Innovation Management in Digital Platforms

Nilam Kaushik

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

UCL School of Management

University College London

June 7, 2019

I, Nilam Kaushik, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Digital platforms are increasingly gaining prevalence as orchestrators of innovation, spawning new value chains, business models and organisational forms. Platforms have also changed the way firms manage various aspects of their innovation process. Improving innovation success is imperative for organizations in a growing platform-based economy. To contribute to our understanding of innovation management in these emerging contexts, I present three original research projects that provide an empirically based understanding of the operational drivers of successful innovation management in emerging digital platforms with econometric studies and field experiments.

In the second chapter, I quantify and characterize product evolution in mobile app development and examine the market performance implications of sequential innovation. I take a "search" perspective on how firms add new features and attributes into their digital products in successive product versions and find conditions under which search is associated with higher market performance.

Increasing participation from marginal users who bring diverse skills, knowledge and experience can be instrumental for innovation success involving distant search through crowdsourcing. In the third chapter, I provide insights from a randomized controlled trial from a technology-based crowdsourcing platform and investigate whether and to what extent gender based preferences explain the underrepresentation of women in technology-based work. My findings provide counterintuitive insights into heterogeneous gender preferences for tech-based work.

In the fourth chapter, using a natural language processing technique, I study the recombinant breadth and atypicality of crowdsourced contributions to complex

problems and explore which patterns and knowledge configurations are more or less likely to be associated with proposal success.

Impact Statement

Consistent with its practical relevance, there has been a growing interest among management scholars to understand the antecedents of innovation and how to manage and organize for innovation in an increasingly digital economy. Through the empirical settings of three different digital platforms I leverage a rich methodological toolbox combining econometrics, text mining and information retrieval and a randomized controlled trial to study topics around innovation management in digital platforms in this thesis. I have linked digital innovation to real operational and societal impact in this thesis and my work provides practical managerial insights as well as policy recommendations.

In the second chapter, I provide an empirically based understanding of sequential innovation by mobile app firms and its performance implications in mobile app development, an important and yet understudied industry in the existing operations management literature. I provide actionable insights into how mobile app development firms can improve their search process through which they incorporate new features and attributes in their products.

Users are an asset in crowd-sourcing and increasing participation from marginal users who bring diverse skills, knowledge and experience can be instrumental for the innovation success of organizations that increasingly rely on crowdsourcing for distant search. Building on the observation that women are under-represented in tech-based crowdsoucing work, in the third chapter, I provide insights from a randomized controlled trial on a technology-based crowdsourcing platform to investigate whether and to what extent gender based preferences may explain the under-representation of women in tech. I suggest policy remedies for

the leaky pipeline.

In the last chapter, I draw from the literatures on search, crowdsourcing, and knowledge recombination to study which kinds of proposals for crowdsourced complex social problems are evaluated successfully by gatekeepers. I explore whether text-based measures of previously established conceptualizations of search and successful knowledge recombination can be used to identify useful solutions in crowdsourcing initiatives. My work can serve as a basis for future research that studies innovation and its management in digital platforms .

Acknowledgements

I would like to thank my dissertation advisor, Dr. Bilal Gokpinar for his support and guidance throughout my doctoral studies and for giving me ample opportunities to develop as a researcher. I am also grateful to Dr. Kevin Boudreau for his mentorship, support and inspiration since 2015. I will cherish our journey together at the IOT Open Innovation Lab. I would like to thank Dr. Kamalini Ramdas and Dr. Sruthi Thatchenkery for their detailed and valuable feedback on my dissertation.

I also thank my fellow colleagues and cohort mates at UCL- Jung Won Lee, Hang Ren, Xiaojia Guo, Prateek Raj, Yufei Huang, Janice Sanchez, Miros Zohrevand, Philip Cornelius, Wenxin Xie, and Poornika Anantharamakrishnan for the wonderful time we have spent at Boston House and later at the Canary Wharf office through these years. I would like to acknowledge my colleagues at the IoT Open Innovation Lab - Abhinav Kharbanda, Patrick McGrath, Austen Keene, Katie Wilhoit, Dyan Khor, and Afan Ahmad Khan for their support. I also would like to thank Tanu Nahar and Pooja Gaikwad for being amazing roommates in Boston. I would also like to thank Enrico Forti, Anil Doshi, Dave Chapman, Milan Miric and Kenny Ching for their advice and fruitful conversations about my research.

I am grateful to my parents, Mohinder and Sunita Kaushik, for their lifelong support and inspiration and would like to dedicate this thesis to them. I also would like to thank my brother, Pashupati, who has been a constant source of strength throughout my studies. This journey would not have been possible without the unwavering support and encouragement of my husband, Subodh Sharma, who has been with me all the way. I also appreciate the support extended to me by my father and mother-in-law, Vishnu Dutt Sharma and Sudha Sharma. I would like to

acknowledge the contributions and support of my uncle and aunt, Shri Laxmi Chand and Smt. Satyavati. I am grateful for the guidance of Swami Tyaganandaji and Pravarajika Vivekpranaji through the last few years of my doctoral studies. Lastly, I would also like to thank the UCL School of Management for their administrative and academic support throughout my studies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Digital platforms[1] integrate digital technologies with new business models and have spawned new products and services by shaping new business and organisational forms and models, thereby transforming entire industries (Parker and Van Alstyne 2017). Compelling disruptive crossovers such as in entrepreneurial finance (e.g. Kickstarter), mobility (e.g Uber), hospitality (e.g. Airbnb) and more complex ecosystems such as the Apple Store which offer digital products (developed by external developers and sold to consumers), are based on the digital platform logic[2]. Platforms have emerged as the locus for innovation activities for firms across various industries. Large firms create value these days by developing platforms and platform ecosystems rather than single products (Yoo and Kim 2012). The most innovative firms today such as Apple, Google, 3M, Amazon, Microsoft, Tesla etc. were pioneers in developing digital platforms as an innovation strategy.

Successful platform ecosystems involve interactions between organizations and entities outside the firm (Boudreau and Jeppesen 2015) and typically create more value than a single firm can create. New organizational forms called "problem solving organizations" (Afuah and Tucci 2012) that involve dynamic pairing of problems and solutions have also emerged. This includes crowd-sourcing platforms where participating entities are external crowds or users that bring diverse knowl-

[1] A platform is defined as a "building block, providing an essential function to a technological systemwhich acts as a foundation upon which other firms can develop complementary products, technologies or services" (Gawer 2009)

[2] These are different from non-digital or traditional platforms which are beyond the scope of this work see (de Reuver et al. 2018)

edge and expertise (Afuah and Tucci 2012, Howe 2008). Crowdsourcing platforms can be large firms or intermediaries between large firms and the crowd, such as in the case of Innocentive, a platform that hosts scientific problem solving challenges on behalf of organizations to a large problem solver network.

Innovation is an imperative for organizations to stay competitive. Traditionally, activities related to new product/service development have been conducted within firm boundaries. Digital platforms facilitate the innovator's search, sourcing and gathering of information from customers, users, suppliers, competitors etc in new ways. Today, platforms have enabled heterogeneous actors (Boudreau 2012b) to participate and create value in the digital economy in unprecedented ways. My research is motivated by the increasing prevalence and economic significance of digital platforms as orchestrators of innovation (Nambisan et al. 2017). How do firms offering digital products (marked by the unique property of reprogrammability) within in today's platforms manage innovation to stay competitive? What are some challenges faced in new organisational forms such as crowdsourcing and how do crowdsourcing platforms organize and manage innovation to create value? In the three chapters of my thesis, I provide an empirically based understanding of the operational drivers of successful innovation management in emerging digital platforms with econometric studies and field experiments. Specifically, in the three chapters of my dissertation, I study i) the management of the sequential innovation process by mobile app firms, ii) user participation as it relates to the effect of gender-based preferences on participation in tech-based crowdsourcing and tech-based work more generally, and iii) whether established measures and patterns of knowledge recombination are more or less likely to be associated with success of ideas in crowdsourcing.

This thesis consists of three main parts. In the first part (Chapter 2), I study how mobile app firms manage product development post-initial launch through sequential innovation. Mobile app development platforms facilitate rapid sequential product development which presents an interesting dynamic for product innovation management as the scope and features of the digital product (mobile app) continue

to evolve. A key operational question is what features and attributes to add into existing products to generate value and attract new customers. I leverage text mining and information retrieval techniques to study the rich information in the release notes of apps and characterize the sequential innovation through a "search" lens (Katz and Allen 1982, Clark and Fujimoto 1999, Cusumano and Nobeoka 1992, Iansiti 1998, Katila 2002) at each version release and link this with market performance in a dynamic estimation model.

Adding new insights into the classical novelty vs risk/uncertainty trade-off in product innovation, I find that in mobile app development greater search dissimilarity (searching far from the neighborhood of search paths explored by the firm previously) and greater market-oriented search (similarity between the focal firms update with respect to the recent updates of its competitors) are associated with higher market performance. My results also suggest that a rapid release of updates dampens potential market benefits that the mobile app developers may gain from market-oriented search while the benefits of search dissimilarity are positively moderated by the functional maturity of the app. These findings offer new and potentially actionable insights to help managers systematically improve the innovation 'search' process by which they incorporate new features and attributes into mobile apps.

In the second part and third part of the thesis (Chapters 3 and 4 respectively), I study issues arising in crowdsourcing pertaining to managing and organizing for innovation success. Crowdsourcing is an appealing avenue for firms seeking innovative ideas and solutions and has the potential for creating competitive advantage (Malhotra and Majchrzak 2014). By approaching problems in unconventional ways, marginal users can prove to be assets for generating novel ideas and solutions in crowdsourced problem solving (Jeppesen and Lakhani 2010). Users are an asset in crowdsourcing and increasing participation from marginal users who bring diverse skills, knowledge and experience can be instrumental in a crowdsourcing platform's growth and success. During my doctoral studies, collaborating with Northeastern University's IoT Open Innovation Lab, I was instrumental in developing a two-

sided crowdsourcing platform that brings together solvers and seekers (companies) to ideate, develop and prototype Internet-of-Things (IoT) applications. I present results from a large-scale field experiment on 112,770 men and women from a wide range of ages and fields of training to study responses of subjects to participate in a platform based technical IoT learning-by-doing program. My experimental findings provide interesting and counter-intuitive insights into heterogeneous gender preferences for tech-based crowdsourcing work, and tech work in general.

Existing views predict that female participation would be lower in a tech-intensive activity. Indeed, I find that females overall participated less than males in this technical program, especially under competition. However, in this large and representative sample, I find no single monolithic gender-based tendency. Most notably, subjects in Engineering have the very opposite tendency; among individuals trained in Engineering fields, the mean female participation was in fact higher than that of males, even under competition. These results are consistent with lifelong gender-based separation and stratification in the workforce producing distinct subgroups of men and women. These findings have implications for the pipeline of technically skilled workers.

The recombination of ideas and knowledge forms the basis for innovation (Fleming 2001). Organizations have taken to externalizing their innovation efforts (typically problem solving and idea generation) by engaging in distant search through crowdsourcing. While crowdsourcing provides firms the potential to access novel solutions, firms are faced with the challenge of sifting through and filtering large numbers of contributions to identify useful contributions. The literature on innovation highlights the benefits of recombinant breadth, distant search and atypicality of knowledge combination in generating impactful inventions and technological innovation. In the third part of the thesis (Chapter 4), I study which kinds of crowdsourcing ideas get successfully evaluated by domain experts on a platform that brings together individuals to design solutions for complex social problems. Through a novel natural language processing method, I decompose proposed ideas into constituent topics or latent themes and operationalize previously established

measures of successful innovation such as atypicality, recombinant breadth to study whether certain patterns or combinations are associated with successful evaluation in crowdsourcing and explore whether computational techniques may be used by organizations to identify useful contributions.

Note: I am the first author on all chapters. Chapter 2 and 4 are joint work with Dr. Bilal Gokpinar and Chapter 3 is joint work with Dr. Kevin Boudreau.

Chapter 2

Sequential Innovation in Mobile App Development

2.1 Introduction

Sequential innovation, that is, the introduction of improved versions of existing products is prevalent in technology-based industries. Existing literature on sequential innovation has predominantly studied established industries such as durable goods, industrial products or consumer electronics with a primary focus on firms' patenting, pricing and product architecture (i.e., modular upgradable design) decisions by also typically considering monopolist firms (Bessen and Maskin 2009, Kornish 2001, Ramachandran and Krishnan 2008, Krishnan and Ramachandran 2011). Today, the quintessential example of rapid sequential innovation is software-based products such as mobile apps where a large number of firms compete with each other through dynamic and mostly transparent (both to customers and to competitors) innovation efforts by continuously adding new features and attributes to existing products. Here, unlike in traditional product development settings, search takes place continuously (MacCormack et al. 2001). Such new forms of sequential innovation introduce unique challenges for innovating firms beyond patenting, pricing and architecture decisions. In developing successive versions of the same product, mobile app developers keep the core innovation and incrementally add new features

and attributes with a hope to generate value and attract new customers.[1]

For example, as of December 2018, the popular social media app Instagram has released 75 major versions, and many smaller versions since its initial launch in October 2010. In Jan 2011, it introduced 'hashtags' to help users discover photos by users. In September 2011, it added live filters, instant tilt shift, a border option and one-click rotation. Lately, in August 2017 it has added multi-image posts, allowing users to combine 10 images or videos at a time. Through updated versions of the same product with additional features and attributes, Instagram generated value and was able to attract significant number of new customers over time (see Online Appendix A.1).

