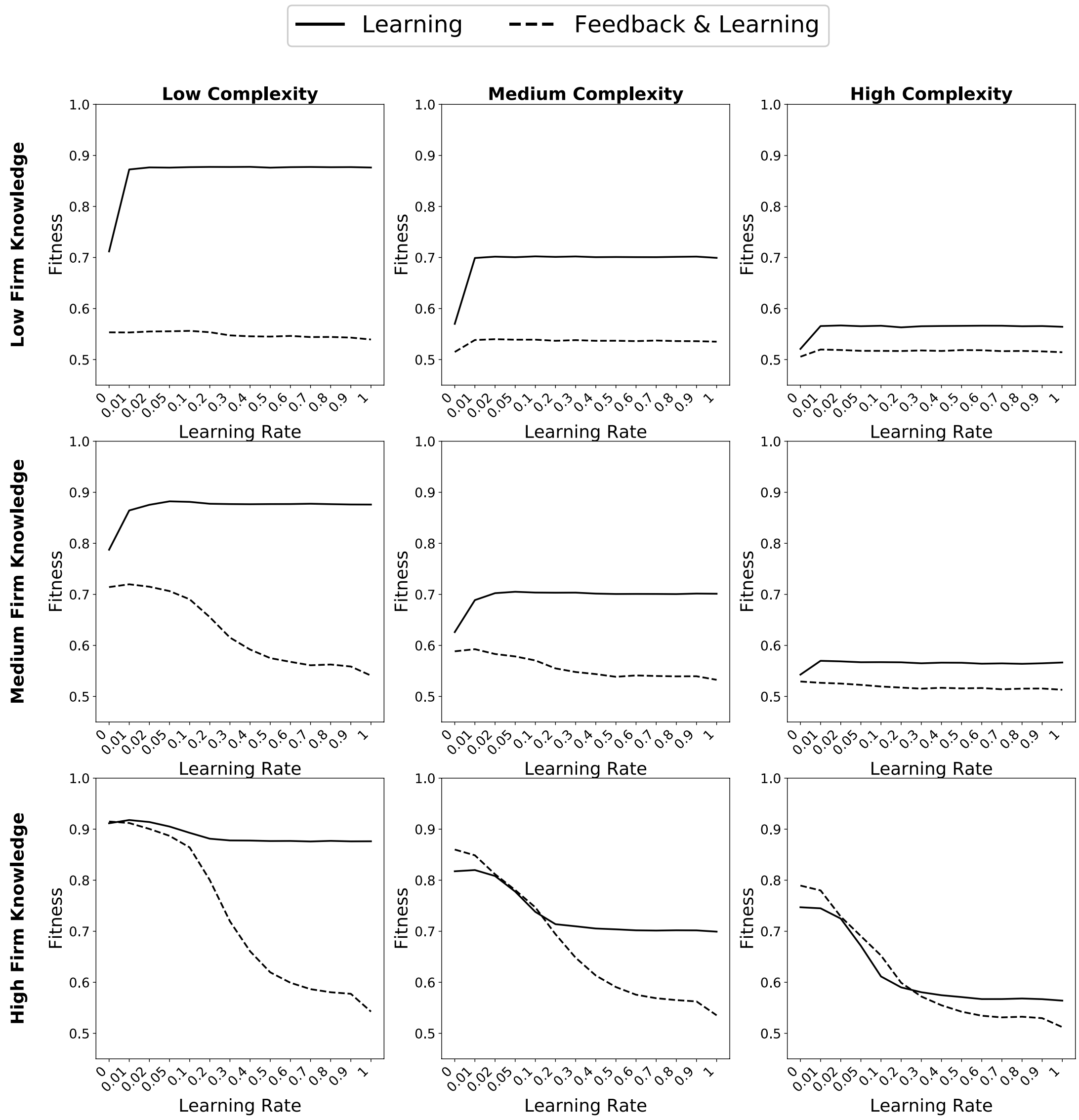

Figure 3.8: Absolute Quality of the Adopted Solution with Learning and Learning and Feedback

solid lines) as baselines.

Counterintuitively, we find that the absolute quality of the adopted solution is worsened when the seeker firm provides feedback and learns from the crowd simultaneously. When the seeker firm has a low or a medium level of knowledge, providing feedback hinders the effectiveness of learning from the crowd (see the dashed and solid lines in the 1st and 2nd rows of Figure 3.8). In other words, simultaneously providing feedback and learning from the crowd is always less effective than solely learning from the crowd for the seeker firm to obtain higher quality solutions. When the seeker firm has a high level of knowledge, learning from the crowd weakens the positive effect of providing feedback on the absolute quality of the adopted solution (i.e., the highest points on the dashed lines are always associated with no learning ($lr = 0$) in the 3rd row of Figure 3.8). Put differently, simultaneously providing feedback and learning from the crowd is always less effective than merely providing feedback. These results suggest the two knowledge flows in the system of crowdsourcing are interdependent. On the one hand, the benefits of learning from the crowd source from the wisdom of the crowd that requires solvers to be as independent as possible. Independent solvers with diverse knowledge can each evaluate particular aspects of a solution. As a result, the seeker firm can regard certain design choices as of high quality if they have been validated and taken by solvers with diverse knowledge. However, the seeker firm's feedback homogenizes solvers' evaluation of solutions and thus harms learning. On the other hand, when the seeker firm has a high level of knowledge, it is likely to increase the absolute quality of the adopted solution by providing feedback on aspects that cannot be captured by any individual solver. Learning from the crowd, however, undermines the value of ongoing feedback by forcing the seeker firm to incorporate solvers' in-progress (and potentially underdeveloped) solutions into its existing knowledge.

Taken together, these results suggest providing feedback to and learning from the crowd may be detrimental to each other if they are implemented simultaneously, necessitating the isolation of feedback and learning in crowdsourcing design. Prior research on crowdsourcing has demonstrated that the seeker firm may provide feedback at different timing (Sanyal and Ye 2023, Jiang et al. 2021), which presents natural opportunities to isolate feedback from learning. As such, we conduct two additional experiments in which the seeker firm first learns from then provides feedback to the seeker firm, or the other way around.[5] Solid lines with vertical line markers and dashed lines with vertical line markers in Figure 3.9 present the absolute quality of the adopted solution when the seeker firm first learns from the crowd then provides feedback, and when the seeker firm first provides feedback then learns from the crowd, respectively. To better highlight the impact of isolating

[5]For simplicity, we divide crowdsourcing into two stages according to the midpoint – i.e., 50th period.

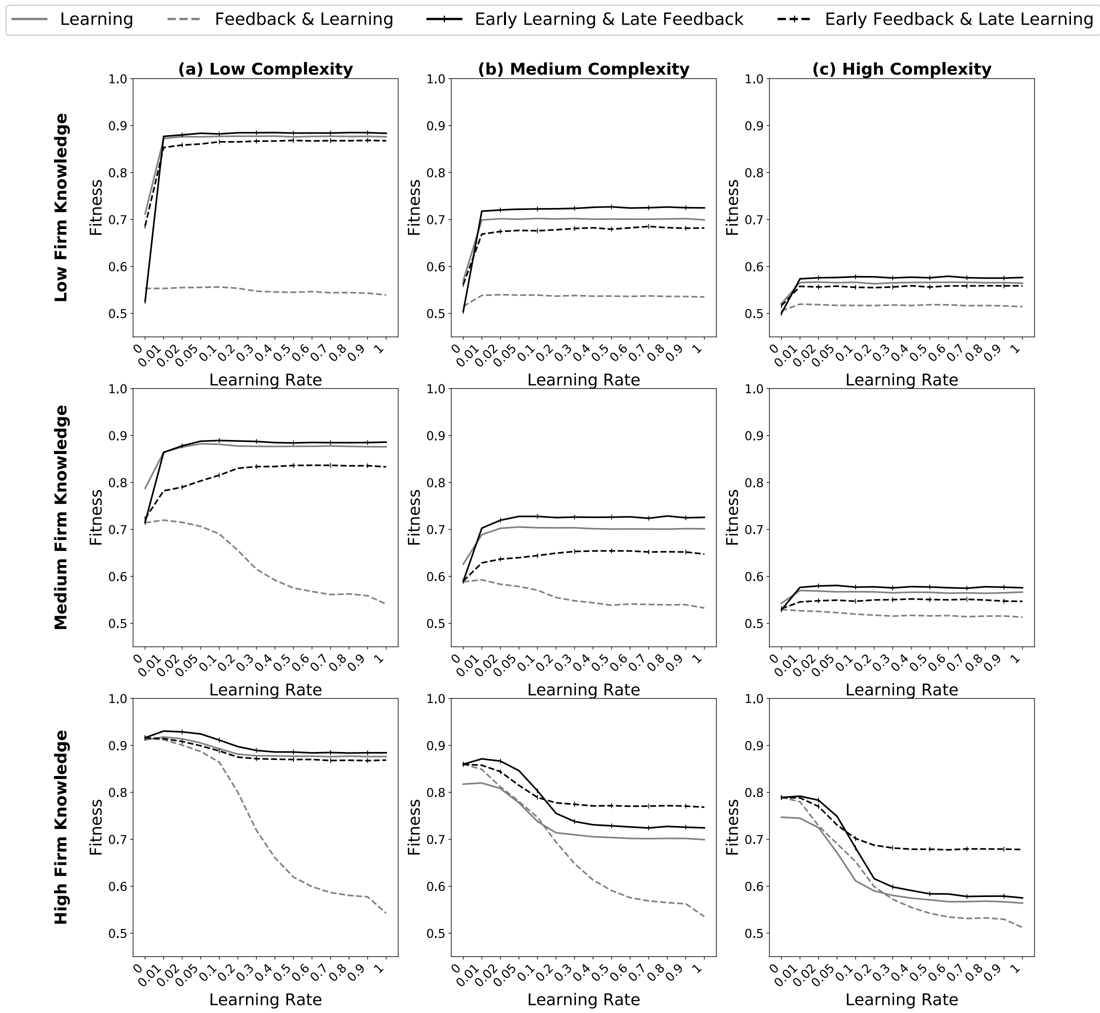

Figure 3.9: Absolute Quality of the Adopted Solution with Learning, Learning and Feedback, Early Learning and Late Feedback, and Early Feedback and Late Learning

feedback and learning, Figure 3.9 also shows previous results (i.e., learning and learning and feedback simultaneously) in solid and dashed lines, respectively.

We find that if the seeker firm first provides feedback and then learns from the crowd, the absolute quality of the adopted solution seems to be nudged up compared to the scenario when the seeker firm provides feedback and learns from the crowd at the same time (i.e., all the dashed lines with vertical line markers are above the dashed lines when $lr > 0$ in Figure 3.9). However, the absolute quality is still less than or equal to when the seeker firm only provides feedback (the leftmost point of the dashed lines in the 3rd row of Figure 3.9) or learns from the crowd (the solid lines in the 1st and 2nd rows of Figure 3.9). The value of learning from the crowd is undermined by the seeker firm's feedback that homogenizes solvers' evaluation. As a result, there is little synergy between the seeker firm's existing knowledge and the wisdom of the crowd, hindering crowdsourcing effectiveness to be further enhanced.

In contrast, we find that the seeker firm may adopt a solution of similar and even higher quality if it first learns from and then provides feedback to the crowd. The solid lines with vertical line markers tend to have the highest performance for problems with medium complexity when the seeker firm has a low or a medium level of knowledge (as in the 2nd column, 1st and 2nd rows of Figure 3.9), and for problems with low complexity when the seeker firm has a high level of knowledge (as in the 1st column, 3rd row of Figure 3.9). These results suggest that the seeker firm is only able to transmit the knowledge learned from the crowd to the crowd under certain conditions. When the seeker firm has limited knowledge about domains relevant to the crowdsourced problem, it mostly relies on solvers' judgments about what design choices constitute a good solution. On the one hand, the quality of the wisdom of the crowd decreases as problem complexity increases (as in Propositions 3.1). On the other hand, to what extent the seeker firm can utilize the wisdom of the crowd to guide the solvers' search also depends on problem complexity. When problem complexity is low, there are few interdependencies among different design elements and as a result, how a particular design element contributes to the quality of the overall solution is less dependent on other design elements. Knowing if a particular design element in the solution fits other solvers' judgments is unlikely to induce solvers to re-configure the remaining elements. Providing feedback in the late stage thus may fail to fully utilize the wisdom of the crowd to guide solvers' search. As problem complexity increases, it is more likely for solvers to search the whole knowledge space again if one particular design element has been adjusted. Therefore, taken these two forces together, when the seeker firm has limited knowledge regarding the domains relevant to the crowdsourced problem, it can further enhance the absolute quality of the adopted solution by first learning from and then providing feedback to the solvers, especially for problems with medium complexity levels. As the seeker firm's knowledge level increases, it relies less on solvers' judgments and is more likely to take advantage of the synergy between its existing knowledge and the wisdom of the crowd (as shown by the inverted-U shape functions across the extent of learning in Figure 3.7). As such, the seeker firm's feedback captures not only whether a design choice fits the wisdom of the crowd but also how it actually performs. However, the benefits of learning from the crowd gradually diminish as problem complexity increases, making providing feedback after learning less valuable for complex problems. The difference between the solid line with vertical markers and the other lines in the 3rd row of Figure 3.9 is most pronounced when problem complexity is low.

In summary, it is possible for the solution seeker to further enhance crowdsourcing from a systems view by first learning from and then providing feedback to the crowd. We summarize our findings into the following propositions:

Proposition 3.2a. When the seeker firm has a low level of knowledge and the problem is of medium complexity, learning first from and then providing feedback to the crowd enhances the absolute quality of the solution adopted from crowdsourcing.

Proposition 3.2b. When the seeker firm has a high level of knowledge and the problem is of low complexity, learning first from and then providing feedback to the crowd enhances the absolute quality of the solution adopted from crowdsourcing.

3.4.5 Sensitivity Analysis

We conducted a number of sensitivity analyses. First, as discussed, we varied the number of solvers $(100, 200, 400)$ for Experiments 2, 3, and 4 and found no substantive differences in the results. Second, we varied the size of the solvers' knowledge set N_c. While results from Experiments 1, 2, and 3 remain qualitatively unchanged, increasing (reducing) the solvers' knowledge level makes learning from the crowd relatively more (less) beneficial than providing feedback to solvers. As a result, our results in Experiment 4 would shift towards more learning (when solvers possess relatively more knowledge) and more feedback (when solvers possess relatively less knowledge), respectively. Table 3.2 summarizes all the simulation results accordance with the framework we proposed in Figure 3.1.

Table 3.2: Summary of Experiments and Results

Experiment	Systems View (as per Figure 3.1)	Simulation Results
1	The solution flow from the crowd of solvers to the seeker firm	**Paradox 3.1.** When the seeker firm has insufficient knowledge about the solution space, crowdsourcing is most needed but also least utilized.
2	The knowledge flow from the seeker firm to the crowd of solvers	**Paradox 3.2.** When the seeker firm has insufficient knowledge about the solution space (i.e., when crowdsourcing is actually most needed), the seeker firm's feedback to solvers worsens crowdsourcing effectiveness even though firm may be more satisfied with the solutions developed via crowdsourcing.
3	The knowledge flow from the crowd of solvers to the seeker firm	**Proposition 3.1a.** Learning from the crowd enhances the absolute quality of the solution adopted from crowdsourcing when the seeker firm has a low level of knowledge and the problem is not complex.
		Proposition 3.1b. As the seeker firm's knowledge level increases, synergy with the wisdom of the crowd becomes more attainable.
		Proposition 3.1c. As the problem complexity increases, the benefits of learning from the crowd diminish; synergy between the seeker firm's knowledge and the wisdom of the crowd becomes less achievable.
4	The interactions of the solution and knowledge flows	**Proposition 3.2a.** When the seeker firm has a low level of knowledge and the problem is of medium complexity, learning first from and then providing feedback to the crowd enhances the absolute quality of the solution adopted from crowdsourcing.
		Proposition 3.2b. When the seeker firm has a high level of knowledge and the problem is of low complexity, learning first from and then providing feedback to the crowd enhances the absolute quality of the solution adopted from crowdsourcing.

3.5 Discussion and Conclusion

This paper offers a novel perspective on how to enhance crowdsourcing effectiveness from a systems view. Drawing on the tension between the parallel path effect and knowledge absorptive capacity theory, we posit that the seeker firm is unlikely to possess the full range of knowledge to evaluate solutions when solvers with diverse (albeit limited) knowledge are engaged in a crowdsourced problem that requires knowledge from multiple domains. As such, crowdsourcing is a system constrained by the seeker firm's knowledge – solvers' parallel search always needs to be evaluated and selected by the firm before being applied to the crowdsourced problem. We extend the *NK* fitness landscape model to examine how the system of crowdsourcing, within which the seeker firm with limited knowledge interacts with solvers via solution and knowledge flows, performs in discovering high quality solutions. Our results uncover two paradoxes related to the solution flow, in which solutions developed by the solvers are submitted to the seeker firm and the knowledge flow, in which the seeker firm provides feedback to the solvers. First, the crowdsourcing system is most likely to overlook high quality solutions developed by the solvers when they are most needed. Second, when the seeker firm provides feedback, although the seeker firm perceives the adopted solution as of higher quality, the absolute performance may be poor. To resolve these paradoxes, we examine the knowledge flow from the crowd of solvers to the seeker firm by enabling the seeker firm to learn from the crowd about what design choices constitute a good solution. We find that learning from the crowd helps the seeker firm obtain higher quality solutions, albeit conditionally depending on the seeker firm's knowledge level and problem complexity. We find that the two knowledge flows can be detrimental to each other such that the system performance can only be enhanced by isolating providing feedback to and learning from the crowd. These findings point to a wide range of theoretical and practical implications.

This study is, to the best of our knowledge, the first holistic examination of crowdsourcing effectiveness from a systems view. It introduces a new and important conceptual framework for understanding crowdsourcing effectiveness by highlighting the existence of the cave of crowdsourcing. In the famous allegory of "The Cave," the Greek philosopher Plato describes a group of people who have been imprisoned all their lives, chained to gaze at the wall in front of them. Behind the prisoners, objects are passed through in front of a fire, projecting shadows on the wall. The shadows are the reality for the prisoners because they have never seen anything else but may not be accurate representations of the real world. This story is often told to describe the chase for truth of the world. When prisoners are born with intelligence incapability in the cave, instead of naming the "shadows" as reality, the only way to approach reality is by escaping from the cave. Along with this anecdote, our study suggests that to fully utilize the innovativeness of crowdsourcing, it is crucial to evade the cave of crowdsourcing. Due to the nature of crowdsourcing to discover solutions for the seeker firm, satisfying the seeker firm has been the reality for solvers in many crowdsourcing practices. However, as the crowdsourced problem becomes increasingly complex and requires multidisciplinary knowledge (Gurca et al. 2023, Majchrzak et al. 2021, Majchrzak and Malhotra 2020), it becomes almost impossible for the seeker firm to possess the full range of relevant knowledge. In turn, limited knowledge hinders the seeker firm's understanding and consequently its ability to evaluate solutions submitted by solvers with diverse knowledge and may not be able to distinguish between a satisfactory solution from an effective one, making satisfying the seeker firm likely to become the "shadow" of reality. The discrepancy between "shadow" (i.e., satisfaction) and "reality" (i.e., effectiveness) requires a clear distinction between satisfying the seeker firm within the system (i.e., naming the "shadows") and enhancing the overall system performance in discovering effective solutions (i.e., chasing reality). For instance, open crowdsourcing platforms have received considerable attention in academia due to the accessibility of data. However, it has primarily (and perhaps implicitly) focused on the seeker firm's satisfaction at the end of crowdsourcing (e.g., Mo et al. 2021, Jin et al. 2021) rather than the absolute quality of the adopted solution, which would better reflect crowdsourcing effectiveness. Without a systems view outside the cave of crowdsourcing, crowdsourcing platforms may constantly host (seemingly) successful crowdsourcing campaigns but generate few valuable solutions to the seeker firm. The conventional wisdom of crowdsourcing and the taken-for-granted approach of studying crowdsourcing with data drawn from crowdsourcing platforms thus need to be carefully reconsidered in terms of whether they indeed help the seeker firm in developing truly innovative solutions. While this study takes a first step by generating theoretical propositions via a novel simulation-based approach, we urge future studies to continue to bridge this discrepancy.

By highlighting the existence of the cave in crowdsourcing, this study extends our understanding of crowdsourcing in several ways. First, this study adds to the growing body of research that seeks to enhance crowdsourcing effectiveness via strengthening the parallel path effect (e.g., Jian et al. 2019, Boudreau et al. 2011). Prior studies have mainly examined various ways to encourage solver participation to increase the likelihood of having one solver with the required knowledge participate in crowdsourcing. Although this perspective is reasonable, it may be untenable when the crowdsourced problems require knowledge beyond the seeker firm's existing knowledge base, the raison d'être of crowdsourcing. Several recent studies have shown that the seeker firm may fail to utilize the parallel path effect and only adopt solution it is familiar with (e.g., Piezunka and Dahlander 2015). Our findings provide evidence that the seeker firm's knowledge impedes the parallel path effect. The seeker firm's existing knowledge base determines its representation of the problem (and solutions) and eventually how the solvers' parallel search is evaluated according to the (potentially) inaccurate mental representation. When the "best" solution defined according to the seeker firm's mental representation happens to be discovered by the solvers, further involving solvers has a marginal impact on overall crowdsourcing effectiveness as the "shadows" have already been fully seized by the solvers. This study thus contributes to the crowdsourcing literature by highlighting the seeker firm's knowledge as a potential boundary condition of the effectiveness of the parallel path effect.

Relatedly, this study also contributes to research on the impact of the seeker firm's feedback in shaping solvers' parallel search (e.g., Jiang et al. 2021, Sanyal and Ye 2023, Mihm and Schlapp 2019) from a normative perspective by showing the potential misalignment between the perceived and absolute quality of the adopted solution when the seeker firm guides solvers with feedback. Prior research has mainly investigated the performance implication of feedback through the lens of the parallel path effect – i.e., whether and to what extent providing feedback would affect solvers' participation level and search behavior. A somewhat overlooked aspect is that feedback, albeit in various forms, is inherently grounded in the seeker firm's existing knowledge about the solution space of the crowdsourced problem. Our simulation results suggest that the seeker firm's feedback could be the "chains" that force solvers to face the "shadows" on the wall, making the seeker firm more likely to obtain solutions of higher perceived quality within the system (i.e., satisfactory solutions). However, to what extent feedback increases the likelihood of finding solutions of higher absolute quality (i.e., effective solutions) depends on whether the "shadows" are accurate representations of reality. When the seeker firm has limited knowledge regarding the domains relevant to the crowdsourced problem, the seeker firm's mental representation may deviate from the landscape and as a result, providing feedback makes the seeker firm satisfied at the cost of worsening overall crowdsourcing effectiveness. This study thus points to an alternative mechanism through which feedback may affect crowdsourcing effectiveness. Future research could investigate the interplay between how feedback affects the parallel path effect and how feedback steers solvers to chase the "shadows."