If new features and attributes can have such significance in providing value and attracting new customers, a key operational question for mobile app developers is what features and attributes to add into existing products in successive versions? More fundamentally, in their 'search' for new features and attributes for greater market success, where to get ideas from? After all, in such a "high-velocity" (Brown and Eisenhardt 1995) and "hyper-competitive"(Datta and Kajanan 2013) environment with high degree of product transparency, low barriers for creating content and a weak IP regime, in their search for new attributes and features, firms can turn to their direct competitors as a source of innovative ideas, they can explore other broad product domains, or they can simply build on their own previous versions by recombining existing features and attributes. While each of these strategies may make the resulting product more appealing to consumers, in contrast, they may be a hindrance for market success.

In this paper, we aim to empirically address the aforementioned question by employing a unique dataset and using novel text mining and information retrieval techniques. Specifically, following the existing literature, we take a "search" perspective of the product innovation process (Katz and Allen 1982, Clark and Fujimoto 1999, Cusumano and Nobeoka 1992, Iansiti 1998, Katila 2002), and examine

[1]While our study focuses on new features and attributes in successive versions of applications, a notable portion of version updates also involve corrective maintenance (i.e., bug fixing), as such, we control for such retrospective changes in our models while examining the market implications of new features and attributes.

the relationship between three types of search activities during sequential innovation and market performance. Namely, we study market performance implications of i) dissimilar search (i.e., incorporating new features and attributes that are distant from existing ones), ii) search breadth (i.e., incorporating new features and attributes from other domains, and iii) market-oriented search (i.e., incorporating new features and attributes from direct competitors) in mobile app development. In addition, we explore the possible effect of an app's functional maturity in governing the relationship between dissimilar search and market performance. Finally, considering the potential interplay between competitor dynamics and release timing decisions, we also explore the follow-up question on whether introducing a rapid release amplifies or dampens any relationship between market-oriented search and performance.

First, in searching for new features and attributes for successive versions of their products, developers can incorporate previously unexplored and distant search paths (i.e., dissimilar search) which may help improve their understanding of the structure of the knowledge landscape (Ahuja and Katila 2004) and potentially make major improvements in the new versions (Katila and Chen 2008). In contrast, considering potential risks and uncertainties associated with market response, developers can resort to variations of existing features and attributes (Katila and Ahuja 2002). Such reuse of familiar components and refinement of familiar combinations (Fleming 2001) can be very effective and it can lead to highly useful innovations as a result of synthesis of well-known technical information (Utterback 1996). Consequently, our study examines these contrasting perspectives and sheds light on the relationship between search dissimilarity and market performance when innovation is sequential in nature.

Secondly, developers can turn to other domains to source innovative ideas for the next version of their products. Searching broadly outside of a firm's usual domain (i.e., search breadth) may increase their available knowledge pool and bring distinctive new variations (Katila and Ahuja 2002). However, such variety can also bring additional uncertainty, reliability concerns and noise for the consumers (Flem-

ing 2001, Martin and Mitchell 1998), and inefficiencies to the innovation process (Henderson and Clark 1990). Considering these, it is not clear whether incorporating features and attributes from other domains would be an asset or liability for mobile app developers.

Thirdly and more importantly, intense competition is one of the defining characteristics of the mobile app industry. In introducing successive versions of their products, developers can either keep a close eye on their competitors and try to incorporate competitors' new features and attributes into their own releases, or instead they can prefer being less reactive and responsive to market changes, and pursue their own search paths. While the former responsive policy (i.e., market-oriented search) could help firms remain competitive and relevant in the marketplace, it may also result in diminished novelty and lack of focus (Köhler et al. 2012) from the consumers' viewpoint. As such, market-oriented search could be a help or hindrance for market success.

Our study contributes to the operations management literature by providing an empirically based understanding of sequential innovation and its market performance implications in mobile app development, a paramount industry in terms of size, scope and potential[2], which has, however, received little attention in the existing operations management literature. While previous studies on sequential innovation mostly analyzed monopolist firms in established markets, unique characteristics of the mobile application industry call for new scholarly attention as innovation efforts are transparent, competition is intense, market players are in flux and there is no clear market or industry structure (Eisenhardt and Martin 2000). In addition, availability of highly granular data for successive versions of mobile apps over time as well as corresponding market data allows us to directly link and empirically investigate sequential innovation efforts with market performance.

Our findings based on text-mining and information retrieval analyses reveal that greater search dissimilarity in successive versions of mobile applications is associated with greater market performance. That is, firms can significantly benefit by

[2]By 2022, the mobile app economy is expected to be worth $6.3 trillion (Molla 2017).

searching far from the neighborhood of the search paths they explored previously. As there may be a limit to deriving benefits from refining or recombining existing features, introduction of original and fresh features result in market success. We also find that an app's functional maturity amplifies this positive effect presumably due to reduced uncertainties and perceived risks of new features during the later stages of an app's life cycle. However, we find no evidence for beneficial effects of searching broadly across other sub-categories. Interestingly, we find that the greater the market-oriented search, that is, the higher the similarity of the focal firm's new features and attributes with respect to the recent additions of its competitors, the higher the market performance. This finding suggests that the market rewards those developers with a responsive policy to its competitors' product innovation efforts. Our results also indicate that rapid introduction of an update reduces potential market benefits that mobile app developers may gain from market-oriented search. That is, while incorporating competitors' recent features into their own apps seems to be beneficial for developers, rushing to incorporate such features dampens the benefits.

2.2 Mobile application development

Mobile applications are 'search' or 'experience' goods. Mobile applications have multiplied the capabilities of the mobile internet and have the potential for transforming business models across a range of industries including retail, banking, media, among others (Kavadias and Ladas 2015). The steady rise in smartphone adoption and usage has also led to the explosive growth of mobile applications. Another important contributor to such growth has been the introduction of mobile app stores, such as Apple's App Store and Android's Google Play store. These mobile application stores have been pivotal to innovation in the mobile apps industry. As of January 2017, there were 2.2 million apps available for download on the Apple App Store and 2.7 million apps available on Android's Google Play store (Vaidos 2017). By 2022, the app economy will be worth $6.3 trillion and the main driver for this growth is expected to be the purchase of goods and services through mobile applications (Molla 2017). As mobile application development becomes a crucial

part of the digital business strategy for firms (Bharadwaj et al. 2013), understanding the process of app development and its implications become essential.

Mobile app platforms are two-sided platforms (Rochet and Tirole 2003, Choudary et al. 2016), with growing demand from consumers on one side and a large number of producers (mobile app developers) on the other side. The availability of developer friendly Software Development kits (SDKs) and decreasing technology costs have lowered the barriers for creating content. Moreover, app distribution efforts are minimized by the platforms, making it seamless for developers to provide sequential updates to their product. However, an explosion of mobile apps and publishers has made this a 'hypercompetitive' setting where it is challenging for developers to attract consumer attention (Datta and Kajanan 2013). Moreover, the effortlessness with which apps can be installed and deleted and the prevalence of free versions of apps makes it easy for customers to try out new apps, which can eventually materialize into substitution.

Software is characterized by 'rapid and sequential' reuse and recombination (Cohen and Lemley 2001). Innovation in the software industry is sequential and cumulative (Parker and Van Alstyne 2017). In fact, in the case of software, the post initial launch phase may account for as much as 90% of the development activities (Bennett 1996). Recombination and reuse of ideas is also prevalent in the mobile app development setting. Waze, a popular navigation app for iPhone, is an example of recombination of two ideas- location sensing and social media. Additionally, mobile app platforms ease the delivery of content by enabling firms to release incremental changes to their apps. The ease of recombining digital modules facilitates the firm's ability to experiment with content by releasing new features. Subsequent to a release, it can learn the market reaction and learn from competitor activities and further experiment.

While the app's description field provides users with a general overview of its functionality, each app 'update' (or version) contains developer reported information about new features, enhancements, and defect fixes in the text-based descriptions of their software release notes. An example of this is given in Online

Appendix A.2. These release notes are publicly visible and can help potential customers in their purchase decisions, and inform existing customers about changes that went into the update. Such text information, though unstructured, contains certain keywords related to software development activities.

Unlike in the case of traditional software development or open source software projects where code-level changes can be observed with ease, software in the mobile apps setting is a black-box as apps are hosted in the ecosystem of proprietary app stores. Therefore, since code-level changes cannot be observed, an alternative way to analyze the nature of product development in this setting is to look at publicly available app release notes.

The software engineering literature has noted that release notes are a good indicator for the maintenance and evolution activities of the underlying software (Baysal and Malton 2007, Yu 2009). However, in order to understand the nature of product development and search in the mobile apps setting, one would have to carefully analyze thousands of version updates individually, as such analyzing each update manually would be a non-trivial exercise. Therefore, we employ novel text-mining and information retrieval techniques to study the rich information in the release notes of apps. We then characterize developers' search behavior at each version release and link this with app performance in a dynamic estimation model.

2.2.1 Basic economics of setting

In what follows, we first highlight the key mechanics of our mobile apps setting, and in the subsequent section we propose an outline of how prospective customers make app purchase decisions. In turn, in reviewing the relevant literature and building our hypotheses, we will refer to these basic economics and mechanisms of our setting.

In comparison to traditional product development settings (such as the automotive, medical, robotics etc.) that have been widely studied in the previous literature, there are a few salient features of the mobile apps setting. The first is the ability of firms (developers) to release digital product updates more easily and frequently in comparison to physical products which require greater amount of resources (capital, labor, time) and greater number of steps including manufacturing.

Innovation in the software industry is sequential and cumulative (Parker and Van Alstyne 2017) in nature. The setting is characterized by rapid sequential innovation. This kind of update practice is not specific to mobile applications but is also commonly seen in desktop computer software market. In the case of the latter, however, revenue is generated upon reaching a certain level of penetration from reselling upgraded software to existing users rather than from new customers (Sankaranarayanan 2007). In the case of mobile apps, though, a common practice is to release updates free of charge to the installed base users. Therefore, a primary goal of product upgrades is to attract new customers rather than generating additional revenue from the current installed base. As such, our focus in this study is the ways through which mobile app developers add new features and attributes to generate value and attract new customers.

In traditional product development environments, new products incur significant financial and non-financial risk both for firms and prospective customers. Importantly, product safety issues, severe malfunction or defects can result in costly product withdrawals or recalls (Shimbun 2013). In contrast, in digital goods, defects mostly manifest as software errors and are generally perceived by consumers as glitches which can be fixable. They also seldom create any permanent or irreversible damage to the product. The ability of fixing software defects post- release is a key distinguishing factor between digital and physical products (Picker 2005). Indeed, while consumers do not like defects, and crashes or errors may result in reduced customer loyalty to an app; developers' ability to fix issues and reverse defects by releasing updates seamlessly through the underlying digital platform is a major virtue of the mobile apps setting. Therefore, for prospective customers, perceived risks associated with updated products are arguably lower in the case of digital goods such as mobile apps than more traditional products. Additionally, the mobile apps setting has lower barriers to entry, and therefore, imitation is prevalent and facilitated by the ease of content dissemination through updates. This sets it apart from other digital goods. We next discuss prospective customers' purchase process for productivity apps, the focus of the present study.

2.2.2 App purchase process

As Productivity apps are utilitarian in nature, we expect that their purchase is driven largely by need rather than want, unlike in the case of consumption of hedonic gaming apps. The app discovery process is facilitated through Apple's search that helps users navigate across a variety of application genres and find relevant apps, and through search keywords. The underlying IOS platform has created broad categories, such as Health and Fitness apps, Gaming apps, Productivity apps, and so on, to help users find relevant products. Developers are enforced to choose a primary category that accurately reflects the app's core experience in order to get listed on the App Store. Search traffic accounts for the majority of app download traffic (Leuwer 2018, Panzarino 2016).

We suggest that in our case of digital utility goods (mobile apps on productivity), consumers have a particular need (e.g., a note taking app) and search for an app in the Appstore to fulfill that need. Online settings such as platforms allow consumers to gather information and learn about the product through content such as detailed descriptions provided by the sellers which help them make purchase decisions (Klein 1998, Lynch Jr and Ariely 2000, Huang et al. 2009). Therefore, without product ownership, knowledge of the listed attributes and functional aspects of a product offering help prospective customers familiarize themselves with the product and reduce any possible concerns and perceived risks they may associate with it (Goering 1985). Through new version release notes, mobile app firms reveal critical and updated information about their products and its key attributes.

When making purchase decision for a particular product category, prospective customers typically consider a small set of competing products that they consider seriously, which is known in the literature as the consideration set (Kotler and Armstrong 2010). A consideration set typically involves 3-5 products left after a person has narrowed down their choices based on their own personal screening criteria. We contend that for a digital product such as a mobile app, after identifying the need (e.g., a note taking app), a prospective customer would undertake a two-stage process, first filtering available alternatives and then undertaking a detailed analysis of

the reduced (consideration) set before making a final decision (Roberts and Lattin 1991). As such, prospective customers browse the overall description and other attributes of the app (e.g., previous download numbers, rating, price, etc.) which are readily available, and review the descriptions of multiple competing products to get a sense of the state-of-art product offering. Such information is provided in an app's update release notes with the latest changes and features.