Finally, this study also contributes to the crowdsourcing literature by showing the possibility of using the wisdom of the crowd to evade the cave of crowdsourcing. The prior literature on the power of crowds could be generally divided into two streams – one stream has mainly focused on the parallel path effect to conduct independent search in mass via crowdsourcing (Boudreau et al. 2011, Jian et al. 2019); the other has focused on how to aggregate crowds' beliefs to improve decision-making accuracy (Surowieck 2004, Da and Huang 2020). The parallel path effect and wisdom of the crowd, albeit focusing on different tasks, both require the participation of solvers (i.e., the crowd) with diverse knowledge. As such, crowdsourcing can naturally benefit from utilizing the wisdom of the crowd. This study integrates these two streams of literature by examining how the seeker firm could learn from solvers in crowdsourcing. Whilst the seeker firm would only adopt one solution with the highest perceived quality, it could learn about the design choices that have been frequently made in solvers' solutions from the union of solvers. Our findings suggest that when the seeker firm has limited knowledge regarding the domains relevant to the crowdsourced problem, learning from the crowd can generally enhance the absolute quality of the solution adopted by the seeker firm. By incorporating the knowledge about whether a mainstream design choice has been incorporated in a solution into the seeker firm's existing knowledge, it is possible to make the "shadows" closer to reality. However, the wisdom of the crowd becomes less informative in problems with high complexity due to interdependencies among different design elements. As a result, learning may backfire as problem complexity increases. Our results also suggest that to fully utilize the wisdom of the crowd, the seeker firm should first "unchain" solvers by ceasing feedback to solvers. Taken together, this study thus points to a contingency view of learning from the crowd to evade the cave of crowdsourcing.

We suggest several directions for future research into this underexplored area. First, although this study, to the best of our knowledge, is the first to examine the wisdom of the crowd in the crowdsourcing context, the seeker firm in practice may have already begun to learn from the crowd implicitly. Some crowdsourcing platforms, for instance, support in-progress ratings given by the seeker firm (Jiang et al. 2021) and two-stage competitions (Hou and Zhang 2021), which provide natural opportunities for the seeker firm to update its understanding of the solution space according to the exposed in-progress solutions. Some crowdsourcing platforms have also implemented ongoing communication channels between the seeker firm and solvers during crowdsourcing. Future research could validate the possibility of learning from the crowd via qualitative interviews with the seeker firm. Second, this study uses computational experiments to generate theoretical propositions of learning from the crowd. Future research could test these propositions via experimental or field studies. These propositions may not be tested in any single study due to the difficulty of simultaneously manipulating or measuring all the variables and solution (knowledge) flows in our simulation model. This study highlights possible boundary conditions about how individual studies' different (sometimes even conflicting) results should be compared and accumulated. Last, prior research has suggested that the seeker firm's attention may be crowded out in crowdsourcing (Piezunka and Dahlander 2015). Learning from the crowd, however, requires the seeker firm to focus on in-progress solutions during crowdsourcing. Future research can thus examine whether and to what extent the seeker firm is able to allocate attention to in-progress solutions to utilize the wisdom of the crowd.

This study points to several practical implications as well. First, this study suggests that the seeker firm's efforts to leverage external knowledge via crowdsourcing must take into account the impact of its existing knowledge base. Although crowdsourcing enables the seeker firm to conduct parallel search of the solution landscape cost-effectively compared to serial trial and error via internal R&D, it still involves a considerable expenditure of time and financial resources to attract solvers. Even if a great number of solutions can be generated by artificial intelligence technologies in a cost-efficient manner (Boussioux et al. 2024), the extent to which these solutions can be effectively leveraged still depends on the firm's existing knowledge base. Considering the seeker firm's knowledge as the bottleneck of the parallel path effect, the seeker firm should carefully evaluate (and enhance) its existing knowledge base to align efforts in attracting solvers. Second, this study suggests the seeker firm should reassess the performance implication of how it steers solvers' parallel search. Some rules of thumb in extant practices, for instance, designing reward structures to guarantee solver participation and offering in-progress feedback, have been shown to have marginal and even adverse impacts on crowdsourcing effectiveness beyond merely satisfying the seeker firm when the seeker firm has limited knowledge regarding the domains relevant to the crowdsourced problem. The seeker firm can thus first evaluate its knowledge base before directing the crowd of solvers. When the crowd of solvers can find its way, less sometimes is more. Third, this study also points to a new direction for the seeker firm to enhance crowdsourcing effectiveness via learning from the crowd. Learning from the crowd requires the seeker firm to engage in the solution development process and update its evaluation criteria correspondingly. Our results provide guidance on to what extent the seeker firm should learn from the crowd and whether the seeker firm should learn from the crowd together with other crowdsourcing policies like providing feedback.

Despite the strengths of this study, some limitations should be noted. Our extended *NK* model, similar to other simulation models, is a simplified abstraction of reality. We only capture the basic properties of crowdsourcing and the interactions between the seeker firm and solvers. Crowdsourcing in the real world is more complicated. For instance, feedback from the seeker firm may take place in various forms (e.g., private vs. public, quantitative vs. qualitative) (Mihm and Schlapp 2019, Sanyal and Ye 2023). Likewise, solution solvers and the seeker may exhibit differences beyond different knowledge levels. In a similar vein, learning from the crowd in reality is also constrained by solver motivation. Solvers may participate in crowdsourcing to merely take a chance on the reward and thus submit less valuable solutions to the seeker firm. Future research could continue to add more details to the developed model according to empirical observations and examine their impact on the effectiveness of the crowdsourcing system. It is worth noting that most observational data is about how the seeker firm perceives a solution developed via crowdsourcing and is thus unlikely to be the ground truth to study how the system consisting of the seeker firm and solvers performs. We encourage more longitudinal studies in the crowdsourcing field to step out of the cave and uncover the long-term performance implications beyond short-term satisfaction and acceptance of solutions by the seeker firm. The simulation methodology equips us with a theoretical framework to abstract reality and to foresee possible outcomes and boundary conditions on this journey.

Chapter 4 Conclusions

4.1 Summary

Crowdsourcing is fast emerging as a mainstream innovation tool for firms. However, despite its wide adoption, crowdsourcing has been challenged with respect to its ability to solve complex problems. This thesis aims to further crowdsourcing for innovative solutions to complex problems. Given the limits of Taylorism assumptions in the extant crowdsourcing research, this thesis first proposes a post-Taylorism theoretical perspective of crowdsourcing that improves crowdsourcing effectiveness with a focus on the (re)combination of expertise across varying levels of problem complexity and then conduct two related studies that develop different aspects of the proposed theoretical perspective through a simulation-based theory building approach (Davis et al. 2007).

The first essay investigates (re)combination through teaming that takes place in a self-selection manner. It focuses on how such self-selected teaming would influence overall crowdsourcing effectiveness. I first find suggestive empirical evidence from Kaggle that solvers are likely to rely on quality signals from publicly available information on crowdsourcing platforms in the teaming process. Based on the empirical observations, I extend the *NK* fitness landscape model and conduct counterfactual experiments to unlock the impact of self-selected teaming on global crowdsourcing effectiveness. I find that the impact of teaming on crowdsourcing effectiveness originates from the immediate returns from identifying other solvers (or their solutions) to integrate with at the time of teaming and potential (but time-consuming) returns from joint teamwork after teaming. The *identification*, *integration*, and *teamwork effects* are moderated by problem complexity and the timing of teaming and may make teaming detrimental to overall crowdsourcing effectiveness under certain conditions.

The second essay examines knowledge (re)combination between the firm (i.e., the solution seeker) and solvers. I first conceptualize crowdsourcing as a system constrained by the firm's existing knowledge base. I further extend the *NK* fitness landscape model and conduct simulation experiments to illustrate how the extant crowdsourcing literature and practice may only contribute to suboptimal problem-solving. I find that when massive parallel search is most needed to develop solutions beyond the firm's knowledge base, the firm is also most likely to overlook high quality solutions. Second, feedback from the firm may make the firm more satisfied at the cost of worsening crowdsourcing effectiveness. To mitigate the identified issues, I examine whether and how the seeker firm could enhance the crowdsourcing system's ability to obtain high quality solutions by proactively learning from the crowd about what constitute a good solution (i.e., the wisdom of the crowd). I find a potential synergy between the seeker firm's existing knowledge and the wisdom of the crowd, albeit depending on problem complexity and the interplay between learning and other interactions between the solution seeker and solvers.

In summary, my dissertation examines the impact of knowledge (re)combination on crowdsourcing effectiveness for complex problems. Overall, our findings suggest that knowledge (re)combination can enhance crowdsourcing effectiveness and help firms discover high quality solutions. However, if not properly managed, it may also backfire. Notably, the impact of knowledge (re)combination may be misleading if the outcome variable is not clearly defined. Our findings suggest that the quality of individual solutions and the seemingly high quality solutions selected by the solution seeker does not always align with overall crowdsourcing effectiveness. Finally, our results highlight problem complexity as a key contingency factor for the above conclusions.

4.2 Contributions

This thesis makes significant contributions to theory, methodology, and practice. Chapter 2 and Chapter 3 have already discussed the contributions within their specific contexts and foci. The following sections will outline and discuss each of the contributions from a broader perspective.

4.2.1 Contributions to Theory

This thesis contributes to the literature on crowdsourcing by proposing a new theoretical framework that relaxes the assumptions grounded implicitly in Taylorism management in existing crowdsourcing research. The proposed framework highlights three essential elements that complement the Taylorism view of crowdsourcing. First, in addition to recruiting as many solvers as possible and matching the crowdsourced problem with suitable expertise from specialized solvers, the crowdsourced problem can be solved in a collaborative manner where the expertise possessed by the seeker and solvers (re)combines each other and ignite exploration into solutions that can hardly be discovered by parallel search alone. However, improper (re)combination cannot compensate for the loss in parallel search and may hamper crowdsourcing effectiveness. The crowdsourcing literature should therefore incorporate knowledge (re)combination as a key theoretical construct. I suggest, in particular, several directions for future research to further our understanding of knowledge (re)combination in crowdsourcing. First, from a descriptive perspective, as suggested by the empirical observation in Essay I, knowledge (re)combination in crowdsourcing, although not formally managed, does not occur in a fully random manner. As such, future research should collect data and investigate how knowledge (re)combination unfolds at a finer level of granularity. While my dissertation focuses on direct knowledge (re)combination among agents, knowledge (re)combination can be even more complex and may be better conceptualized as a dynamic network in which agents interact with one another over time. For instance, a solver may build its solution on top of others' solutions, which themselves were built upon prior solutions submitted by earlier participants. The solutions submitted by earlier participants, in turn, may be influenced by how the seeker firm steer the massive parallel search. The recent emergence of AI tools further exacerbates the tendency for knowledge (re)combination to concentrate along specific edges in the network (e.g., human solvers using a public AI agent). I therefore encourage future research to examine knowledge (re)combination in crowdsourcing by incorporating both the contemporaneous and temporal interdependencies among agents. Theoretical lenses such as network dynamics and complex adaptive systems might be leveraged. Such descriptive studies will further refine the micro-level theoretical assumptions about whether and how knowledge (re)combination takes place in crowdsourcing. While my dissertation offers a way to understand the impact of micro-level (re)combination on macro-level outcomes (i.e., global crowdsourcing effectiveness) from a normative perspective, the extent to which knowledge (re)combination is beneficial depends critically on its micro-level foundations. Relatedly, from a prescriptive perspective, I hope my dissertation can also inspire future research to design crowdsourcing systems that better incentivize knowledge (re)combination in a desired manner. A two-stage approach should be employed to examine the effectiveness of different design strategies. Namely, how a design influences micro-level behaviors and in turn, how global crowdsourcing effectiveness is achieved.