2.3 Related literature

Our work builds upon and extends existing research on sequential innovation. Previous work on sequential innovation has predominantly focused on moderately dynamic markets where change occurs mostly along predictable paths in a relatively stable industry structure with defined market boundaries, making it conducive for firms to turn to strategies such as modular upgradability. While introducing advanced technologies to the market is unavoidable, firms have to bear in mind consumers' disinclination to pay for new products every time technology gets replaced. In order to minimize the cost incurred by the customer in keeping up with new technology, firms design their products to be modularly upgradeable (Ramachandran and Krishnan 2008), allowing customers to select desired components for upgrade. Such an upgrade path reduces the cost incurred by customers for replacing the entire product and may also potentially reduce waste (McDonough and Braungart 2002). A stream of literature in product innovation has studied the decisions made by firms managing rapid sequential innovation through modular upgradability, such as those related to product design (Ramachandran and Krishnan 2008), release timing, pricing (Kornish 2001), and assessing the appropriateness of modular upgradability for different types of markets and products (Krishnan and Ramachandran 2011), mostly with analytical models.

A stream of literature in operations management explores the strategies faced by firms in managing innovation in dynamic environments faced with uncertainty (Sommer and Loch 2004, Sommer et al. 2009, Pich et al. 2002), namely trial and error learning and selectionism, both through analytical models and through em-

pirical examination of novel startup ventures. Our mobile app development setting is indeed dynamic and characterized by uncertainty, however, rather than focusing on the "search for new products we explore the series of activities performed by the app developers past the initial launch of the product to keep it competitive. In a sense, our setting is consistent with trial and error learning but involves incrementally improving the same product as adjustments are made to the development process based on new information.

A relatively understudied facet of sequential innovation is high-velocity and dynamic environments such as mobile app development where modular upgradability may not be feasible. Mobile app firms release sequential app updates involving 'rapid and sequential' reuse and recombination (Cohen and Lemley 2001) in which it is easy to have the entire product upgraded (as opposed to modular upgrades) at no additional cost to existing customers. Furthermore, since mobile apps tend to be stand-alone products, this also minimizes backward compatibility, unlike in the case of products such as cameras that consist of several components such as the main body and lenses.

In sequential innovation, successive invention builds on the preceding one. Much of the work around sequential innovation in the economics literature has focused on models of a single invention with an infinite sequence of quality improvements examining issues such as the division of profit between the initial inventor and subsequent follow-on inventors (Green and Scotchmer 1995), and the role of patents in encouraging invention (Bessen and Maskin 2009), and protecting early innovators (Scotchmer 1996).

While the above-mentioned mostly analytical studies in the product development and economics literature have improved our understanding of sequential innovation, there is limited empirical evidence linking the nature of sequential innovation efforts with market performance. Employing a search perspective with a focus on emerging dynamics of sequential innovation including fast-paced, intense, and transparent competition; our study provides new empirical evidence of the consequences of sequential innovation efforts in terms of market success.

2.4 Hypotheses development

How firms obtain the requisite knowledge to innovate has been an important line of inquiry in innovation research (Eisenhardt and Schoonhoven 1990, Katila 2002, Rao and Drazin 2002) as knowledge forms a basis for differential firm performance (Grant 1996). Firms engage in various types of search activities, such as the search for better manufacturing methods (Bohn and Jaikumar 1992), the search for new products (Katila and Ahuja 2002) and the search for ways to incorporate new innovations (von Hippel and Tyre 1995). In addition, contingent on the firm's environment, one type of search may benefit more than another type (Sidhu et al. 2007).

Sequential innovation is an ongoing a process in which firms engage in sequential experimentation (Davis and Yin 2014, Thomke 2003) by introducing multiple versions of the same product over time (Brown and Eisenhardt 1995). In between consecutive versions, firms search to improve their products' features, learn from the market and incorporate new ideas from competitors. In this manner, firms can keep up with market trends and prevent obsolescence. Such versioning can help preserve the product's core features but at the same time generate new value for the consumers (Kim and Kogut 1996).

We first focus on how different the search is relative to search paths explored by the firm before. In this sense, our perspective is related to the characterization of 'local search' where a firm's R&D activity is closely related to previous activities (Nelson and Winter 2002), but with different (i.e., inverse) operationalization. In our work, we conceptualize *Search Dissimilarity* of a focal firm as the dissimilarity of the search activities (e.g., features/attributes) with respect to its own past search activities. In other words, in sequential innovation for mobile app development, we are interested in characterizing how far the new version is from the neighborhood of the existing versions (i.e., what the firm has already done). While minor improvements to existing features and functionality could be considered as search along previously explored paths, addition of new and novel attributes and functionality could be viewed as 'dissimilar search' from the previous search paths.

First, in searching for new features and attributes for successive versions of

their products, developers can incorporate previously unexplored and distant search paths (i.e., dissimilar search) which may help improve their understanding of the structure of the knowledge landscape (Ahuja and Katila 2004) and potentially make major improvements in the new versions (Katila and Chen 2008). In contrast, considering potential risks and uncertainties associated with market response, developers can resort to variations of existing features and attributes (Katila and Ahuja 2002).

While the reuse of familiar components and refinement of familiar combinations (Fleming 2001) can be effective and it can lead to highly useful innovations as a result of synthesis of well-known technical information (Utterback 1996), there is also a limit to using the same set of knowledge elements (Katila and Ahuja 2002) (i.e. improvising existing features) to enhance innovation performance. Due to the low development cost in mobile apps, new features can be introduced with ease. Given these development factors and that consumers' switching costs are low, we posit that an increase in novelty through previously unexplored elements (new attributes/features), would make the product more appealing and lead to an increase in market performance.

The key premise of our analysis is that the churn of words (i.e., terminology) used in app updates reflects the activities in the underlying software too. The act of adding new functionality would introduce new words whereas improving existing functionality would tend to reuse previously introduced words.

Hypothesis 1 *The greater the search dissimilarity, the higher the app's market performance.*

New and novel features in a product are a major driver of sales. As such, dissimilar search (increase in novelty through previously unexplored elements (new attributes/features)) would make a mobile app more appealing for prospective customers. However, marketing literature also highlights consumers' possible resistance and anxiety about new innovative products which leads them to delay or abandon their purchase (Castao et al. 2008). One mechanism through which firms reduce consumer anxiety and resistance for new products is to provide additional informa-

tion and make them familiar with the product (Goering 1985, Heiman and Muller 1996). Through new version release notes, mobile app firms not only communicate to their existing users, but more importantly, they inform prospective customers about the features and attributes of their products and their latest changes.

We propose that before making a purchase, prospective customers consider the expected additional utility they would get from the new features/attributes, but in doing so, they also take into account possible risks and uncertainty associated with those new features. For a new and novel feature introduced, the anxiety and risk aversion of prospective customers would be particularly high for early versions of products. However, such concerns would be mitigated as the product gets more mature and hence more stable and dependable over time. That is, while dissimilar search (i.e. offering new features/attributes) to existing features is in general beneficial for market performance, the positive effect would be greater during the later stages of an apps life cycle and lower during its early stages. As such, we propose that:

Hypothesis 2 *The relationship between search dissimilarity and an app's market performance is positively moderated by the app's maturity.*

Innovation draws on various sources of ideas, as such, firms can improve their innovation success by leveraging a large number of knowledge sources. Search breadth in terms of the information sources utilized also affects innovation outcomes (Rosenkopf and Nerkar 2001, Katila and Ahuja 2002, Laursen and Salter 2006). Successful search combines knowledge known to the firm and knowledge outside the firm (Katila and Ahuja 2002). While firms vary in the degree of use and reuse of their existing knowledge, they also vary in the way they explore new knowledge (Katila and Ahuja 2002). Search with greater breadth can enrich the firm's knowledge pool by adding distinctive new variations (Katila and Ahuja 2002). Recombination of different kinds of knowledge enhances innovation (Fleming and Sorenson 2004) and can produce novel ideas of high economic value (Ahuja and Lampert 2001). Although greater variety as a result of broad search can also bring additional uncertainty or confusion to consumers (Fleming 2001, Martin and Mitchell 1998),

we expect such detrimental effects to be modest for the mobile apps due to the sequential nature of the innovation efforts and transparent nature of product changes. Overall, for a focal app, we expect that incorporating elements and features from other domains will enhance the market performance of the focal app through useful re-combinations.

Recombination of different knowledge types enhances innovation (Fleming and Sorenson 2004) and can produce novel ideas of high economic value (Ahuja and Lampert 2001). Indeed, we argue that an app which makes recombination of a variety of attributes by drawing from other app categories would generate greater value (utility) for the consumer as the knowledge gained via broad search could be an input for successful recombination (Jung and Lee 2016).

Hypothesis 3 *The greater the search breadth, the higher the app's market performance.*

An important source of firms' knowledge search is the product market (Köhler et al. 2012). Market-driven knowledge search, such as the knowledge of competitor activity, is embodied in the products and services available in the market. In the mobile apps market, unlike in other traditional industries, IP regimes are weak, making barriers to imitation low. Moreover, customers have low switching costs due to ease of substitution of apps. The ease of inspecting competitor apps makes it easy for firms to identify features that are worth imitating or incorporating into their existing features. This makes the setting conducive for firms to be responsive and reactive to market changes. Furthermore, the value of imitation increases in dynamic 'high velocity' environments (Brown and Eisenhardt 1995) where industry structure is unclear, there are no dominant models of success, and there is a constant flux of firms and products. Indeed, such environmental uncertainty has been suggested to be conducive for the efficacy of imitation (Ross and Sharapov 2015). Considering these, we expect firms that actively pursue market driven search and incorporate competitors' recent attributes/features into their own products to exhibit greater performance.

Indeed, in our particular category of productivity apps where consumers' pur-

chase decisions are typically need based, after identifying the need and quickly forming a consideration set of a small number of competing products (Armstrong et al. 2014), prospective customers would then undertake a more detailed analysis of these apps (Roberts and Lattin 1991) based on their descriptions to get a sense of the state-of-art product offering. As this plays an important role in users decision making, it is important for an app to stay up-to-date and be responsive to market competition by tracking and incorporating the latest features and functionality introduced by competitors.

Hypothesis 4 *The greater the market-driven search, the higher the app's market performance.*

While market search may be helpful in driving performance, another factor critical for performance is related to the timing of such updates. On one hand, by delaying an update, the firm can learn more about innovations introduced by competitors and observe market reaction. In their study of the branded drug industry, Ethiraj and Zhu (2008) found a greater imitation time lag to be associated with quality improvement of imitator products. The time lag is essentially a window to learn more about the innovation and to ensure high quality implementation. In dynamic environments where product features constantly evolve and market response may be uncertain, it may seem sensible to delay market search until uncertainty is resolved.

On the other hand, in learning from competitors, speed also matters in that competitor knowledge may lose its value quite fast. A slow incorporation of competitor knowledge was found to deteriorate innovation performance in the R&D activities of robotics firms and optical disc firms (Katila 2002, Rosenkopf and Nerkar 2001). While this was found to be the case in the development of new products, it is not obvious whether this finding will hold in a high-velocity sequential innovation setting.

In other words, whether a rapid update would amplify or dampen any beneficial effects of market-driven search on performance is unclear. By rapidly updating the product, a firm may avoid the risk of losing market opportunity and lagging behind competitors. In contrast, a rapid update may result in a rushed inadequate product

which may limit the value generated by market-driven search. We therefore present two competing hypotheses:

Hypothesis 5a *The relationship between market-driven search and the app's market performance is negatively moderated by a rapid update.*

Hypothesis 5b *The relationship between market-driven search and the app's market performance is positively moderated by a rapid update.*

2.5 Data and Empirical Approach

Our sample consists of data from 588 mobile applications in the 'productivity' category that appear in the top 500 Paid apps list for the iPhone in the US Apple App Store. The observation period spans across 23 months, from January, 2013 to December, 2014. There are two types of data available for apps- 1) static data that typically does not change over time such as the app name, category, app publisher name, and 2) dynamic data that changes over time such as the app rank, version description, price, ratings, etc. Since Apple only displays data for the most recent update of apps, we obtain historical data on daily ranks over our observation period from a mobile analytics firm. Download rank data is available for ranks 1-1500. We exclude from our sample those apps for which download rank data was not available beyond rank 1500. Apart from the paid model, freemium is also a popular monetization strategy that relies on an ad-supported model. For consistency, we chose to focus only on *paid* apps rather than *free* apps because additional or premium functionality in free apps is charged via in-app purchases and free apps heavily rely on advertising for revenue generation. These additional features are otherwise typically accessible for paid apps.

Productivity apps are a suite of apps aimed at enhancing individuals' productivity in life and work. There are various sub-categories within the productivity category, such as note-taking, calendar, organizer and password management, pomodoro apps and so on. Examples of some popular productivity apps include Scanner Pro and WunderList (Heath 2015). A WordCloud depicting the most frequently used words across all the apps in our sample is given in Appendix 2.7.4. The reason

for our choice of the productivity category is the utilitarian nature of the apps in this category (in contrast to the hedonic nature of apps, for example, in the gaming category). We would expect that within the productivity category, features would be utilitarian and a common vocabulary would be shared across firms. Indeed, we find that features across apps tend to use similar terminology, such as "Dropbox support", "organizer", "password manager" and so on. Categories such as gaming, on the other hand, tend to have more creative content and developers tend to use self-generated creative names for features, which would have made it difficult to find a common feature vocabulary for our study.

2.5.1 Main variables of interest

We provide a brief description of our variables in Table 4.1.

Dependent Variable: Our dependent variable is the sales performance of the apps as measured by *Downloads*, the log of estimated weekly download volume for the mobile apps. Although download volume data is not publicly available in the App store, rank information based on product downloads (sales) is released publicly. Apple releases sales rank information for the top few hundred apps in each category. Following the work of Goolsbee and Chevalier (2003) and Brynjolfsson et al. (2003), we assume that the relationship between sales rank and sales volume follows a Pareto distribution. This relationship has been shown to hold in various settings such as books, software, and electronics where sales volume information is not reported publicly, but rank information is available. Furthermore, Garg and Telang (2013) provide a methodology to calibrate the relationship between daily app download volume and app download rank. Based on their approach, we calibrate the relationship between the sales rank and download volume in our study using the actual sales data that we obtained from a select group of developers. By contacting developers and using their actual rank and download data, we were able to calibrate the scale and shape parameters of the distribution [3]. Based on the estimates of the parameters of the Pareto distribution, we infer sales (i.e., download)

[3]Consistent with the findings of Garg and Telang (2013), we estimate the shape parameter to be 1.04

quantities of the apps over time [4]. This strategy is very similar to the one employed by Ghose and Han (2014).

Independent Variables: For examining H1 and H2, we construct the variable *SearchDissimilarity* which refers to the dissimilarity of the firm's search activities with respect to its past activities. For H3, the variable *SearchBreadth* measures the extent to which the firm searches broadly. Lastly, in investigating H4 and H5, we create *MarketSearch* which is the extent to which the firm searches close to its competitors. Our key independent variables are derived using text mining and information retrieval techniques. We further discuss operationalization of each one of our key independent variables in the next section.

Variable	Definition
Downloads	Log of the estimated total number of downloads in the week
Price	Price of the app in the week
AgeofApp	Number of weeks since the app was launched
UpdateSize	Proxy for the extent of the update
NumRetrospective	Number of maintenance related changes in the update
TimeSinceLastVersion	Number of weeks since the last version update was released
SearchDissimilarity	Distance of focal update from past updates of the firm
SearchBreadth	Breadth of the search undertaken by the firm in the update
MarketSearch	Average similarity of content between the focal firm's update and competitor updates
RapidUpdate	Indicator capturing whether the update was released rapidly in comparison to competitors
HasUpdate	Indicator whether there was an update in the app during the week
VersionsTillNow	Proxy for the maturity of an app based on the number of distinct versions since initial release
Rating	App's rating based on customer reviews
TimeTrend	Quarter of the year

Table 2.1: Variable Definitions

Control variables: Sales performance can be affected by several app-specific characteristics. We include the following control variables in our model:

[4] this includes new purchases only

Price: As changes in the price of apps can affect consumers' purchase decision, which in turn would affect downloads, we control for the price of an app in a given week.

AgeofApp: This variable controls for the endogenous time effects of an app's survival in the market, and is measured as the number of weeks elapsed since the app's initial release.

UpdateSize: The overall size[5] of the app update could be expected to have an impact on app performance. We use an indicator *UpdateSize* which takes on the value of 1 if the size of the focal update, operationalized as the total count of words in the update, is greater than the mean size of its previous updates.

NumRetrospective: This variable is a measure for the number of retrospective activities in the app update. This essentially captures the corrective maintenance (bug-fixing) changes, which can affect an app's market performance.

TimeSinceLastVersion: We also control for the time since an app update was last made. For this we use, *TimeSinceLastVersion*, which is a count for the number of weeks since the last app update was released.

Rating: We also control for the app's overall review rating from user reviews from the previous week as this can influence a prospective customer's decision to purchase the app.

RapidUpdate: This is an indicator variable which captures whether the app update is introduced rapidly or not in comparison to its competitors. In order to operationalize this, we first calculate the release frequency (based on TimeSince-LastVersion) of all version updates of the focal app's competitor group. We then create an indicator variable which is equal to 1 if the focal app's update is in the top quartile of its competitor group. Since we are not interested in the specific release time of a version update per se (which is noisy as it is influenced by many external factors), we rather try to identify whether an update was released rapidly or not in comparison to competitors (as we are primarily concerned about the moderating effect of RapidUpdate), we operationalize this as a discrete rather than a continuous

[5]The actual size of updates in bytes is not known, therefore, we use the number of words in the update as a proxy

variable.

HasUpdate: This controls for the effect of an update on the performance in a particular week.

VersionsTillNow: This is a proxy for the degree of maturity of the app operationalized as the number of distinct versions of an app since its initial release.

2.5.2 Operationalization of Independent Variables

It is a standard practice in the software industry to distribute 'release notes' with every new software update to summarize the nature of changes made in that particular version such as enhancements, new features, bug fixes and so on. In the software engineering literature, the content in non-source code documents, such as release notes, has been shown to be a good accurate indicator of activities at the underlying software level. Moreover, release notes reflect the maintenance and evolution activities of the software (Baysal and Malton 2007, Yu 2009). Following this literature, we call the keywords related to software maintenance activities (bug-fixing) as 'maintenance' or 'retrospective' keywords and keywords related to new features and functionality as 'prospective' keywords (Yu 2009). Examples of maintenance keywords are terms such as "bug", "fix", "error", "crash" etc. whereas terms such as "improve", "feature" etc. (see section 2.7.5) are prospective keywords.

In order to quantify the prospective (new features/attributes/enhancements) and retrospective changes (corrective maintenance related) in an app update, we leverage text mining techniques. Text mining is the process of extracting information of value from unstructured or semi-structured text such as descriptions, comments and news articles. It involves collecting data in a paresable format, pre-processing (to remove noise), extracting the items of interest and analytics (i.e., modeling) (Fan et al. 2006). More details can be found in section 2.8.

While there would be heterogeneity in the level of detail used by firms to describe their changes, we assume that for the same firm, the reporting style would remain consistent over time.

2.5.2.1 Operationalization of Search Dissimilarity:

We have historical version update data for each app in our dataset from the time it was initially released (i.e. prior to the observation period). Our analysis is based on the idea that new functionality and attributes (prospective changes) will introduce new terminology which would not have appeared in previous updates. On the contrary, improvements to existing features would re-use terminology that has previously been used.

One of the simplest and intuitive measures of similarity between documents is the Jaccard index, which measures document similarity as the intersection divided by the union of the objects (Huang 2008). The Jaccard distance is given as 1 minus the Jaccard index, given in 2.1, is our measure for *SearchDissimilarity*.

$$\text{SearchDissimilarity} = 1 - \frac{|X \cap Y|}{|X \cup Y|} \tag{2.1}$$

where X is the set of words in the current update and Y is the set of words in the past updates of the app.

2.5.2.2 Identifying sub-categories using topic modeling

In order to operationalize Market-oriented search in the next section, we first need to identify sub-categories of apps (in which apps compete with each other) within the 'Productivity' category. Since no such classification existed at the time of collecting this data, we employ a topic modeling approach to determine sub-categories of productivity apps.

As of 2016, the Apple App Store had 24 app categories, ranging from Games to Utilities and Social Networking (Apple 2016) in which apps are grouped based on content. To facilitate app discovery for users, Apple requires developers to declare a primary category for their app. While some very popular categories, such as Games have sub-categories such as Action, Adventure, Puzzle etc, the Productivity category does not have explicit sub-categories. To understand the effect of market-driven search on app performance, it is important to identify sub-categories with similar competing contents. As a starting point, we generated a WordCloud from

the text in app descriptions in Section 2.7.4 to visually assess whether any themes emerge. It is evident from the WordCloud that certain keywords are frequently used, such as "password", "notes", "calendar" etc. We subsequently conduct more formal analysis by leveraging Latent Dirichlet Allocation (LDA) (Blei et al. 2003) or topic modeling, which is a statistical modeling technique premised on the idea that documents are mixtures of topics, where a topic is a probability distribution over words. In LDA, the choice of the number of topics can affect the quality of topics that emerge. Choosing very few topics a-priori yields very general topics whereas choosing too many topics may result in uninterpretable topics. Using previously suggested sub-category figures by practitioners as a guide (Kimura 2014), we use K = 15, which gives the most appropriate set of topics [6]. We have used a majority rule to assign a topic (category) to an app.

Each topic is labeled with representative words such as password/security management, calendar, goal/habit tracking, note-taking, file transfer, timer apps. A list of keywords associated with the topics and the full set of topics is given in Appendix 2.7.6. Our identified topics are qualitatively similar to those listed by Sensor Tower. (Kimura 2014), a mobile analytics company.

2.5.2.3 Operationalization of Search Breadth:

In order to measure search breadth of an app update in Hypothesis 3, we assess the extent to which there was variety in an app update in the form of discriminating keywords from other domains. For this, we use an entropy based measure of diversity (Harrison and Klein 2007) given in Equation 4.1 to calculate *SearchBreadth*.

$$\text{SearchBreadth} = -\sum_{i=1}^{K} p_i \times \ln(p_i) \tag{2.2}$$

where p_i is the weight of topic i. Updates in which topics are spread evenly across a greater number of clusters will have a higher value of Search Breadth.

[6]For robustness, we do sub-sample analysis on apps having a single topic with 50 percent or more assigned weight, see Table 2.9

2.5.2.4 Operationalization of Market-oriented Search:

In order to capture similarity between a focal firm's update and those of its competitors, we use a keyword similarity approach, such as in Criscuolo et al. (2007), Haas and Criscuolo (2015). Since apps in the Productivity category are utilitarian in nature, a common vocabulary to describe features like "Dropbox support", "Retina display" etc. would be prevalent. Our similarity measure captures how similar the keywords in the focal firm's update are to the keywords used by its competitors in the period between the focal firm's current update and its last update.

For each competitor category, we first construct an App-by-Keyword asymmetric matrix X where element $ij = 1$ if App i has keyword j. From the content of all the apps in our sample, we extracted keywords excluding stopwords. We construct a keyword-by-keyword matrix, K, the ijth cell of which is a measure of similarity between keywords i and j. We use the Salton cosine metric in Equation 2.3 to measure the similarity between keywords. The numerator denotes the co-occurrence of each pair of keywords in the focal firm's update and the update of the competitor, and the denominator is the product of the square root of the respective frequencies of the two keywords across all the updates. In essence, pairs of keywords that co-occur frequently would have similarity closer to 1, while keywords that rarely co-occur would have a value close to 0.

The ijth element of the App-by-App matrix is the similarity between keywords in App i's change and in App j's change. Each ijth element is divided by the product of the total number of keywords in App i's change and in App j's change. From this, we derive the average similarity between App i's current update and updates made by all other competitors since App i's last update, as given by Equation 2.4. A focal firm's update that contains keyword combinations that are similar, on average, to those keyword combinations in competitor changes will have a higher average similarity score. We define this as *MarketSearch*.

For the purpose of illustration, consider Figure 2.5 in Section 2.7.3, a timeline of competitor activities between focal Firm 1's update, v3, and its latest update, v4. Firms 2, 3, 6, 4, 5, 9 and 7 release updates over time during this period.

s_{1i} is the similarity between Firm 1's update content and Firm i's update, where $i\varepsilon(2,3,4,5,6,7,9)$. We are interested in obtaining the average similarity, or *MarketSearch* given by Equation 2.4. A higher value of *MarketSearch* implies that v4's content (keywords) is closer, on average, to the content of competitor updates.

$$\text{cosine}(\text{i},\text{j}) = \frac{\text{coocurrence}(\text{i},\text{j})}{\sqrt{\text{occurrence}(\text{i}) \times \text{occurrence}(\text{j})}} \tag{2.3}$$

$$\text{MarketSearch}_\text{i} = \frac{\sum_{\text{j}=1}^{\text{J}} \frac{\text{q}_{\text{ij}}}{\text{kw}_\text{i} \times \text{kw}_\text{j}}}{\text{J}} \tag{2.4}$$

2.5.3 Estimation Strategy

Correlations between variables are reported in Table 2.2 and descriptive statistics are reported in Table 2.3. We log-transform our dependent variable, the weekly downloads, since this variable is highly skewed (Afifi and Clark 1999).

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(Downloads)	1									
Price	-0.1067	1								
AgeofApp	0.0484	0.1572	1							
TimeSinceLastVersion	-0.1075	-0.0008	0.2287	1						
NumRetrospective	0.0451	0.0148	-0.0498	-0.0582	1					
MarketSearch	0.0404	-0.0015	-0.0451	-0.0037	0.4394	1				
SearchDissimilarity	0.0209	-0.0126	-0.1569	-0.0317	0.2718	0.4246	1			
SearchBreadth	0.0394	0.0117	-0.0236	-0.0396	0.2863	0.3817	0.1884	1		
VersionsTillNow	0.1387	0.1865	0.6182	-0.164	0.0525	0.0344	-0.117	0.0569	1	
Rating	0.0983	0.0032	0.0387	-0.093	0.0085	0.0175	-0.0024	0.018	0.109	1

Table 2.2: Correlations matrix, N = 49114, unit of observation is appid-week

Our data consists of an unbalanced[7] panel of weekly data of firms' market performance (i.e., downloads) and covariates outlined earlier. Because the performance of an App in one week would be highly correlated with its performance in the previous week, we estimate a dynamic autoregressive model in the equation 2.5:

$$\text{Y}_{it} = \beta_0 + \beta_1 \text{Y}_{it-1} + \beta_2 \text{X}_{it} + \beta_3 \delta_i + \beta_4 \gamma_{it} + \varepsilon_{it} \tag{2.5}$$

[7]Our panel is unbalanced because app launch times are different, some apps were launched before the start of our observation period, and some were released during our observation period

Variable	Mean	Std. Dev	Min	Max
Log(downloads)	7.09	1.29	3.96	12.8
Price	4.26	4.9	0.99	99.99
AgeofApp	114	68	0	331
TimeSinceLastVersion	16	19	0	213
NumRetrospective	0.08	0.55	0	21
SearchDissimilarity	.012	.084	0	1
MarketSearch	0.017	0.076	0	0.76
SearchBreadth	.50	3.95	0	254
RapidUpdate	0.28	0.45	0	1
VersionsTillNow	13.8	9.27	1	57
Rating	3.63	1.5	1	5
UpdateSize	0.08	0.28	0	1
HasUpdate	.06	.25	0	1

Table 2.3: Descriptive Statistics for all App firms, N = 49,114, unit of observation is appid-week

$$\Delta Y_{it} = \beta_1 \Delta Y_{it-1} + \beta_2 \Delta X_{it} + \beta_4 \Delta \gamma_{it} + \Delta \varepsilon_{it} \tag{2.6}$$

where Y_{it} is the log of aggregated weekly downloads of App i in week t and Y_{it-1} is the log of aggregated weekly downloads of the same App in the previous week. X_{it} is a set of our covariates, δ_i is the firm fixed effect, γ_{it} is the set of controls and ε_{it} is the error term. Included in X_{it} are our variables of interest, *SearchDissimilarity*, *SearchBreadth* and *MarketSearch*. Any seasonal variation is also controlled with quarterly time dummies. We also add controls for the age of the app, time since the last update, size of the update and the number of retrospective changes. We also include a dummy, *HasUpdate*, that captures whether or not there is an update in a particular week to discern effects of updating on app performance.