The proposed theoretical framework also takes a contingency view regarding complexity. Early theoretical perspectives of crowdsourcing have been proposed and validated on simple tasks. This thesis further refines them by including complexity as a boundary condition. The results of this thesis highlight that the same attempt to increase crowdsourcing effectiveness may have divergent impacts in problems with varying complexity levels. Therefore, the development of crowdsourcing theory should be prudent specifically with regard to levels of problem complexity levels. Here I suggest two directions for future research related to the sources and channels through which complexity influences crowdsourcing. First, as explored in Essays I and II, the complexity associated with problem solving can stem from two sources – the problem itself and the integration of the solution into the firm. Seemingly simple crowdsourced problems can exhibit hidden complexity when substantial interdependencies exist between the seeker firm and crowdsourcing solvers. A more nuanced view of the pattern and distribution of interdependencies among design elements between the seeker firm and crowdsourcing solver can therefore help advance our understanding of the moderating role of complexity. Second, complexity can act as a contingency factor through different channels. It may influence crowdsourcing effectiveness by directly affecting the likelihood of finding high quality solution, or by affecting how knowledge (re)combination occurs and how such knowledge (re)combination influences crowdsourcing effectiveness. Future research therefore should differentiate among these channels, as the way complexity exerts its influence may require different remedies to fully leverage crowdsourcing for complex problem solving.

Finally, the proposed framework focuses on crowdsourcing effectiveness rather than efficiency, which was the primary focus of most prior crowdsourcing studies. While an efficiency-oriented theoretical framework helps crowdsourcing to produce effective solutions by encouraging solver participation and submission, this thesis further complements this view by arguing and showing that the crowdsourcing effectiveness can be further enhanced by active (re)combination. Crowdsourcing research can better underpin crowd-based innovation activities by integrating these two views. The move from efficiency to crowdsourcing effectiveness requires an ontological shift – from studying crowdsourcing solvers (and their solutions) to examining the performance of crowdsourcing systems. This shift therefore points to a more system-oriented epistemology in which interactions among agents are central to understanding crowdsourcing effectiveness. It opens up a range of research questions, some of which have been discussed above using the example of knowledge (re)combination. For example, how do micro-level interactions aggregate to influence crowdsourcing effectiveness? What types of interaction patterns best support the discovery of high quality solutions? And how can platforms be better designed to improve the performance of crowdsourcing systems? Answering these questions, of course, requires innovative research designs to capture and evaluate crowdsourcing outcomes. In addition to the simulation-based approach employed in this dissertation, large-scale field experiments and longitudinal studies examining whether the firm benefit from crowdsourcing could provide richer insights into the the effectiveness of crowdsourcing beyond the quality of individual solutions or short-term satisfaction.

4.2.2 Contributions to Methodology

This thesis contributes to the research methodology by introducing and extending the basic *NK* fitness landscapes model to crowdsourcing research. The proposed simulation model captures massive parallel search for problems with different complexity levels and interdependency patterns. More importantly, it enables counterfactual experimentation on crowdsourcing effectiveness. Future research can easily extend the proposed model by changing some components to study other phenomena relevant to crowdsourcing.

Second, this thesis also brings novel insights into the extant methods in crowdsourcing research. The mainstream empirical methods should adapt to investigate crowdsourcing effectiveness rather than efficiency. For those studies that approximate crowdsourcing effectiveness with solvers' performance or the quality of the winning solutions, this thesis offers a more comprehensive picture of the underlying data generating process and how the observed data may be biased representations of crowdsourcing effectiveness. For analytical methods used in the extant crowdsourcing literature, this thesis provides an alternative perspective to formalize the seeker and solvers according to the assumptions of bounded rationality and anticipate possible outcomes when a crowd of actors interact with each other. Given constraints of time, imperfect information, and insufficient knowledge at both the seeker and solver ends, analytical methods should be more prudent in making assumptions such as simplifying crowdsourcing as a one-shot game between several decision makers.

4.2.3 Contributions to Practice

Practical contributions from this research are many-fold. First, this research offers guidance to firms on whether and how they should adopt crowdsourcing approach to pursue innovative solutions. In deciding whether to adopt a crowdsourcing approach, Essay I highlights the importance of understanding the solver population in terms of whether solvers are inclined toward collaboration or tend to avoid it. Likewise, Essay II points to the firm's existing knowledge base as a key factor that influences the effectiveness of crowdsourcing. The results from these two essays also suggest that crowdsourcing effectiveness largely depend on how proactively firms manage the crowd. For instance, decisions about whether to allow team formation and what information to provide to nudge the team formation process, as well as whether and to what extent in-progress feedback and evaluation updates should be offered. This research also suggests that there is no one-size-fits-all approach to running crowdsourcing. It identifies several important boundary conditions, such as the number of solvers recruited and the complexity of the problem, for firms to consider when strategically opting in and managing crowdsourcing.

This thesis also has important implications for crowdsourcing platforms. It contributes to the design of crowdsourcing platforms. The results of this thesis should help platforms to design and examine different features from both the parallel path and (re)combination perspective. Meanwhile, crowdsourcing platforms should also commit to improving the innovativeness of crowdsourcing rather than merely improving the average quality of individual solutions or satisfying the seeker, which may hamper the value of crowdsourcing and crowdsourcing platforms in the long run.

4.3 Limitations

There are several limitations for this research that need to be carefully considered when interpreting the findings. First and foremost, this thesis draws on a simulation-based approach. The computation models developed are still simplified abstractions of the real world based on certain assumptions about how crowdsourcing operates. The models do not reflect all possible factors that might affect crowdsourcing effectiveness. For example, the models do not account for solver motivation, which is an important factor influencing the level of effort solvers contribute in the crowdsourcing process. Interpretation of the simulation results should be made with caution, as the model's abstraction defines the contextual boundaries of the findings. Future research could relax the assumptions made in this thesis and examine the boundary conditions of my findings.

Second, this thesis considers only two representative cases of knowledge (re)combination – i.e., team formation between formerly independent solvers, and knowledge (re)combination between the seeker and solvers via feedback and learning. There are various ways in which knowledge possessed by different actors is combined in crowdsourcing. For example, the crowd of solvers can co-develop solutions in a community-based manner. Likewise, the solution seeker may influence solvers' parallel search by providing exemplars. While the theoretical insights can, to some extent, be applied to these variations, future research should incorporate additional contextual factors that influence the effectiveness of knowledge (re)combination among solvers, and between the solution seeker and solvers.

The conclusions derived from the simulation experiments are also not empirically tested. The key variable in this thesis, namely crowdsourcing effectiveness, poses challenges for empirical validation. Each crowdsourcing competition typically yields only one observation (i.e., the quality of the best solutions from all discovered ones), and there is no counterfactual for how the same competition would perform under alternative teaming policies and firm interventions. I encourage future research to bridge this gap through creative empirical strategies.

Bibliography

Afuah A, Tucci CL (2012) Crowdsourcing as a solution to distant search. *Academy of Management Review* 37(3):355–375.

Alavi M, Tiwana A (2002) Knowledge integration in virtual teams: The potential role of KMS. *Journal of the American Society for Information Science and Technology* 53(12):1029–1037.

Amatrian X, Basilico J (2012) Netflix recommendations: Beyond the 5 stars. Retrieved (November 7, 2025), https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429.

Baumann O, Schmidt J, Stieglitz N (2018) Effective search in rugged performance landscapes: A review and outlook. *Journal of Management* 45(1):285–318.

Benbya H, Nan N, Tanriverdi H, Yoo Y (2020) Complexity and information system research in the emerging digital world. *MIS Quarterly* 44(1):1–17.

Bernstein E, Shore J, Lazer D (2018) How intermittent breaks in interaction improve collective intelligence. *Proceedings of the National Academy of Sciences* 115(35):8734–8739.

Boudreau KJ, Lacetera N, Lakhani KR (2011) Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science* 57(5):843–863.

Boudreau KJ, Lakhani KR (2013) Using the crowd as an innovation partner. *Harvard Business Review* 91(4):60–69.

Boudreau KJ, Lakhani KR, Menietti M (2016) Performance responses to competition across skill levels in rank-order tournaments: Field evidence and implications for tournament design. *The RAND Journal of Economics* 47(1):140–165.

Boussioux L, Lane JN, Zhang M, Jacimovic V, Lakhani KR (2024) The crowdless future? Generative AI and creative problem-solving. *Organization Science* 35(5):1589–1607.

Butler BS, Bateman PJ, Gray PH, Ilana DE (2014) An attraction-selection-attrition theory of online community size and resilience. *MIS Quarterly* 38(3):699–728.

Cao F, Wang W, Lim E, Liu X, Tan CW (2022) Do social dominance-based faultlines help or hurt team performance in crowdsourcing tournaments? *Journal of Management Information Systems* 39(1):247–275.

Chan TH, Mihm J, Sosa M (2021) Revisiting the role of collaboration in creating breakthrough inventions. *Manufacturing & Service Operations Management* 23(5):1005–1024.

Chesbrough HW (2003) The era of open innovation. *MIT Sloan Management Review* 44(3):35–41.

Christensen M, Knudsen T (2010) Design of decision-making organizations. *Management Science* 56(1):71–89.

Cohen WM, Levinthal DA (1990) Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35(1):128–152.

Cramton CD (2001) The mutual knowledge problem and its consequences for dispersed collaboration. *Organization Science* 12(3):346–371.

Csaszar FA, Eggers JP (2013) Organizational decision making: An information aggregation view. *Management Science* 59(10):2257–2277.

Cyert RM, March JG (1963) *A Behavioral Theory of the Firm* (Englewood Cliffs, NJ: Prentice-Hall).

Da Z, Huang X (2020) Harnessing the wisdom of crowds. *Management Science* 66(5):1847–1867.

Dahlander L, Piezunka H (2014) Open to suggestions: How organizations elicit suggestions through proactive and reactive attention. *Research Policy* 43(5):812–827.

Dahlander L, Piezunka H (2020) Why crowdsourcing fails. *Journal of Organization Design* 9(24).

Davis JP, Eisenhardt KM, Bingham CB (2007) Developing theory through simulation methods. *Academy of Management Review* 32(2):480–499.