In an autoregressive model like ours, inclusion of the lagged dependent variable may give rise to biased estimates and lead to endogeneity concerns. Standard OLS estimates of the lagged dependent variables in a dynamic panel are biased and inconsistent due to the correlation between the individual fixed effect and the lagged dependent variable (Nickell 1981). This bias is not caused by an autocorrelated error process ε, but arises even if the error process is iid. Furthermore, the lagged dependent variable y_{it-1} is mechanically correlated with the ε_i for time period $s < t$,

such that the standard fixed effects estimator is also biased (Wooldridge 2001). Anderson and Hsiao (1981) were the first to demonstrate that the issue with the within estimator can be solved by differencing to eliminate unobserved effects. As differencing shown in equation 2.6 induces a correlation with the error term and the lagged dependent variable, instrumental variables in the form of differences of the lagged dependent variable, such as ΔY_{it-2}, can be constructed in order to obtain unbiased estimates. Dynamic panel estimators based on the lagged instruments have been widely used in recent economics and management research to address similar endogeneity concerns, see for example Haselhuhn et al. (2012), Senot et al. (2015), Chung (2015), Suarez et al. (2013), Acemoglu et al. (2008), Bhargava and Mishra (2014), Narayan and Kadiyali (2015). As such, we opt for the Anderson-Hsiao estimator in a 2SLS procedure which is based on the first-differencing approach to instrumental variables in a GMM framework. This is due to certain characteristics of our sample: 1) the need to control for fixed app-level effects, 2) a linear functional relationship between the outcome variable and regressors, 3) a dynamic outcome variable (current performance depends on its own past realizations), and 4) challenge of finding exogenous shocks or purely external instruments for our regressors.

2.6 Results

Table 2.4 summarizes the results for our hypotheses. Each of the regressions includes an indicator variable, *HasUpdate*, which is set to 1 for weeks when there was an update. Although the use of a dynamic panel estimator is more appropriate for estimation due to the inclusion of lags in our model, we also replicate the analysis with simpler estimators such as OLS and Fixed effects.

We expected corrective maintenance activities to impact performance, for which we added a control *NumRetrospective*. Furthermore, an increase in price is expected to have a negative impact on downloads which is confirmed by the negative and significant coefficient of *Price*.

In H1, we proposed that the greater the *Search Dissimilarity* with respect to the

Table 2.4: Results of estimation

	Dependent Variable = Performance in log(downloads)								
Variables	OLS (1)	FE (2)	A-H (3)	OLS (4)	FE (5)	A-H (6)	OLS (7)	FE (8)	A-H (9)
$\log(\text{Downloads})_{t-1}$	0.5012***	0.3945***	0.4723*	0.5844***	0.3945***	0.4794*	0.5845***	0.3946***	0.4631*
	(0.0088)	(0.0111)	(0.2458)	(0.0072)	(0.0111)	(0.2466)	(0.0072)	(0.0111)	(0.2456)
Price	-0.0035***	-0.0401***	-0.1828***	-0.0036***	-0.0401***	-0.1828***	-0.0037***	-0.0400***	-0.1824***
	(0.0008)	(0.0081)	(0.0241)	(0.0009)	(0.0081)	(0.0241)	(0.0008)	(0.0081)	(0.0240)
AgeOfApp	0.0001*	0.0009	0.0000	0.0001	0.0009	0.0000	0.0001	0.0009	0.0000
	(0.0001)	(0.0007)	(0.0000)	(0.0001)	(0.0007)	(0.0000)	(0.0001)	(0.0007)	(0.0000)
TimeSinceLastVersion	-0.0001	-0.0005**	-0.0017**	-0.0002	-0.0005**	-0.0016**	-0.0002	-0.0005**	-0.0021**
	(0.0001)	(0.0003)	(0.0008)	(0.0001)	(0.0003)	(0.0008)	(0.0001)	(0.0003)	(0.0008)
SearchBreadth	-0.0002	-0.0000	0.0004	-0.0006	-0.0007	-0.0002	0.0002	0.0000	0.0004
	(0.0011)	(0.0008)	(0.0005)	(0.0011)	(0.0008)	(0.0006)	(0.0011)	(0.0008)	(0.0005)
HasUpdate	0.0264	0.0105	-0.0050	0.0211	-0.0058	-0.0189	0.0500*	0.0183	0.0005
	(0.0267)	(0.0259)	(0.0242)	(0.0284)	(0.0277)	(0.0253)	(0.0271)	(0.0263)	(0.0243)
SearchDissimilarity	0.1237*	0.1612***	0.0800	-0.0889	-0.0004	-0.0745	0.0857	0.1568**	0.0703
	(0.0658)	(0.0623)	(0.0665)	(0.0821)	(0.0798)	(0.0822)	(0.0593)	(0.0620)	(0.0662)
MarketSearch	0.1489**	0.1380**	0.1559**	0.1186*	0.1176*	0.1358**	0.1953***	0.1791***	0.1911***
	(0.0680)	(0.0634)	(0.0649)	(0.0644)	(0.0618)	(0.0641)	(0.0694)	(0.0657)	(0.0677)
RapidUpdate	0.0092	0.0135**	-0.0023	0.0085	0.0146**	-0.0014	0.0129**	0.0177***	0.0041
	(0.0058)	(0.0067)	(0.0134)	(0.0056)	(0.0067)	(0.0134)	(0.0056)	(0.0068)	(0.0137)
NumRetrospective	-0.0060	-0.0046	-0.0026	-0.0057	-0.0044	-0.0026	-0.0032	-0.0023	-0.0017
	(0.0050)	(0.0050)	(0.0042)	(0.0052)	(0.0050)	(0.0043)	(0.0052)	(0.0049)	(0.0042)
UpdateSize	-0.0352**	-0.0255	-0.0505***	-0.0428**	-0.0268*	-0.0521***	-0.0434**	-0.0268*	-0.0458**
	(0.0165)	(0.0162)	(0.0184)	(0.0169)	(0.0163)	(0.0185)	(0.0169)	(0.0162)	(0.0182)
VersionsTillNow	0.0012***	0.0074***	0.0385***	0.0014***	0.0075***	0.0403***	0.0015***	0.0073***	0.0332**
	(0.0004)	(0.0019)	(0.0145)	(0.0005)	(0.0019)	(0.0146)	(0.0005)	(0.0019)	(0.0147)
Rating	0.0102***	0.0144***	0.0023	0.0103***	0.0145***	0.0025	0.0102***	0.0144***	0.0023
	(0.0014)	(0.0022)	(0.0033)	(0.0015)	(0.0022)	(0.0033)	(0.0015)	(0.0022)	(0.0033)
SearchDissimilarity X VersionsTillNow				0.0547***	0.0478**	0.0424***			
				(0.0210)	(0.0199)	(0.0139)			
MarketSearch X RapidUpdate							-0.3487***	-0.2811***	-0.1871**
							(0.0954)	(0.0924)	(0.0839)
_cons	0.3224***	2.0888***	-0.0296***	0.4291***	2.0861***	-0.0300***	0.4284***	2.0880***	-0.0290***
	(0.0273)	(0.1680)	(0.0067)	(0.0335)	(0.1679)	(0.0067)	(0.0335)	(0.1679)	(0.0068)
N	44005	44005	41716	45372	44005	41716	45372	44005	41716
R^2	0.8417	0.3839	.	0.8324	0.3842	.	0.8324	0.3841	.

[i] Notes: Standard errors are clustered at the app-level. [a]
[ii] ***p < 0.01, **p < 0.05, *p < 0.1

firm's past activities, the higher the performance. We operationalize the dissimilarity of content in terms of the words used in the update description. The key idea is that while any updates to existing features will use terminology that has previously been introduced; more novel words, or words that have not been used previously are associated with new attributes and functionality. To test this, we examined the coefficient of *Search Dissimilarity* across the OLS, Fixed Effects and Anderson Hsiao specifications in Table 2.4. We find partial support for H1 with the positive and significant coefficient across Columns 1-2 in the OLS and Fixed effects specification. In terms of the magnitude of the effect, a one unit increase in search dissimilarity results in a 16% increase in estimated downloads (β * 100).

We find that the app's maturity is positively associated with market performance, as indicated by the positive and significant coefficient on *VersionsTillNow* in Columns 1-9. In H2, we proposed that the app's maturity, operationalized as *VersionsTillNow*, positively moderates the relationship between *Search Dissimilarity* and an app's market performance. We find support for this in Columns 3-5 with the positive and significant coefficient on the interaction term *SearchDissimilarity x VersionsTillNow* in Table 2.4.

In H3, we proposed that greater *SearchBreadth* would be associated with higher performance. We do not, however, find support for this hypothesis in any of the specifications (Columns 1-9). While search breadth (or scope) has been found to affect innovation and market performance positively in the extant product development literature, this finding appears to not hold in our rapid sequential innovation setting. We discuss this result along with our main findings in Section 7.

In H4, we proposed that the higher the *MarketSearch*, in other words, the greater the similarity of a focal firm's changes with respect to competitor changes since its last update, the higher its market performance. Estimates of OLS, Fixed Effects and Anderson-Hsiao are presented in Columns 1-3. We find a positive and significant effect of *MarketSearch* on app's performance. This suggests that closely following and incorporating competitors' recent features and attributes into own products is associated with increased market success. In terms of magnitude, a

one unit increase in *Market Search* results in a roughly 15% increase in estimated downloads (β * 100).

In H5, we proposed that the relationship between *MarketSearch* and performance is moderated by rapid updating. We find that the coefficient on the interaction term *MarketSearch x RapidUpdate* in Columns 7-9 in Table 2.4 is negative and significant. That is, our results indicate that a rapid update dampens the positive relationship between market search and performance.

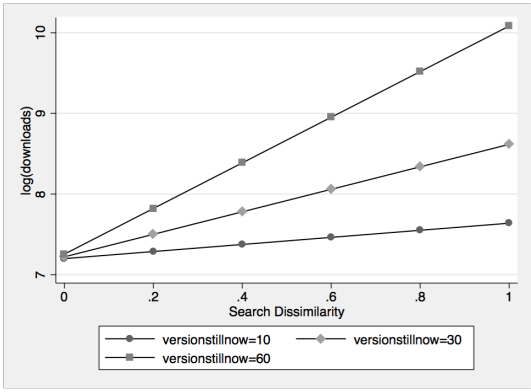


Figure 2.1: Interaction plot between VersionsTillNow and Search Dissimilarity

Figure 2.1 graphs the margins plot for the interaction of *Search Dissimilarity* and *VersionsTillNow*, which shows that the slope gets steeper as the app matures, implying that greater benefits are accrued from dissimilar search as the app matures.

Figure 2.2 graphs the margins plot for the interaction of *MarketSearch* and *RapidUpdate*, and shows that the slope is positive when *RapidUpdate* is 0, whereas the slope is negative when *RapidUpdate* is 1, implying that firms which do not update rapidly derive greater benefits from market search, all other things held constant.

We have controlled for immediate past performance by including the lagged dependent variable in our models. For performance in the subsequent period, we only look at the effects of the main variables *SearchDistance* and *MarketSearch* that are not already explained by the lagged performance, and find these to be significant. The estimates of the lagged dependent variable from OLS and the Fixed-effects estimator can be used as bounds for the true parameters to assess the validity of the

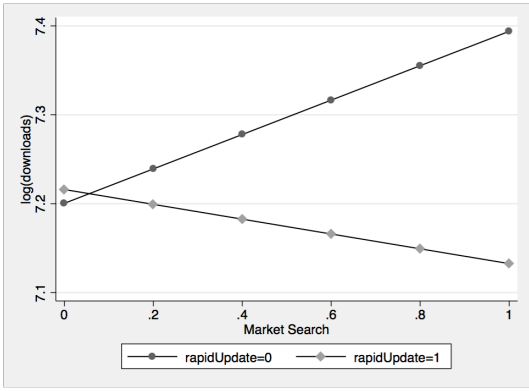


Figure 2.2: Interaction plot between RapidUpdate and MarketSearch

Anderson Hsiao estimator. The OLS estimate is expected to be positively correlated with the error term, resulting in an upward bias. On the other hand, the lagged dependent variable should be negatively related to the error term in a Fixed-effects specification, biasing the coefficient downward (Roodman 2009, Bond 2002). In other words, if the Anderson Hsiao model is valid, the coefficient on the lagged dependent variable from FE should be a lower bound and the coefficient of the OLS regression should be an upper bound. The Anderson-Hsiao estimate for the coefficient of the lagged dependent variable across the different specifications is indeed within the credible range. For example, the OLS and FE estimates for the lagged dependent variable are in the range of 0.50 and 0.39. This supports the validity of our estimation strategy.

2.6.1 Alternative Explanations and Robustness Checks

In this section, we consider and rule out potential alternative explanations to our findings and conduct additional robustness checks. First, the included fixed effects in our model allow us to control for all observed and unobserved time invariant heterogeneity across apps (e.g., inherent appeal or quality of the app, management team skill, etc). That is, we are only exploiting within app variation over time. However, one may argue that an unobserved time varying factor may simultaneously affect both a firm's market oriented search strategy (or search dissimilarity) and its market performance so that $Cov(x_{it}, u_{it} \neq 0)$. This factor could be due to significant changes in management or organizational structure, or new hires which may

improve the market oriented search (or search dissimilarity) of the firm and at the same time independently improve the market performance of the app. We anticipate such an unobserved change would unlikely bias our results as our dynamic panel is weekly and each week is in effect a control for the subsequent week. That is, such unobserved factors would be already accounted for via the lagged dependent variable which is included in our dynamic specification.

It may also be possible that the effect of market-oriented search on performance is amplified at certain times such as in the period after a new IOS release. In response to an IOS release, firms upgrade their apps to become 'IOS compatible'. Therefore, around IOS releases, the activities across firms would appear similar and this can be misconstrued as market oriented search, when actually firms are responding to the IOS event. We control for such an effect by using a dummy variable *IOSRelease* which takes on a value of 1 for a period of 4 weeks (1 month) after an IOS update. After explicitly incorporating IOS update into our Anderson-Hsiao model, our results remained the same. We present the results in Table 2.5, Column 1.

Another potential alternative explanation for the effect of market-oriented search could be that one firm may be an innovator in its sub-category and introduces new features, while others in the same category are imitators. In this case, the innovating firm's subsequent incremental updates might appear to be similar to other apps' features due to overlapping terminology (as our approach hinges on text similarity). As such, the innovator firm's own incremental changes might be incorrectly perceived as its market-oriented search (i.e., imitation of other firms). In order to address this, in the models where we estimate the effect of market-oriented search, we control for *SearchDissimilarity*.

One could also argue that over time, an apps developers may improve app quality as they get better at understanding their market or implementing functionality in better ways, reducing defects, or get better at marketing the product as the product matures. This could then improve market performance in a way not accounted for in our analysis. To investigate this, we regress the count of defects (*NumRetro-*

Table 2.5: Robustness Tests

	Dependent Variable = Performance in log(downloads)			
Variables	AH	AH	AH	AH
log(downloads)$_{t-1}$	0.4751*	0.4780*	0.4719*	0.4573*
	(0.2474)	(0.2458)	(0.2454)	(0.2414)
NumRetrospective	-0.0027	-0.0049	-0.0025	-0.0022
	(0.0042)	(0.0044)	(0.0042)	(0.0042)
SearchBreadth	0.0004	-0.0003		0.0005
	(0.0005)	(0.0006)		(0.0005)
HasUpdate	-0.0050	-0.0062	-0.0033	0.0027
	(0.0243)	(0.0244)	(0.0240)	(0.0247)
Price	-0.1829***	-0.1826***	-0.1828***	-0.1823***
	(0.0242)	(0.0241)	(0.0241)	(0.0240)
TimeSinceLastVersion	-0.0017**	-0.0017**	-0.0017**	
	(0.0008)	(0.0008)	(0.0008)	
MarketSearch	0.1575**	0.1277**	0.1560**	0.1336**
	(0.0649)	(0.0637)	(0.0649)	(0.0658)
RapidUpdate	-0.0023	-0.0021	-0.0023	
	(0.0134)	(0.0134)	(0.0134)	
SearchDistance	0.0795		0.0800	0.0765
	(0.0666)		(0.0665)	(0.0658)
UpdateSize	-0.0504***	-0.0508***	-0.0502***	-0.0424**
	(0.0184)	(0.0186)	(0.0183)	(0.0175)
IOSRelease	-0.0235			
	(0.0204)			
VersionsTillNow	0.0387***	0.0397***	0.0383***	0.0292**
	(0.0146)	(0.0146)	(0.0145)	(0.0137)
Rating	0.0023	0.0022	0.0023	0.0023
	(0.0033)	(0.0033)	(0.0033)	(0.0032)
AgeofApp	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
NewWords		0.0033***		
		(0.0012)		
SearchBreadth_HHI			0.00003	
			(0.0000)	
RelativeTimeToUpdate				-0.0027***
				(0.0009)
MarketSearch X RelativeTimeToUpdate				0.0033*
				(0.0018)
_cons	-0.0324***	-0.0300***	-0.0296***	-0.0283***
	(0.0072)	(0.0067)	(0.0067)	(0.0067)
N	41716	41716	41716	41716

[i] ***p $<$ 0.01, **p $<$ 0.05, *p $<$ 0.1

spective), on the number of versions so far (*VersionsTillNow*) and find a U-shaped relationship between defect count and version count. We interpret this as follows: In the early stages of an apps life, we may expect more defects, as the app stabilizes gradually the defects reduce over time. However, in later stages there is an increase in defect count again. Since defect count increases in later stages, we can rule out improvement in handling defects as a potential explanation for the positive moderation effect we observe. In fact, we observe a positive moderating effect of later versions on the relationship between market performance and search dissimilarity despite the increase in defects in the later versions. Going back to our proposed mechanisms based on consumer anxiety for new and novel features earlier, we contend that defects in this setting are not perceived as product failures as they are in traditional product development settings such as in the automotive industry. Instead, defects are likely to be viewed by customers as glitches that are resolvable in a relatively short span of time and they do not cause irreversible damage. That is, despite the increasing trend in defects, consumers would have greater confidence in the products at later stages and therefore experience lesser purchase related anxiety when they face new and novel features in a product.

It is worth noting that our sample in this study consists of top-performing apps (ranked between 1-1500). Therefore, we would be unable to observe cases where an app in our sample might imitate features of a poorly performing app. While such an imitation may occur in our setting, it would make our current estimates for *MarketSearch* more conservative.

A limitation of our work is that our dataset includes only paid apps. One possibility is that an app may copy or mimic a 'free' app out of the existing sample. First, conceptually this would not change our conclusions for search dissimilarity as this variable focuses on the extent to which an update is dissimilar to an app's own previous features/attributes, and it is not concerned about the source of the new attribute (i.e., how or where it was acquired). Second, with regards to market search, similar to the previous point on potentially imitating a poorly performing app, presence of such imitation would only increase the true effect size of *MarketSearch*.

Another limitation is that developers may incorporate features and ideas from other categories outside Productivity, such as Social Networking, Utilities, Weather, Lifestyle etc. Consequently, future research could examine how this factor affects performance.

Additionally, to ensure the robustness of our results for *SearchDissimilarity*, we consider an alternative measure. One of the best performing sentence novelty measures uses a simple word novelty measure that assigns a score based on the number of words a new document contains that have previously not appeared (Allan et al. 2003). This can be extended to document novelty, given by Equation 2.7 where W_{d_i} is the set of words in document d_i. We call this new variable *NewWords*.

$$\mathrm{NewWords}(\mathrm{d_i}|\mathrm{d_1},....\mathrm{d_{i-1}}) = |\mathrm{W_{d_i}} \cap \overline{\bigcup_{\mathrm{j=1}}^{\mathrm{i-1}} \mathrm{W_{d_i}}}| \tag{2.7}$$

The estimate of *NewWords* in Table 2.5, Column 2 is positive and significant consistent with that of *SearchDissimilarity*, which further supports our second hypothesis. We attribute the difference in the magnitude of the two variables to their different levels of granularity of measurement. While *NewWords* is a raw measure of the number of new words in a new document, *SearchDissimilarity* is based on Jaccard similarity between two sets of words.

We also use an alternative measure for Search Breadth based on the Herfindahl Hirschman Index in Equation 2.8.

$$\mathrm{SearchBreadthHHI} = \sum_{\mathrm{i=1}}^{\mathrm{K}} \mathrm{p_i^2} \tag{2.8}$$

where p_i is the proportion of keywords in cluster i. The estimate of *SearchBreadthHHI* in Table 2.5, Column 3 is insignificant. For robustness, we introduce a continuous measure for speed of updating, *RelativeTimeToUpdate*, which is operationalized as difference between the time since the apps last update, *TimeSinceLastVersion* and the mean time of updating of rival apps in the same category. The positive and significant coefficient on the interaction term MarketSearch x RelativeTimeToUpdate in Column 4 of Table 2.5 shows that the relative timing of the update

compared to rival apps positively moderates the relationship between *MarketSearch* and performance.

Lastly, since sales is imputed as a function of downloads, this leads to measurement error in the dependent variable. This random measurement error is uncorrelated with the independent variables and does not bias the estimates. However, the error results in larger estimated variance. This problem may be mitigated in future studies through the use of larger data sets.

2.7 Discussion and Conclusion

In today's highly dynamic and flexible product environments such as the mobile app development, products evolve with new features and attributes throughout their life cycle in unprecedented speed. During such sequential innovation, firms constantly search for new ideas to generate new value and to make their products appealing to consumers. Our study provides insights into mobile app developers' various search strategies and corresponding market performance implications. We find empirical evidence on the positive effects of (i) search dissimilarity and (ii) market-oriented search on market performance. That is, our results suggest that (i) greater dissimilarity between the contents of the focal firm's update and the content of its previous updates, and ii) greater similarity between the focal firm's update with respect to the recent updates of its competitors are associated with higher number of downloads. Additionally, the functional maturity of an app positively moderates the relationship between search dissimilarity and market performance. Put differently, there are greater benefits accrued to making novel content at later stages, rather than early stages of the app's life cycle. Moreover, we find that rapid update negatively moderates the relationship between market-oriented search and market performance, indicating that a rapid release dampens potential benefits of market-oriented search.

At the same time, our findings suggest that it does not help to be a lone-wolf in this environment. We find that monitoring and keeping up with competitor ideas and incorporating these into your own products can help performance by mitigating obsolescence. This is consistent with and complements the earlier work which

showed that in mobile app platforms with high number of producers, overall innovation occurs through the diversity of the population of producers and not by the heroic efforts of any one innovator (Boudreau 2012b). Also, our results indicate that while incorporating competitors' recent features into their own apps seems to be beneficial for developers, rushing to incorporate such features is not a good idea as it may limit the value generated by market-oriented search.

While we find support for our hypotheses on search dissimilarity and market-oriented search, we do not find support for our hypothesis on search breadth (this is consistent across different operationalizations of search breadth). This may be due to several reasons. Incorporating broad knowledge sources may increase the complexity to manage the variety and relationships between the sources (Leiponen and Helfat 2010). Moreover, it may become difficult to combine technology components as the number of interactions between them grows (Fleming and Sorenson 2004). As a result, firms may avoid searching too broadly. Moreover, broad search may be confusing for customers and risky in a setting with a fluid customer base, as customers may download competitor apps quickly. Broad search may also be more applicable in settings where a firm has more time to search and incorporate knowledge from diverse sources. In rapid sequential innovation, the quick implementation and time pressure for frequent product release may make broad search less conducive for market performance.

In an earlier review article on product development decisions, Krishnan and Ulrich (2001) observed that product development is highly contingent on market uncertainty and other environmental characteristics, and they called for new research to provide insights on customizing product development practices to diverse environments such as small entrepreneurial firms and varied industries. Responding to this call, our study investigated sequential product development and its market implications in the mobile apps industry, to our knowledge the first time in the operations management literature. Unlike traditional product development environments that have been studied (Sommer and Loch 2004, Sommer et al. 2009, Pich et al. 2002) where R&D activities are kept secretive and revealed only upon product

launch, in this dynamic apps setting, the activities of one's immediate competitors are visible through updates. Our access to detailed longitudinal data and use of novel text mining and information retrieval techniques allowed us to monitor and quantify both continuous development (i.e. product evolution) and the corresponding market performance of the mobile apps over the course of their sequential updates. Also, while existing work in product development with the search perspective has primarily relied on patent-based data where patent citations represent technological search, we take a different perspective and focus on search over ideas or features deployed in the app and embodied in the text descriptions of app releases.

The traditional NPD setting is that of a large firm which has multiple products and sufficient time between successive releases. In emerging hyper-competitive settings such as the mobile apps, the typical firm is a resource-constrained entrepreneurial venture with no established innovation processes where knowledge management activities change throughout the venture's life (Gaimon and Bailey 2013). In such settings, it becomes even more critical to identify appropriate search strategies to generate value and to ultimately attract new customers. Our study sheds new light into this by identifying several conducive search strategies (e.g., dissimilar search, market-oriented search) for mobile app developers to achieve greater market success.

In addition, the mobile app development industry has been described with the analogy "let a thousand flowers bloom" (Boudreau 2012b) where large numbers of producers are brought to the platform to stimulate innovation. Despite the significance of competition in mobile app development and in other emerging product development settings, prevalent competition dynamics such as imitation have received limited attention in the existing operations and product development literatures. As such, our study makes an important contribution by taking an operational lens to competition dynamics and documenting a positive association between incorporating competitors' recent attributes into one's own products and market performance.