Devine DJ, Philips JL (2001) Do smarter teams do better? A meta-analysis of cognitive ability and team performance. *Small Group Research* 32(5):507–532.

Dissanayake I, Nerur S, Zhang J (2019) Team formation and performance in online crowdsourcing competitions: The role of homophily and diversity in solver characteristics. *Proceedings of the 40th International Conference on Information Systems (ICIS 2019)* (Munich, Germany).

Dissanayake I, Zhang J, Gu B (2015) Task division for team success in crowdsourcing contests: Resource allocation and alignment effects. *Journal of Management Information Systems* 32(2):8–39.

Edmundson AC (2012) *Teaming: How Organizations Learn, Innovate, and Compete in the Knowledge Economy* (San Francisco, CA: Jossey-Bass).

Espinosa JA, Slaughter SA, Kraut RE, Herbsleb JD (2007) Team knowledge and coordination in geographically distributed software development. *Journal of Management Information Systems* 24(1):135–169.

Faraj S, Sproull L (2000) Coordinating expertise in software development teams. *Management Science* 46(2):1554–1568.

Fleming L, Sorenson O (2003) Navigating the technology landscape of innovation. *MIT Sloan Management Review* 44(2):15–23.

Gavetti G, Helfat CE, Marengo L (2017) Searching, shaping, and the quest for superior performance. *Strategy Science* 2(3):194–209.

Gavetti G, Levinthal D (2000) Looking forward and looking backward: Cognitive and experiential search. *Administrative Science Quarterly* 45(1):113–137.

Girotra K, Terwiesch C, Ulrich KT (2010) Idea generation and the quality of the best idea. *Management Science* 56(4):591–605.

Goldenberg S (2011) BP's oil spill crowdsourcing exercise: 'A lot of effort for little result'. Technical report, The Guardian, retrieved (November 7, 2025), https://www.theguardian.com/environment/2011/jul/12/bp-deepwater-horizon-oil-spill-crowdsourcing.

Gomez-Zara D, Guo M, DeChurch L, Contractor N (2020) The impact of displaying diversity information on the formation of self-assembling teams. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, Hawaii).

Gomez-Zara D, Paras M, Twyman M, Ng J, DeChurch L, Contractor N (2019) Who would you like to work with? Use of individual characteristics and social networks in team formation systems. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK).

Gompers PA, Huang K, Wang SQ (2017) Homophily in entrepreneurial team formation, available at SSRN: https://ssrn.com/abstract=2973329.

Gurca A, Bagherzadeh M, Velayati R (2023) Aligning the crowdsourcing type with the problem attributes to improve solution search efficacy. *Technovation* 119(102613).

Hahn J, Lee G (2021) The complex effects of cross-domain knowledge on is development: A simulation-based theory development. *MIS Quarterly* 45(4):2023–2054.

Hahn J, Moon JY, Zhang C (2008) Emergence of new project teams from open source software developer networks: Impact of prior collaboration ties. *Information Systems Research* 19(3):369–391.

Hamman JR, Martínez-Carrasco MA (2023) Managing uncertainty: An experiment on delegation and team selection. *Organization Science* 34(6):2272–2295.

Hou T, Zhang W (2021) Optimal two-stage elimination contests for crowdsourcing. *Transportation Research Part E: Logistics and Transportation Review* 145(102156).

Hu M, Wang L (2020) Joint vs. separate crowdsourcing contests. *Management Science* 67(5):2711–2728.

Huang P, Ceccagnoli M, Wu D (2022) IT knowledge spillovers, absorptive capacity, and productivity: Evidence from enterprise software. *Information Systems Research* 33(3):908–934.

Javadi H Khasraghi, Hirschheim R (2021) Collaboration in crowdsourcing contests: How different levels of collaboration affect team performance. *Behaviour Information Technology* 41(7):1566–1582.

Jeppesen LB, Lakhani KR (2010) Marginality and problem-solving effectiveness in broadcast search. *Organization Science* 21(5):1016–1033.

Jian L, Yang S, Ba S, Lu L, Jiang LC (2019) Managing the crowds: The effect of prize guarantees and in-process feedback on participation in crowdsourcing contests. *MIS Quarterly* 43(1):97–112.

Jiang Z, Huang Y, Beil DR (2021) The role of feedback in dynamic crowdsourcing contests: A structural empirical analysis. *Management Science* 68(7):4858–4877.

Jiang ZZ, Huang Y, Beil DR (2020) The role of problem specification in crowdsourcing contests for design problems: A theoretical and empirical analysis. *Manufacturing & Service Operations Management* 23(3):637–656.

Jin Y, Lee HCB, Ba S, Stallaert J (2021) Winning by learning? Effect of knowledge sharing in crowdsourcing contests. *Information Systems Research* 32(3):836–859.

Johnson SL, Faraj S, Kudaravalli S (2014) Emergence of power laws in online communities: The role of social mechanisms and preferential attachment. *MIS Quarterly* 38(3):795–808.

Kauffman SA (1993) *The Origins of Order: Self-Organization and Selection in Evolution* (New York: Oxford University Press).

Knudsen T, Levinthal DA (2007) Two faces of search: Alternative generation and alternative evaluation. *Organization Science* 18(1):39–54.

Koh TK (2019) Adopting seekers' solution exemplars in crowdsourcing ideation contests: Antecedents and consequences. *Information Systems Research* 30(2):486–506.

Koh TK, Cheung MYM (2022) Seeker exemplars and quantitative ideation outcomes in crowdsourcing contests. *Information Systems Research* 33(1):265–284.

Korpeoglu G, Korpeoglu E, Tunc S (2020) Optimal duration of innovation contests. *Manufacturing & Service Operations Management* 23(3):657–675.

Lazer D, Friedman A (2007) The network structure of exploration and exploitation. *Administrative Science Quarterly* 52(4):667–694.

Lee HCB, Ba S, Li X, Stallaert J (2018) Salience bias in crowdsourcing contests. *Information Systems Research* 29(2):401–418.

Lee HCB, Mishra A (2025) Lost time in crowdsourcing contests. *Information Systems Research* 36(3):1652–1669.

Levinthal DA (1997) Adaptation on rugged landscapes. *Management Science* 43(7):934–950.

Levinthal DA, Workiewicz M (2018) When two bosses are better than one: Nearly decomposable systems and organizational adaptation. *Organization Science* 29(2):207–224.

Lifshitz-Assaf H (2018) Dismantling knowledge boundaries at nasa: The critical role of professional identity in open innovation. *Administrative Science Quarterly* 63(4):746–782.

Liu J, Kim K (2023) Designing contests for data science competitions: Number of stages and the prize structures. *Production and Operations Management* 32(11):3752–3772.

Liu TX, Yang J, Adamic LA, Chen Y (2014) Crowdsourcing with all-pay auctions: A field experiment on taskcn. *Management Science* 60(8):2020–2037.

Loch CH, Terwiesch C, Thomke S (2001) Parallel and sequential testing of design alternatives. *Management Science* 45(5):663–678.

Lykourentzou I, Robert Jr LP, Barlatier PJ (2021) Unleashing the potential of crowd work: The need for a post-taylorism crowdsourcing model. *M@n@gement* 64–69.

Lysyakov M, Viswanathan S (2020) Synthesizing winning strategies: What differentiates experienced solvers in crowdsourcing markets? *Proceedings of the 41th International Conference on Information Systems (ICIS 2020)* (Hyderabad, India (Virtual Conference)).

Lysyakov M, Viswanathan S (2023) Threatened by AI: Analyzing users' responses to the introduction of AI in a crowd-sourcing platform. *Information Systems Research* 34(3):1191–1210.

Majchrzak A, Malhotra A (2013) Towards an information systems perspective and research agenda on crowdsourcing for innovation. *Journal of Strategic Information Systems* 22(4):257–268.

Majchrzak A, Malhotra A (2020) *Unleashing the Crowd: Collaborative Solutions to Wicked Business and Societal Problems* (London: Palgrave Macmillan).

Majchrzak A, Malhotra A, Zaggl MA (2021) How open crowds self-organize. *Academy of Management Discoveries* 7(1):104–129.

Majchrzak A, More PHB, Faraj S (2012) Transcending knowledge differences in cross-functional teams. *Organization Science* 23(4):951–970.

Malhotra CK, Malhotra A, Bayus BL (2021) Get more ideas from the crowd. *MIT Sloan Management Review* 62(4):15–17.

Manshadi V, Rodilitz S (2022) Online policies for efficient volunteer crowdsourcing. *Management Science* 68(9):6572–6590.

Marks MA, Mathieu JE, Zaccaro SJ (2001) A temporally based framework and taxonomy of team processes. *Academy of Management Review* 26(3):356–376.

Mason W, Watts DJ (2012) Collaborative learning in networks. *Proceedings of the National Academy of Sciences* 109(3):764–9.

Mihm J, Schlapp J (2019) Sourcing innovation: On feedback in contests. *Management Science* 65(2):559–576.

Millhiser WP, Coen CA, Solow D (2011) Understanding the role of worker interdependence in team selection. *Organization Science* 22(3):772–787.

Miranda S, Berente N, Seidel S, Safadi H, Burton-Jones A (2022) Computationally intensive theory construction: A primer for authors and reviewers. *MIS Quarterly* 46(2):iii–xviii.

Mo J, Sarkar S, Menon S (2021) Competing tasks and task quality: An empirical study of crowdsourcing contests. *MIS Quarterly* 45(4):1921–1948.

Mollick E, Nanda R (2016) Wisdom or madness? Comparing crowds with expert evaluation in funding the arts. *Management Science* 62(6):1533–1553.

Muller-Trede J, Choshen-Hillel S, Barneron M, Yaniv I (2017) The wisdom of crowds in matters of taste. *Management Science* 64(4):1779–1803.

Nagaraj A (2021) Information seeding and knowledge production in online communities: Evidence from openstreetmap. *Management Science* 67(8):4908–4934.

Nan N, Tanriverdi H (2017) Unifying the role of IT in hyperturbulence and competitive advantage via a multilevel perspective of IS strategy. *MIS Quarterly* 41(3):937–958.

Niculescu MF, Wu D, Xu L (2018) Strategic intellectual property sharing: Competition on an open technology platform under network effects. *Information Systems Research* 29(2):498–519.

Nunley D (2020) Crowdsourcing innovation [podcast episode]. Small Steps, Giant Leaps, retrieved (November 7, 2025), https://www.nasa.gov/podcasts/small-steps-giant-leaps/small-steps-giant-leaps-episode-36-crowdsourcing-innovation.

Owens DA, Mannix EA, Neale MA (1998) Strategic formation of groups: Issues in task performance and team member selection. *Research on Managing Groups and Teams: Composition* 1:149–165.

Park S, Piezunka H, Dahlander L (2022) Co-evolutionary lock-in in external search. *Academy of Management Journal* 67(1):262–288.

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

Reagans R, Zuckerman E, McEvily B (2004) How to make the team: Social network vs. demography as criteria for designing effective teams. *Administrative Science Quarterly* 49(1):101–133.

Reagans R, Zuckerman EW (2001) Networks, diversity, and productivity: The social capital of corporate R&D team. *Organization Science* 12(4):502–517.

Riedl C, Seidel VP (2018) Learning from mixed signals in online innovation communities. *Organization Science* 29(6):1010–1032.

Riedl C, Woolley AW (2017) Teams vs. crowds: A field test of the relative contribution of incentives, member ability, and emergent collaboration to crowd-based problem solving performance. *Academy of Management Discoveries* 3(4):382–403.

Rivkin JW, Siggelkow N (2003) Balancing search and stability: Interdependencies among elements of organizational design. *Management Science* 49(3):290–311.

Rivkin JW, Siggelkow N (2007) Patterned interactions in complex systems: Implications for exploration. *Management Science* 53(7):1068–1085.

Robert LP, Dennis AR, Ahuja MK (2008) Social capital and knowledge integration in digitally enabled teams. *Information Systems Research* 19(3):314–334.

Robert LP, Dennis AR, Ahuja MK (2018) Differences are different: Examining the effects of communication media on the impacts of racial and gender diversity in decision-making teams. *Information Systems Research* 29(3):525–545.

Ruef M, Aldrich HE, Carter NM (2003) The structure of founding teams: Homophily, strong ties, and isolation among us entrepreneurs. *American Sociological Review* 68(2):122–195.

Sanyal P, Ye S (2023) An examination of the dynamics of crowdsourcing contests: Role of feedback type. *Information Systems Research* 35(1):394–413.

Siggelkow N, Rivkin JW (2005) Speed and search: Designing organizations for turbulence and complexity. *Organization Science* 16(2):101–122.

Simon HA (1962) The architecture of complexity. *Proceedings of the American Philosophical Society* 106(6):467–482.

Surowieck J (2004) *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations* (New York: Doubleday; Anchor).

Terwiesch C, Xu Y (2008) Innovation contests, open innovation, and multiagent problem solving. *Management Science* 54(9):1529–1543.

van Knippenberg D, De Dreu CK, Homan AC (2004) Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology* 89(6):1008–22.

Wegner DM (1987) Transaction memory: A contemporary analysis of the group mind. Mullen B, Goethals GR, eds., *Theories of Group Behavior*, 185–208 (New York: Springer-Verlag).

Wooten JO, Ulrich KT (2016) Idea generation and the role of feedback: Evidence from field experiments with innovation tournaments. *Production and Operation Management* 26(1):80–99.

Yang Y, Chen Py, Banker R (2010) Impact of past performance and strategic bidding on winner determination of open innovation contest. *Workshop of Information System and Economics* (St. Louis: MO).

Yang Y, Chen PY, Pavlou PA (2009) Open innovation: Strategic design of online contests. *Proceedings of the 30th International Conference on Information Systems (ICIS 2009)* (Phoenix: AZ).

Zaggl MA, Malhotra A, Alexy O, Majchrzak A (2023) Governing crowdsourcing for unconstrained innovation problems. *Strategic Management Journal* 44(11):2783–2817.

Zenger TR, Lawrence BS (1989) Organizational demography: The differential effect of age and tenure distribution on technical communication. *Academy of Management Journal* 32(2):353–376.

Zhang S, Singh PV, Ghose A (2019) A structural analysis of the role of superstars in crowdsourcing contests. *Information Systems Research* 30(1):15–33.

Appendix A Appendices

A.1 The *NK* Fitness Landscape Model

The creation of new technologies, products and services involves solving complex problems that require configuring a large number of interdependent elements (Simon 1962). The attempt to find high quality configurations of interdependent elements can be illustrated using the metaphor of a "rugged landscape" (Levinthal 1997). Each location on the landscape represents one possible configuration of elements, and its height represents the performance (or fitness) of that configuration. Finding a high quality configuration is thus akin to reaching a high peak on the rugged landscape. The *NK* fitness landscapes model provides a formal framework for constructing such landscapes computationally and for modeling agents' behavioral rules aimed at reaching high peaks.

A.1.1 Modeling the Problem Space

In our context, the fitness landscape represents the problem space that solvers must search for high quality solutions to configuring interdependent elements. The problem space involves N elements $d_1, d_2, \ldots, d_N$, and K interdependencies between each element and other elements. A configuration is thus represented by a N-element vector $\mathbf{d} =< d_1, d_2, \ldots, d_N >$, where d_i can take the value of 0 or 1. The interdependencies among N elements can be represented by an $N \times N$ influence matrix (INF), where $INF_{i,j} = 1$ if the fitness contribution of the i^{th} element c_i is influenced by the j^{th} element, or 0 otherwise. Figure A.1 shows an example of an influence matrix when $N = 12$ and $K = 3$. The fitness contribution of each element is influenced by its own specification (i.e., all diagonal cells (i.e., $i = j$) in Figure A.1 have the value of 1) and the specification of three other elements. For instance, the fitness contribution of the first element c_1 is influenced by the specifications of the second, fifth, and last elements according to Figure A.1.

$$
\begin{bmatrix}
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

Figure A.1: Example of an Influence Matrix (*INF*) when *N=12*, *K=3*

Given an influence matrix, fitness values are randomly generated. Specifically, for each possible specification of element d_i and K other interdependent elements, c_i is randomly drawn from a uniform distribution $U(0,1)$. For instance, the first element d_1 in Figure A.1 will have a total of 16 possible fitness contributions randomly drawn from the uniform distribution (i.e., $d_1 \in \{0,1\}$, $d_2 \in \{0,1\}$, $d_5 \in \{0,1\}$, $d_{12} \in \{0,1\}$). The overall fitness value of a configuration $\mathbf{d}$ is the average of all N fitness contributions c_i, namely $F(\mathbf{d}) = \frac{1}{N}\sum_i c_i$ where $c_i = c_i(d_i|k \text{ other } d_j\text{'s})$. Given the influence matrix as per Figure A.1, a configuration $\mathbf{d} = \langle 0,1,0,1,0,1,0,1,0,1,0,1\rangle$ has a fitness value:

$$
\begin{aligned}
F(\mathbf{d}) = \frac{1}{12}[& c_1(d_1 = 0|d_2 = 1, d_5 = 0, d_{12} = 1) + c_2(d_2 = 1|d_4 = 1, d_8 = 1, d_9 = 0) + \\
& c_3(d_3 = 0|d_6 = 1, d_8 = 1, d_{10} = 1) + c_4(d_4 = 1|d_2 = 1, d_7 = 0, d_{11} = 0) + \\
& c_5(d_5 = 0|d_6 = 1, d_9 = 0, d_{10} = 1) + c_6(d_6 = 1|d_1 = 0, d_3 = 0, d_8 = 1) + \\
& c_7(d_7 = 0|d_3 = 0, d_4 = 0, d_{11} = 0) + c_8(d_8 = 1|d_1 = 0, d_4 = 1, d_6 = 1) + \\
& c_9(d_9 = 0|d_2 = 1, d_6 = 1, d_{12} = 1) + c_{10}(d_{10} = 1|d_3 = 0, d_5 = 0, d_7 = 0) + \\
& c_{11}(d_{11} = 0|d_2 = 1, d_6 = 1, d_9 = 0) + c_{12}(d_{12} = 1|d_1 = 0, d_4 = 1, d_7 = 0)]
\end{aligned}
$$

This procedure of assigning fitness values to each configuration enables us to tune the ruggedness of the generated landscapes by modifying the degree (i.e., K)

and pattern (i.e., *INF*) of interdependencies.[1] When there are few interdependencies among elements (i.e., low K), the generated landscape is more smooth; conversely, when elements are highly interdependent (i.e., high K), the resulting landscape is more rugged. Figures A.2a and A.2b illustrate a smooth landscape with a single peak (i.e., the global optimum) and a rugged landscape with multiple peaks (i.e., a global optimum and numerous local optima), respectively. The ruggedness of the fitness landscape influences the difficulty of finding the global optimum. As shown by Figure A.2a, when the landscape is smooth, incremental local search enables agents to eventually reach the global optimum. However, when the landscape is rugged (see Figure A.2b), incremental local search tends to lead agents to local optima. What's worse, agents are likely to get stuck at those local optima because incrementally changing their configurations will always yield fitness values lower than the status quo.

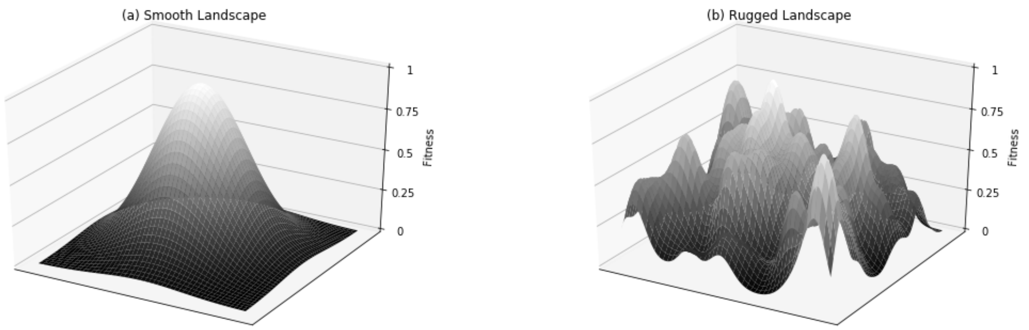

Figure A.2: Illustrative Examples of Fitness Landscapes

Notes: The real fitness landscape is a $N+1$ dimensional space (N dimensions for N elements and 1 dimension for fitness values). For illustrative purpose, Figures A.2a and A.2b are 3-D simplifications of the real fitness landscape. The x-axis and y-axis have no specific meaning other than to indicate the proximity of configurations.

Figure A.3 shows the average number of peaks on the fitness landscape across different levels of complexity (i.e., K) when $N = 12$ (i.e., the setup in our main analysis). Not surprisingly, when $K = 1$, there are only a few peaks on the landscape. As K increases, the number of peaks on the landscape significantly increases. When

[1] Whilst it is possible to specify the influence matrix to explore the impact of various interdependency structures (e.g., Rivkin and Siggelkow 2003, 2007), this study adopts a more general approach by randomizing these effects – i.e., we randomly generate influence matrices according to predefined parameters N and K and then use the stochastic procedure described above to generate fitness values for all possible configurations (i.e., the fitness landscape).

$K = 11$, there are around 300 peaks on the landscape. In other words, around 7.3% ($=\frac{300}{2^{12}}$) configurations are local peaks that hinder agents from reaching the global optimum. As such, the degree of interdependencies (i.e., K) translates into the complexity of searching for high fitness configurations on the landscape.