Moreover, while our conceptualization of search breadth and search dissimilarity is in line with measures of search scope and search depth in Katila and Ahuja

(2002), our results differ. This can be attributed to differences in the nature and complexity of search in R&D intensive industries such as robotics (the empirical setting the work in Katila and Ahuja (2002)), pharmaceuticals, telecom and computer. Search has largely been studied in R&D intensive settings where product development activities are much less transparent to outsiders. Through our novel empirical context, we examine established search constructs in an important yet understudied setting of mobile app development.

Developing an effective imitation strategy is a problem common to firms across industries (Lieberman and Asaba 2006) which gets exacerbated due to high levels of uncertainty. Firms when faced with pressures to imitate upon finding that new feature introduction through upgrades increases sales and customer base. The literature on competitive dynamics and first mover advantage suggest that fast imitators generally do better than firms that are slow to imitate (Giachetti et al. 2017). However, the typical settings in which this is studies are technology intensive firms that face competitive pressures due to technological uncertainty. In the mobile app development settings, uncertainty arises due to other factors such as the hypercompetitive nature of the industry with players in flux and no dominant market players. Moreover, the stakes and investment associated with releasing app updates in response to competitor moves are significantly lower as compared to the intensive research and development and product development cycles for typical technology products.

Our study comes with limitations, which in turn provide new research opportunities. First, while Productivity is one of the most important app categories and this setting is ideal for our text (semantics)-based analysis, generalization of our results to other categories might be limited. Future work could investigate to what extent the results reported in this study would hold in other categories. With advances in image recognition and other similarity techniques, one could explore market implications of feature level similarity in subsequent app releases in other categories such as gaming, health and fitness etc.

In addition, although we included strong control variables, used a dynamic model with a robust estimation approach and ruled out alternative explanations,

one has to be cautious with a strong causal interpretation of our results. In the absence of field experiments (which are difficult to implement in highly competitive and dynamic business settings such as mobile app development) or truly exogenous variation in search strategies due to external factors (which was not available in our data), we addressed our research questions and examined the relationship between various search strategies and market performance by exploiting the longitudinal nature of our data set and within app variation over time. Future work could explore feasibility of aforementioned quasi-experimental strategies to study market implications of sequential development efforts.

As search is a highly useful and previously established perspective to conceptualize product development and innovation efforts in the existing literature, we also employed a search lens in our study of sequential innovation in mobile app development. However, we recognize that 'adding a new feature' may be somewhat different from 'searching for that feature' in certain cases. As such, we are careful not to overclaim theoretical contributions to the search literature itself. We use the search perspective as a tool to help us conceptualize firms' incorporation of new features and attributes into their existing apps. Additionally, as an initial attempt to explore the link between app features and market success, the scope of search in our paper is limited to firms' own previous activities and that of their competitors. We do not investigate particular reasons behind firms' dissimilar search or their market-oriented search efforts. For instance, customers may play a role in influencing various search strategies with their comments and feedback (e.g., "your competitor has added this feature, you should do it, too"). Future research could deepen our understanding by building on our work and exploring implications of such mechanisms.

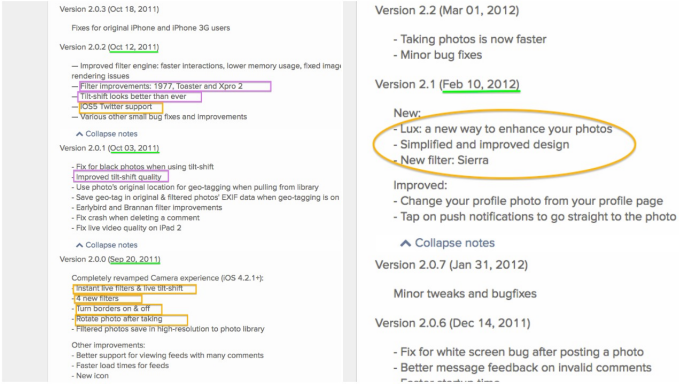


Figure 2.3: Version updates example in Instagram - with periods of feature introduction followed by improvements

Appendix

2.7.1 Example of version updates in Instagram

2.7.2 Version updates example with prospective and maintenance related keywords

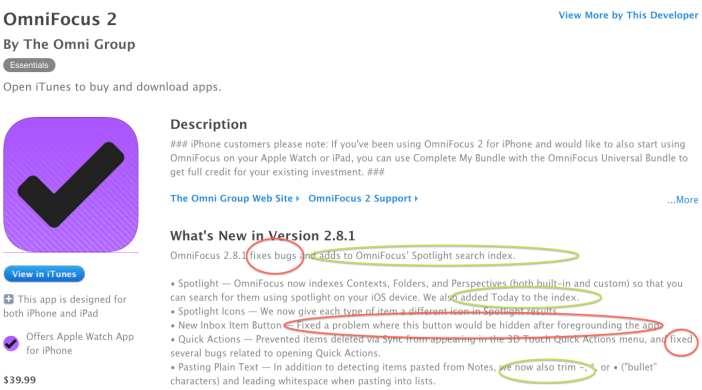


Figure 2.4: Example of version update for the Omni Focus app highlighting prospective and maintenance related keywords

2.7.3 Figure illustrating Market Search

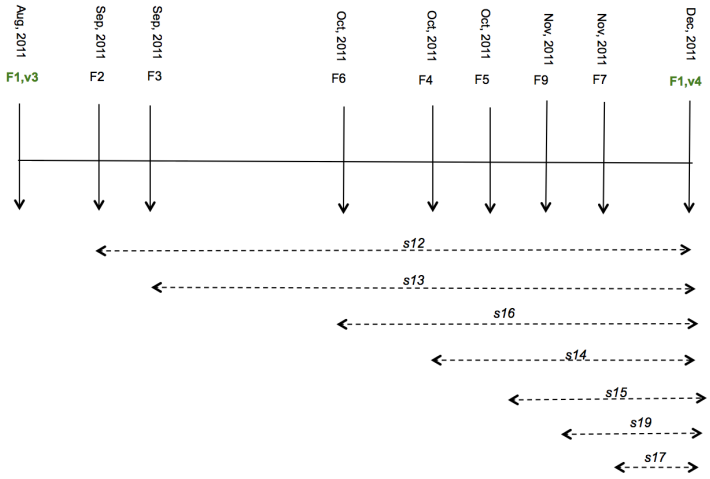


Figure 2.5: Similarity between current change of focal firm with respect to competitor changes

2.7.4 WorldCloud of Productivity app words in sample

2.7.5 Examples of Prospective and Maintenance related keywords

Prospective or Evolutionary Change Keywords	Corrective maintenance Change Keywords
feature, add, improve, optimization	crash, error, memory leak, bug, fix, defect, failure

2.7.6 Topics and associated keywords (unstemmed)

Topic	keywords
Topic 0	file, download, browser, transfer, mac, computer, video, view, open, manage, connect, photo, web, server
Topic 1	language, translate, text,voice, english, spanish, french german, recognition, keyboard, italian, speech, chines, speak
Topic 2	item, list,shop, scan, data, code, pro, barcode, script, store, category, track, collect, inventory
Topic 3	goal, habit, track, daily, time,day, activity, tracker, achieve, progress, start, manage, break, work
Topic 4	data, report, planner, meet, time, display, confer, control, plan, student, chart, view, battery, inform
Topic 5	time, work, custom, create, automatic, quick, design, avail, simple, touch, number, change, note, application
Topic 6	task, list, sync, remind, manage, project, item, gift, organize, todo, date, mac, home, clean
Topic 7	read, record, alarm, clock, sound, mind, article, time program, wake, audio, hypnosis, play, vokul
Topic 8	calculate, unit, tool, conversion, analytical, load, google, size, flight, refer, metric, ticket
Topic 9	contact, message, group, text, send, call, phone, address, number, backup, mail, google, book
Topic 10	password, security, data, manage, generate, encrypt, backup, inform, sync, vehicle, dropbox, protect, record, safe
Topic 11	color, design, draw, photo, image, create, home, palette camera, tool, sketch, shape, pro, room
Topic 12	calendar, event, day, remind, view, schedule, week, month google, shift, holiday, date, sync, time
Topic 13	note, write, text, dropbox, evernote, sync, record, notebook, outline, idea, create, tag, organ, map
Topic 14	document, pdf, page, file, print, text, convert, printer edit, scan, image, dropbox, word, read

2.8

App developers follow the convention of denoting each update with a unique version id, where version ids appear in increasing order. For example, Version 2.0 would be succeeded by Version 3.0, and so on. Another convention used by developers in this setting is to list changes line by line where longer changes are separated by delimiters such as “*” or “-”. Developers may repeat text from a previous major update in a subsequent update, clearly marking what was changed in each version. We filter out this repeated version info from previous updates and only keep the text that is relevant to the most recent release. Regular expressions (regexes) are used to split text by delimiters. In order to improve accuracy,we also use standard natural language processing (NLP) techniques such as stopword removal and stemming to pre-process the data. Stopword removal consists of removing common

Table 2.6: Variance Inflation Factor (with OLS estimation)

Variance Inflation Factor	
Variables	VIF
log(Downloads)$_{t-1}$	6.05
Price	1.06
Timesincelast	1.46
QuarterDummy1	
QuarterDummy2	2.11
QuarterDummy3	2.22
QuarterDummy4	2.31
QuarterDummy5	2.35
QuarterDummy6	2.35
QuarterDummy7	2.34
QuarterDummy8	1.97
HasUpdate	6.15
AgeofApp	2.02
Versionstillnow	1.99
SearchDissimilarity	1.33
SearchBreadth	1.29
RapidUpdate	1.29
MarketSearch	3.28
NumRetrospective	1.41
UpdateSize	3.6
Rating	1.03
_cons	

English words such as articles, prepositions and conjunctions, (such as “the”, “a”, “of”) as these words do not contribute to the semantics of the text and therefore removing these can reduce noise. It is also a common practice in the NLP area to filter domain-specific stopwords that do not have much discriminative power. We appended commonly occurring words in the context of mobile apps, such as “app”, “upgrade”, “customer”, “Apple”, “version” to our stopword list. Stemming reduces a word to its stem. For example, words such as “fix”, “fixing” and “fixed” are reduced to the stem “fix”. In this way, words with the same meaning appearing in different forms (e.g. tense or plural) are reduced to their stem. This significantly improves classification and matching accuracy. We use the classic Porter Stemmer to stem words.

2.9 Robustness- more tables

Results with apps having one major topic (50 percent or more)

Variables	OLS	FE	IV
	(1)	(2)	(3)
$\log(\text{Downloads})_{t-1}$	0.4961**	0.3908**	0.3087
	(0.0125)	(0.0136)	(0.2922)
Price	-0.0030*	-0.0502**	-0.2108**
	(0.0013)	(0.0112)	(0.0493)
AgeOfApp	0.0001	0.0009	0.0000
	(0.0001)	(0.0008)	(0.0000)
TimeSinceLastVersion	-0.0002	-0.0005	-0.0008
	(0.0002)	(0.0004)	(0.0015)
SearchBreadth(keywords)	-0.0135	-0.0082	-0.0054
	(0.0124)	(0.0119)	(0.0157)
HasUpdate	-0.0224	-0.0276	-0.0672
	(0.0357)	(0.0348)	(0.0444)
SearchDissimilarity	0.1758+	0.2045*	0.2020*
	(0.0917)	(0.0874)	(0.0995)
MarketSearch	0.2423**	0.2009*	0.2567*
	(0.0915)	(0.0826)	(0.1034)
RapidUpdate	0.0120	0.0165+	0.0315+
	(0.0076)	(0.0089)	(0.0181)
NumRetrospective	-0.0104	-0.0089	-0.0056
	(0.0071)	(0.0073)	(0.0067)
UpdateSize	0.0067	0.0101	0.0223
	(0.0313)	(0.0293)	(0.0340)
VersionsTillNow	0.0011*	0.0083**	0.0055
	(0.0005)	(0.0028)	(0.0178)
Rating	0.0084**	0.0121**	0.0048**
	(0.0017)	(0.0028)	(0.0012)
_cons	0.2783**	2.2588**	-0.0420**
	(0.0329)	(0.1606)	(0.0112)
Day Dummies	Yes	Yes	Yes
N	25045	25045	24423
R^2	0.8559	0.3972	.

[i] Notes: Standard errors are clustered at the app-level. [8]

[ii] ***p $<$ 0.01, **p $<$ 0.05, *p $<$ 0.1

Chapter 3

Gender-Based Preferences for Tech Work? Field Experimental Evidence from an Internet-of-Things Platform

Crowd-sourcing is generative for innovation through expertise diversity derived from users' knowledge domains, product usage, expertise etc. Managing participation of crowds in these platforms is important. A larger crowd size increases the likelihood of an idea or solution to be generated by the crowd (Boudreau 2012a), in turn increasing the probability of novel, valuable ideas and approaches that the firm may implement (Poetz and Schreier 2012, Jeppesen and Lakhani 2010). Users are an asset in crowd-sourcing and increasing participation from marginal users who bring diverse skills, knowledge and experience can be instrumental in a crowd-sourcing platform's growth and success. Despite similar Internet access, gender disparities exist in participation in online communities. However, female solvers, who have typically been in the "outer circle" in the sciences, outperform men in contributing successful solutions (Jeppesen and Lakhani 2010).

This experiment is layered on a newly created tech-based crowd platform during outreach to members. In this experiment, the goal is to understand gender disparities in participation in crowd-sourcing in general, and more specifically supply-side gender based preferences for tech-based work.