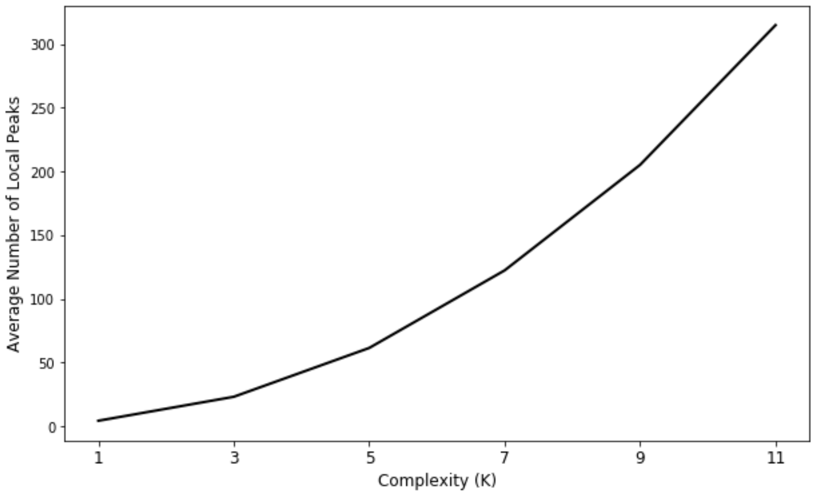

Figure A.3: Number of Local Peaks when *N=12*

A.1.2 Modeling Agents with Diverse Knowledge

As discussed, when the crowdsourced problem becomes broad in scope and requires the integration of knowledge from many different domains, there may be no individual solver who understands all aspects of the problem. Imperfect knowledge possessed by solvers can hinder solvers from fully understanding the problem, leading to incomplete mental representations of the problem. Imperfect knowledge also prevents solvers from developing complete solutions beyond the scope of their individual knowledge. To faithfully model these characteristics, we need to model agents' knowledge in a way that the lack of knowledge about elements outside their knowledge domain would hamper a focal agent's search on the landscape. To do so, we adopt Gavetti and Levinthal's model (2000) of cognitive simplification, which allows agents to have a simple, low-dimensional representation of a more complex, high-dimensional fitness landscape based on their imperfect knowledge. The actual modeling approach is to set an agent's perceived fitness value $\tilde{F}(\mathbf{d})$ of solution $\mathbf{d}$ by averaging the fitness values across all possible configurations of the elements outside of its knowledge domain D. Table A.1 provides a numeric example of how cognitive simplification is modeled. When there are 8 design elements (i.e., $N = 8$), an agent with knowledge about the first 6 elements would perceive a solution $\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, 0, 1 \rangle$ as $\langle 0, 1, 0, 1, 0, 1, *, * \rangle$, where $*$ indicates unknown elements. The perceived value equals to the average fitness across all the possible configurations of the unknown elements – i.e., all possible combinations of the last two elements in the third column of Table A.1.

Table A.1: An Illustrative Example of Cognitive Simplification

Agent's Knowledge	Configuration	Fitness across Unknowns	Perceived Fitness
$D = \{1, 2, 3, 4, 5, 6\}$	$\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, *, * \rangle$	$F(\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, 0, 0 \rangle) = 0.82$ $F(\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, 0, 1 \rangle) = 0.75$ $F(\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, 1, 0 \rangle) = 0.94$ $F(\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, 1, 1 \rangle) = 0.68$	$\tilde{F}(\mathbf{d}) = 0.7975$

This modeling approach does not mean that an agent actually knows the fitness values of all solutions and is able to calculate the average fitness values across configurations of unknown elements. Instead, the average value is simply an unbiased expected value that an agent would perceive a solution configuration given its imperfect knowledge. As illustrated in the third and fourth columns of Table A.1, the perceived fitness values of different configurations may be higher or lower than the actual fitness values due to a lack of knowledge. Yet, how the agent perceives all configurations on the landscape as a whole would be largely correlated with the actual fitness landscape.

Figure A.4 shows how the generated cognitive simplification correlates with the actual fitness landscape when $N = 12$ and $|D| = 6$ (i.e., the setup in our main analysis). Whilst the correlation between randomly generated fitness landscapes (the dashed line) is always around 0, the correlation between the cognitive simplification and the actual fitness landscape is above 0.6 when problem complexity is low (i.e., the leftmost point on the solid line). As the problem complexity increases, their correlation declines but remains significant and positive. These results are consistent with conventional wisdom. That is, as elements become tightly interdependent, a lack of knowledge of certain elements would exacerbate the imperfect evaluation of solution configurations.

To capture solvers' diverse knowledge, in each replication of the simulations, we always randomly set each agent's decision set D. Each design element has a equal

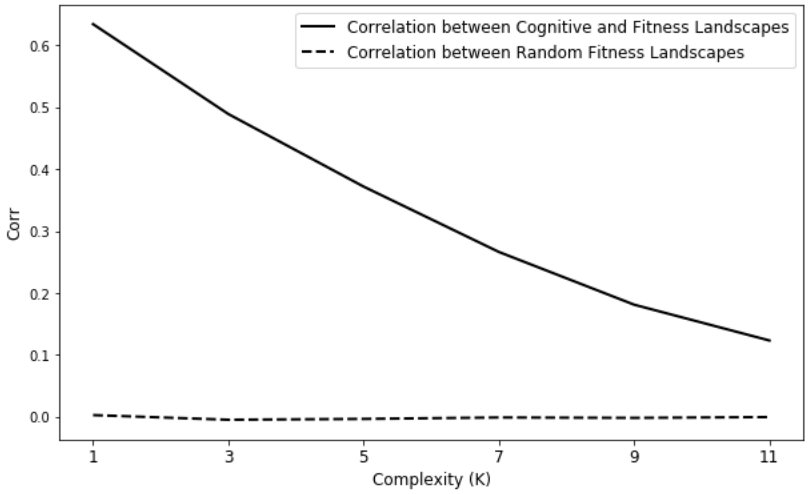

Figure A.4: Correlation between Cognitive Landscapes and Fitness Landscapes when *N=12*, $|D|$=*6*

probability of being selected into agents' decision sets. It is possible that certain elements being more commonly represented in the population of agents and others being less so. Therefore, we also experimented with alternative setups in which elements are chosen according to a nearly normal distribution (i.e., $Beta(2,2)$) and found that the alternative distribution only affects the size of the proposed effects without qualitatively altering our conclusions.

With imperfect evaluation, agents search the problem landscape in an iterative and incremental manner. Continuing with the example above, let's suppose that an agent's current solution configuration is $\mathbf{d} = \langle 0,1,0,1,0,1,*,*\rangle$, with knowledge about how to configure the first 6 design elements (i.e., $D = \{1,2,3,4,5,6\}$). During search, the agent would randomly change one design element within its decision set D and adopt the change if the new configuration is perceived as being better than the status quo based on the imperfect evaluation. In other words, a new configuration $\mathbf{d}^{'}$ would be selected from the 6 neighboring configurations:

$$\{\langle \underline{1},1,0,1,0,1,*,*\rangle, \langle 0,\underline{0},0,1,0,1,*,*\rangle, \langle 0,1,\underline{1},1,0,1,*,*\rangle,$$
$$\langle 0,1,0,\underline{0},0,1,*,*\rangle, \langle 0,1,0,1,\underline{1},1,*,*\rangle, \langle 0,1,0,1,0,\underline{0},*,*\rangle,\}$$

The new configuration $\mathbf{d}^{'}$ would be adopted to replace the current configuration $\mathbf{d} = \langle 0,1,0,1,0,1,*,*\rangle$ if $\tilde{F}(\mathbf{d}^{'}) > \tilde{F}(\mathbf{d})$. Otherwise, the current configuration $\mathbf{d}$ is retained.

In the extended model, we also examine the scenario where agents exhibit heterogeneous search capability and are able to consider more than one alternative configuration in each iteration. Continuing with the $N = 8$ and $D = \{1, 2, 3, 4, 5, 6\}$ example, suppose that the agent can consider 8 alternative configurations and the current configuration is $\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, *, * \rangle$. The agent would first consider all 6 neighboring configurations listed above and 2 additional configurations that involve changing two decisions – e.g., $\{\langle \underline{1}, \underline{0}, 0, 1, 0, 1, *, * \rangle, \langle 0, \underline{0}, \underline{1}, 1, 0, 1, *, * \rangle\}$. Among all considered configurations, the one with the highest perceived value $\tilde{F}(\mathbf{d}')$ will be eventually selected. The selected configuration $\mathbf{d}'$ will be adopted to replace the current configuration $\mathbf{d}$ if $\tilde{F}(\mathbf{d}') > \tilde{F}(\mathbf{d})$. In this way, the number of alternative configurations the agent can consider reflects its problem-solving capability. An agent with a higher level of problem-solving capability is able to consider more alternatives and assess the ramifications of changing more design elements at once within a given time frame (Rivkin and Siggelkow 2003).

In the extended model, we randomly vary the number of alternatives an agent can consider from 1 to $2 \times N_D$ (i.e., 12 in our main analysis). It is worth noting that the upper bound is set to sufficiently capture the advantages of agents with higher levels of search capability in altering and evaluating multiple design elements simultaneously. Further increasing the upper bound would only add to the computational burden without qualitatively changing the results.

A.1.3 Modeling Teaming

At the time of teaming t, each agent has a probability (p_{teamup}) to initiate a team-up invitation. If an agent u decides to initiate a team-up invitation as per p_{teamup}, it sorts other agents according to the quality signal on which it relies and considers possible teammates one by one. The invitation process stops if an agent on the ordered list accepts the invitation (i.e., a team is formed) or if no agents accept the invitation (i.e., no team is formed). The recipients v of team-up invitations also sort other agents according to the public quality signal and have a probability (p_{accept}) to accept the team-up invitation. When a team is formed, agents in the team are excluded from subsequent matches.

In our main analysis, we set p_{teamup} as a constant equal to 0.2 to approximate the proportion of teams on Kaggle. As for p_{accept}, we specify it as an exponential decay function of the extent to which the sender is favored according to the public quality signal. It takes the form of $\alpha \cdot (e^{-\beta x} + \gamma)$, where x is the relative order of the sender agent according to different quality signals. When rank information is used, x represents the relative order of agents' performance at the timing of teaming (i.e., $F(\mathbf{d})$; x takes the value of 1 if the sender agent is with the highest $F(\mathbf{d})$ among all solvers, so on and so forth). When knowledge-based information is used, x represents the relative order of the knowledge overlap between the sender u and recipient v (i.e., $|\boldsymbol{D}_u \cap \boldsymbol{D}_v|$; x takes the value of 1 if the sender agent is with a modest level of knowledge overlap when teaming takes place according to moderate knowledge overlap).[2] When capability information is used, x represents the relative order of the sender's search capability (i.e., the number of alternatives it can consider; x takes the value of 1 if the sender agent has the greatest search capability among all solvers).

To calibrate hyperparameters α, β, and γ, we conducted a grid search on the parameter space (i.e., $\alpha = \{0.1, 0.3, 0.5, 0.7, 0.9\}$, $\beta = \{0.1, 0.3, 0.5, 0.7, 0.9\}$, $\gamma = \{0.001, 0.003, 0.005, 0.007, 0.009\}$) with 200 populations of 100 agents to minimize the mean squared error (MSE) of the following terms: 1) the average proportion of teams in the population of crowd is 0.2; 2) the distribution of team member's relative order is as closed as to our empirical observation (i.e., Figure 2.2a); 3) the likelihood of teaming probability across x closely matches Figure 2.2b. We specify $\alpha = 0.5$, $\beta = 0.1$, and $\gamma = 0.001$ for our main analysis.

Unlike rank- and capability-based information, the extent of knowledge overlap varies among agents. An agent with an unfavored extent of knowledge overlap for a focal agent might be perceived as having a favored extent of knowledge overlap for another agent (i.e., a greater value of p_{accept}). As a result, agents are more likely to team up when they rely on the extent of knowledge overlap to determine possible teammates. α thus is normalized by multiplying it by 0.14 when agents team up according to knowledge overlap. In a similar vein, when we raise the number of agents in the crowd, the proportion of teams increases because agents each will

[2] In our primary analyses, agents are sorted by the extent of knowledge overlap (i.e., $\langle 6, 5, 4, 3, 2, 1, 0 \rangle$ for maximal knowledge overlap; $\langle 3, 2, 4, 1, 5, 0, 6 \rangle$ for modest knowledge overlap; and $\langle 0, 1, 2, 3, 4, 5, 6 \rangle$ for minimal knowledge overlap).

receive more invitations. As such, we normalize α by $\frac{100}{CrowdSize}$.

It is important to note that the specification of the probability of teaming might vary on different crowdsourcing platforms for different contests. As such, our main thesis, namely how teaming impacts the overall crowdsourcing effectiveness through affecting the interplay between the parallel path effect and the *identification*, *integration*, and *teamwork effects*, is built upon random team formation as a counterfactual baseline. Different specifications of the probability of teaming would only alter the size of different effects without qualitatively changing our conclusions.

Once a team is formed, agents in the team stick together for the remainder of the simulations. The next section explains how agents in teams jointly search the landscape after the timing of teaming.

A.1.4 Modeling Teamwork

Prior research on the *NK* fitness landscape model has documented various ways to model how a complex problem is decomposed and how subsets of design elements are allocated to two or more agents to collectively search the landscape (e.g., Rivkin and Siggelkow 2003, Siggelkow and Rivkin 2005). A key modeling decision is to determine the extent of specialization and integration (Baumann et al. 2018). Specialization refers to the labor division among agents. A team's specialization is jointly determined by agents' decision sets D, which are randomly generated as per section A.1.2, and how agents with diverse decision sets are teamed together (i.e., the teaming process in section A.1.3). Integration refers to the process through which agents with different labor divisions collaborate. Unlike groups in traditional organizations where hierarchies play an important role in collaboration, self-organized teams in crowdsourcing are often small and relatively more decentralized and as a result, all team members are more likely to share authority over accepting and rejecting what should be done. Therefore, we adopt the "liaison archetype" as per Siggelkow and Rivkin (2005). That is, a team member agrees to a new solution configuration only if, based on its evaluation, a payoff at least as high as that of the status quo is yielded.

To illustrate how agents in a team jointly search the problem landscape, let us suppose again that the solution configuration is $\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, 0, 1 \rangle$. Suppose that agent u with decision set $D_u = \{1, 2, 3, 4, 5, 6\}$ and agent v with decision set $D_v = \{3, 4, 5, 6, 7, 8\}$ team up. In this case, agent u would perceive the solution as $\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, *, * \rangle$, whilst agent v would perceive the same solution as $\mathbf{d} = \langle *, *, 0, 1, 0, 1, 0, 1 \rangle$. In each iteration, agent u would select an alternative solution configuration $\mathbf{d}_u^{'}$ following the rule described in section A.1.2 (i.e., changing one or two of the first 6 design elements). If the alternative is perceived as better than agent u's current solution $\mathbf{d}_u$ (i.e., $\tilde{F}_u(\mathbf{d}_u^{'}) > \tilde{F}_u(\mathbf{d}_u)$), the alternative would be proposed to agent v; otherwise, the status quo $\mathbf{d}_u$ would be proposed to agent v.

Once agent v receives the proposal, it would evaluate whether the proposed configuration leads to a perceived value at least as high as its current configuration. The proposal would be approved if $\tilde{F}_v(\mathbf{d}_u^{'}) \geq \tilde{F}_v(\mathbf{d}_v)$. If the proposal is approved by agent u, the solution configuration $\mathbf{d}$ is updated according to agent v's proposal. In order to keep the number of alternatives the team can consider identical before and after teaming, we assume that agent u and agent v each has a chance to propose and evaluate the other's proposal. Continuing the example above, after updating the solution configuration, agent v takes a turn to search for an alternative configuration according to its imperfect evaluation and propose it for agent u's evaluation. See Table A.2 for a numerical example of the whole procedure when agent v is with higher search capability (i.e., agent v is able to change two elements at once) and agent u is with lower search capability (i.e., agent u is only able to change one element at once).

Table A.2: An Illustrative Example of Joint Search

<table>
<tr><td rowspan="2">Input</td><td colspan="2">Solution Configuration $\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, 0, 1 \rangle$</td></tr>
<tr><td>Agent u
$D_u = \{1, 2, 3, 4, 5, 6\}$, $\mathbf{d}_u = \langle 0, 1, 0, 1, 0, 1, *, * \rangle$</td><td>Agent v
$D_v = \{3, 4, 5, 6, 7, 8\}$, $\mathbf{d}_v = \langle *, *, 0, 1, 0, 1, 0, 1 \rangle$</td></tr>
<tr><td>Step 1.1</td><td>Agent u proposes a new solution configuration based on its evaluation $\mathbf{d}_u^{'} = \langle 0, \underline{0}, \underline{1}, 1, 0, 1, *, * \rangle$</td><td></td></tr>
<tr><td>Step 1.2</td><td></td><td>Agent v accepts the proposed solution configuration because $\tilde{F}_v(\langle *, *, \underline{1}, 1, 0, 1, 0, 1 \rangle) \geq \tilde{F}_v(\langle *, *, 0, 1, 0, 1, 0, 1 \rangle)$</td></tr>
<tr><td rowspan="2">Step 2.0</td><td colspan="2">Solution Configuration $\mathbf{d} = \langle 0, 0, 1, 1, 0, 1, 0, 1 \rangle$</td></tr>
<tr><td>Agent u
$\mathbf{d}_u = \langle 0, 0, 1, 1, 0, 1, *, * \rangle$</td><td>Agent v
$\mathbf{d}_v = \langle *, *, 1, 1, 0, 1, 0, 1 \rangle$</td></tr>
<tr><td>Step 2.1</td><td></td><td>Agent v proposes a new solution configuration based on its evaluation $\mathbf{d}_v^{'} = \langle *, *, 1, 1, 0, \underline{0}, 0, 1 \rangle$</td></tr>
<tr><td>Step 2.2</td><td>Agent u rejects the proposed solution configuration because $\tilde{F}_u(\langle 0, 0, 1, 1, 0, \underline{0}, *, * \rangle) < \tilde{F}_u(\langle 0, 0, 1, 1, 0, 1, *, * \rangle)$</td><td></td></tr>
<tr><td rowspan="2">Outcome</td><td colspan="2">Solution Configuration $\mathbf{d} = \langle 0, 0, 1, 1, 0, 1, 0, 1 \rangle$</td></tr>
<tr><td>Agent u
$\mathbf{d}_u = \langle 0, 0, 1, 1, 0, 1, *, * \rangle$</td><td>Agent v
$\mathbf{d}_v = \langle *, *, 1, 1, 0, 1, 0, 1 \rangle$</td></tr>
</table>

We also assume that during joint search, agents build and update their meta-knowledge of knowledge possessed by others, which serves as a guide for identifying the knowledge needed (Wegner 1987). Meanwhile, team members will gradually become better aware of their teammates' knowledge and construct a shared understanding of the problem (Alavi and Tiwana 2002, Robert et al. 2008, 2018), which helps to align each other's search behaviors. Agents' evaluation of solution configurations gradually becomes more coherent because shared team cognition unifies the divergent cognitions towards a coherent understanding of the problem. The extent to which such a coherent understanding can be built depends on the degree of knowledge overlap. Greater common ground among team members' respective knowledge bases (i.e., more knowledge overlap) facilitates intra-group communication and coordination and reduces inconsistencies and conflicts (Espinosa et al. 2007, Cramton 2001). In the above example (i.e., Table A.2), greater common ground increases the likelihood of agents u and v working around the same design elements. When inconsistencies arise on these design elements, agents are more likely to be aware of the configurations of those design elements outside their knowledge domains that lead to the inconsistencies. Conversely, teams formed by members with no overlapping knowledge might miss the opportunity to build a shared understanding because each of them independently works on a subset of design elements.

As such, in our simulation model, an agent u has a probability of $p_{learn} \times |\boldsymbol{D}_u \cap \boldsymbol{D}_v|$ to incorporate the configuration of one element from its teammate v's decision set in evaluating alternatives $\mathbf{d}$. Continuing the example in Table A.1, if an agent u with decision set $D_u = \{1, 2, 3, 4, 5, 6\}$ would perceive a solution $\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, 0, 1 \rangle$ as $\langle 0, 1, 0, 1, 0, 1, *, * \rangle$. Now if agent u knows its teammate v has configured the seventh element as 1, the perceived value would equal to the average fitness across the possible configurations of the last element. Table A.3 provides a numeric example.

Table A.3: Cognitive Simplification Given Knowledge Overlap

Agent's Knowledge	**Configuration**	**Fitness across Unknowns**	**Perceived Fitness**
$D = \{1, 2, 3, 4, 5, 6\}$	$\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, \underline{1}, * \rangle$	$F(\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, \underline{1}, 0 \rangle) = 0.94$ $F(\mathbf{d} = \langle 0, 1, 0, 1, 0, 1, \underline{1}, 1 \rangle) = 0.68$	$\tilde{F}(\mathbf{d}) = 0.81$

A.1.5 Model Validation

In this validation experiment, we investigate the likelihood of finding the global optimum of the problem landscape only through parallel search (i.e., agents search independently) across varying levels of problem complexity. We vary the number of agents in the simulation from 0 to 500 so as to benchmark our model to the classical *NK* model and validate our conceptualization of crowdsourcing as landscape search. Figure A.5 shows the results.

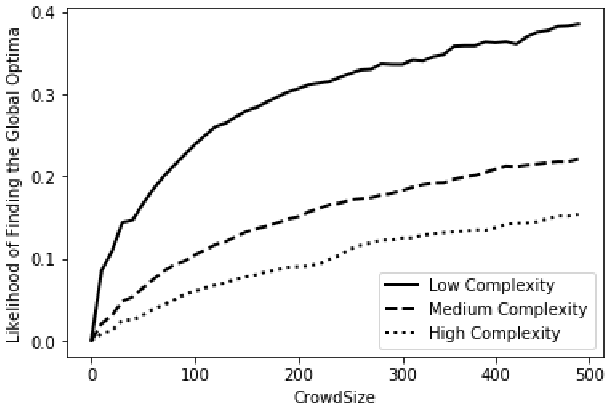

Figure A.5: Likelihood of Finding the Global Optima in Parallel Search

As the classical *NK* model would predict (Levinthal 1997), individual agents suffer a performance decline as problem complexity increases (see the performance gap between the solid line (low complexity) and the dotted line(high complexity)). A crowd of 500 solvers has a likelihood of 40% of generating the optimal solution for the solution seeker when the complexity is low, but the probability of finding the optimal solution reduces to about 10% for high-complexity problems. Moreover, we find that the likelihood of finding the global optima increases as the number of agents, suggesting the efficiency of the parallel path effect for crowdsourcing. Consistent with the crowdsourcing literature (Boudreau et al. 2011, Afuah and Tucci 2012), the simulation results also suggest that the parallel path effect would be more critical for problems with high complexity. The marginal return of increasing the number of solvers keeps decreasing when the complexity is low (see the slope of the solid line). Doubling the number of solvers from 200 to 400 only increases the likelihood of finding the best solution by around 30% (= $(0.39 - 0.3)/0.3$) when problem complexity is low. Conversely, the marginal return of increasing the number of solvers almost remains the same for problems with high complexity (see the slope of the dotted line). Doubling the number of solvers from 200 to 400 can almost double the likelihood of find the best solution when problem complexity is high.