More males than females participate in technology-intensive occupations, to-

day. Across universities in the United States, females make up about one in five of all Computer Science and Engineering enrollees (National Student Clearinghouse Research 2017). The topic of under-representation of females in tech (i.e., applied sciences in Information Technology and Engineering) has attracted prominent popular attention and interest (Turk 2017, Bowles 2017), given the tech industry's central importance in generating employment, growth and innovation (Wolf and Terrell 2016a). Tech industry executives, themselves, routinely describe their industry as an open meritocracy that is starved for talent from any and all sources (Arrignton 2010, Ebertz 2017). Acknowledging disparities between females and males, prominent Silicon Valley companies have initiated diversity programs and taken measures to seek applicants and promote under-represented demographic groups, including females (Huet 2017). Nonetheless, labor market participation by females in technical occupations remains less than one in five (Wolf and Terrell 2016b). In addition, the tech industry today finds itself beset with claims that its work culture and work conditions contribute to these disparities[1] (e.g., Chang (2018)). Gaining a better understanding of sources of disparities between the sexes and how they might be resolved might help in expanding the supply of skilled tech workers–while also better identifying which particular actors or organizations in society might best implement solutions. This could be especially productive in countries such as the U.S. whose domestic technical workforce only fills about one in two of the nation's technical jobs. In this paper, we make progress in understanding possible sources of disparities of the sexes in tech based on the results from a large-scale field experiment focused on gender-based [2] preferences and sorting into tech work.

Disparities among males and females in labor market participation in different fields have been studied by researchers in many areas of social science [3]. This

[1]See (McMahon 2017, Seetharaman 2017, Benner 2017, Mundy 2017)

[2]Gender here is meant to refer not only to biological differences, but also any socialized differences associated with sex

[3]For example, in Sociology, see: (England 2010, Fernandez and Sosa 2005, Ridgeway 2011, Doering and Thébaud 2017, Browne and Misra 2003); in Economics, see: (Blau and Kahn 2017, Bertrand and Mullainathan 2004, Marianne 2011, Buser et al. 2014, Akerlof and Kranton 2002, 2005) in Management, see: (Blau and Kahn 2007, Hakim 1979, Blackburn et al. 2001, Bradley 1989); in Psychology and Social Psychology see (Glick et al. 1988); in Feminist and Social Theory see, for example (Walby 1989, Hartmann 1979, Averett et al. 2017). See (Cohoon and Aspray 2006,

past research collectively points to a long list of potential demand-side explanations (related to tendencies on the hiring side of the labor market) and supply-side explanations (related to tendencies on the worker side of labor market) resulting in the disproportionate representation of the sexes in many fields. Demand-side explanations include such things such as informal or legal and policy-based discrimination (Becker 2010, Aigner and Cain 1977), social network affinities and homophily in hiring practices, the existence (or lack of) role models and peer networks (Findlay et al. 2009, Reskin and Maroto 2011, Reskin and Roos 2009), and the structure of work conditions and work environment (Goldin 2014). Supply-side explanations include constraints created by family and home life, gender roles and expectations (Polachek 1976), gender-based differences in self-confidence and self-efficacy beliefs (Bandura et al. 2003, Marianne 2011), or simply gender-based preferences and tastes for certain types of work or work environments (Hakim 2002, Anker 1998). Of course, each of these factors is hardly independent and may be closely related to these and other mechanisms, to socialized norms and stereotypes (Eagly 1983, Walker and Fennell 1986), or to educational and formative institutions such as schools, and influential individuals such as educators, parents, and role-models (Bleeker and Jacobs 2004). See Azmat and Petrongolo (2014), Shurchkov and Eckel (2017), Blau and Kahn (2017) for recent surveys.

The above ideas and theories on labor market disparities have also begun to be studied in technical occupations and tech industries, with particular emphasis on Computer Science and Engineering (e.g. Beyer et al. (2004), Bandura et al. (2003), Terrell et al. (2017), Kahn and Ginther (2017))[4]. Thus far, however, the bulk of this work is descriptive and conjectural. An exception is the Murciano-Goroff (2018) study of gender-based differences in the behavior of applicants and hiring managers on an online platform, featuring controlled comparisons. This study notes that females on this platform under-represent their skills on their publicly-observed profile relative to independently-collected objective measures of their experience;

Anker 1997) for reviews

[4]Given its distinct institutions and career paths, a separate literature studies the differential representation of males and females in the sciences (e.g., Ceci (2018))

employers then recruit female participants less than men, controlling for observable measures of skills and experience.

Important clues also come from broadest descriptive patterns of the pipeline of technical workers-beginning with earliest education. Girls, on average, have tended to lag behind boys in sciences and mathematics in adolescence, after having performed at least as well as boys in these topics in earlier ages. This gap then continues to widen through high school (Bacharach et al. 2003, Perez-Felkner et al. 2012). Even among high-achieving math students in high school in the U.S., girls more frequently go on to pursue degrees outside of engineering, computing or physical sciences (Lubinski and Benbow 2006, Ceci et al. 2009). Therefore, while more females have pursued a college education in the U.S. than males since the 1980s, gender differences in enrollment across different higher education programs remains a stark first-order fact. As noted earlier, the roughly one-in-five females in the labor market is mostly mirrored by and largely explained by similar ratios of those attending technical programs in college. Thus, if there is a leaky pipeline (Seymour 2002) in the development and supply of female technical workers in the U.S., there is some consensus that an important point of leakage is in the proportion of females electing to pursue science, technology, math and engineering (STEM) programs[5]. Our emphasis here on patterns of development and supply of female technical workers in the United States, in particular, is important and deliberate. Female-to-male ratios in technical fields vary substantially from country to country, with the proportion of female technical workers considerably higher in a number of other countries (Stoet and Geary 2018).

There is no clear evidence that these stark patterns can be explained by innate aptitude differences (Spelke 2005, Hyde et al. 2008, Ceci et al. 2009). However, the differential interests of women and men are cited an important factor contributing to differential career choices (Ceci et al. 2009, Eccles 1994). For example, studies

[5]Among those in technical programs, it is also the case that female graduates in the U.S. disproportionately shift to non-technical careers after graduation (Ashcraft et al. 2016). Attrition of females from technical occupations remains differentially higher compared to other professional occupations (Glass et al. 2013), and has been attributed to factors such as dissatisfaction with pay lack of opportunities for advancement (Hunt 2016)

have shown that women have higher propensity towards people and lower towards things (Woodcock et al. 2013). Interest in itself can be shaped by several factors, such as the person's expectancy for success in the occupation (Eccles 1994), self-confidence and so on. We begin to sort-out demand from supply-side causes by focusing particularly on gender-based (supply-side) preferences for tech. However, there are fundamental challenges in drawing empirically-grounded inferences even on this most basic question. The participation in tech occupations we typically observe reflects two-sided matching between workers and employers, and therefore a combination of both supply-side and demand-side factors at once. Usual gender representation comparisons, even those attempting to account for many explanations with say control variables, will not attend to this problem of distinguishing causes.

The research design here attends to this most basic inferential challenge. We devise a large-scale field experiment design in which participants choose whether or not to participate in a tech learning and work activity- a one-sided supply-side choice by workers. The idea here is to expose comparable males and females to the same opportunity and by observing differential responses under controlled conditions gain insight into differences in willingness to participate.

Therefore, the research design here essentially follows approaches used in prior lab experiments estimating gender-based preferences for different types of organizational conditions such as high-powered incentive systems, intensive levels of competition, risk, social interactions and other economic phenomena (see Croson and Gneezy (2009) for a survey). Here, however, the primary stimulus against which differential gender-based responses are measured relates to the nature of the work, itself. In inferring gender-based preferences and differential sorting, the research design also seeks to exploit representative work and a representative cross-section of workers. Here, the work opportunity relates to a new mainstream area of today's technological innovation: the Internet-of-Things. The experimental population and risk set for participation is 112,770 alumni and students from a large American university, from all fields and career stages.

The research design is also geared to attempting to discern whether preferences for tech work in this instance relate to the nature of the tech work, itself, or instead could reflect expectations about the tech work environment. To do so, we randomly assign subjects to programs that either emphasize competitive or collaborative interactions with other participants.

In a seminal paper by Niederle and Vesterlund (2007), women were found to "shy away from competition when given a choice to compete while men were found to compete too much, conditional on ability, risk attitudes and beliefs about relative performance. If indeed women are less willing to compete in comparison to men, this may have implications for women's career choices, entry and performance in tech work environments.

The finding that women prefer competition less than men appears plausible in light of the observed patterns in labor market outcomes, education, underrepresentation of women in competitive roles (CEOs etc), and has prompted a stream of work in a very short span of time (see (Niederle and Vesterlund 2007, 2011, Croson and Gneezy 2009, Kuhn and Villeval 2014, Lee et al. 2016). The evidence, however, largely comes from relatively low-stakes lab experiments where participants are given a synthetic task such as puzzles or math problems with low-stake incentives. With the exception of the lab study by Kamas and Preston (2012), that found no signicant gender dierences among STEM students in choosing tournaments, most studies confirm gender dierence in taste for competition tend to typically recruit from business schools. While we may generalize from the stylized lab result that women have a distaste for competition, due to the minimal evidence from the field (Niederle and Vesterlund 2011, Garratt et al. 2013), it is unclear whether ndings from the lab would generalize to naturally occurring competitive environments.

While we will report important effects of the organization of the work, we find no evidence that these effects interact with or explain preferences to participate in the tech work, here. Our main findings run counter to usual notions of general gender-based preferences for tech and associated stereotypes. We find that

among individuals in non-technical fields females are significantly less willing than are their male counterparts to participate, and especially so in the competitive treatment. However, the results reverse for individuals in technical fields, wherein females (and especially those in Engineering) are at least as willing as their male counterparts to participate, and even more so in the competitive treatment. Results hold across different ages and cohorts from university students through the senior workers. The population-wide patterns follow those in non-technical fields, simply given the weights of individuals going into technical and non-technical fields (Note here, too, that tastes for competition versus collaboration and tastes for tech are closely related).

As we discuss herein, the differential sorting patterns are inconsistent with simple stereotypical notions of (general) gender-based preferences. Further, the results are consistent with gender-sorting processes and preferences emerging early in life, and prior to university. In as much as large differences exists among those in technical and non-technical fields of training, and are stable across age cohorts within our sampling frame from university through to senior workers, we appear to observe the "sorting on the (already) sorted. Also consistent with this interpretation, the differential gender sorting within the experiment is largely statistically explained by the proportional gender representation within different fields. Therefore, more broadly, the results are consistent with the leaky pipeline largely occurring at early ages. Years of separation and stratification–and possibly the effects of distinct socialization and human capital development lead to enduring differences among those sorting to technical and non-technical fields.

The paper proceeds by introducing the research setting in Section 2. In Section 3, we describe the experimental design. The data are described in Section 4. Results are reported in Section 5. Section 6 concludes with possible intepretations and avenues for future work.

3.1 Experimental Setting

3.1.1 A Representative Area of Tech Work

To draw generalizable conclusions, we exploit a tech opportunity, which to all practical extent, is representative of tech opportunities beyond the experiment. Here, we conform to a variety of notions of tech work. These include the practical application of scientific principles for practical (and often commercial) purposes; the application of specialized skills (as those in sciences, technology, engineering and mathematics); or technological development in the sense of creating novel working systems and products. Modern popular usage of tech also tends to be synonymous with novel applications of modern information technologies.

The opportunity in the experiment that we focus on for the experiment is consistent with each of these ideas and relates to the Internet of Things or IoT. The Internet of Things or IoT is not a single technology but a convergence of pre-existing sensing, networking, and web technologies (Puliafito et al. 2012) that brings numerous business opportunities in the form of products and systems enabled by recofinguring and connecting networks, sensors, software, data science methods, big data, cloud, Internet security, robotics, control systems, etc.

The "things" in Internet of Things refers to Internet enabled physical objects, sensors, devices etc. Since IoT is a synergy of pre-existing technologies, industry observers note that today's technically-trained workers, particularly those in Applied Sciences and Engineering, will grasp important aspects of IoT (e.g., software, networking, control algorithms, etc.). Moreover, there are also considerable opportunities for workers from non-technical areas.

In representing an outgrowth of today's Information Technology and Internet industry, it appears the Internet of Things, at least today, reflects the kinds and sources of gender disparities that exist in today's tech industries. For example, IoT Now, an IoT industry trade magazine carried out a survey of gender disparities, interviewing large numbers of current industry participants and confirmed the existence of lingering disparities, likening the industry to other areas of software,

engineering and information technology in this respect[6].

3.1.2 A Meaningful and Widely Accessible Tech Opportunity

To systematically study whether females and males have different inclinations to participate in learning and project development work in the Internet of Things, we collaborated within a program at Northeastern University's IoT Open Innovation Lab, an organization whose goal is to provide such opportunities for its alumni and students. The program's goal is to create on-ramps to skills development and realization of new projects through a series of learning-by-doing activities, taking the form of a series of discrete problem-solving events in which participants work through early-stage ideation and conceptualization of projects through to design and development. The key features of this program from an experimental design perspective are to (1) instantiate an IoT learning and development opportunity, (2) reflect a one-sided choice, and (3) to reflect a realistic and meaningful choice across a wide population of prospective participants.

The program is designed to accommodate a wide range of types and skills of potential participants, an approach within Learning Science (e.g. Boaler (2015b,a)) referred to as a principle of low-floors and high ceilings to remove learning barriers but also to maintain suitable challenges. This is accomplished in several ways. First, the focus on developing IoT applications, or working systems, will itself lead to high ceilings. Although developing and reconfiguring existing technologies to implement working applications does not require fundamental advances, this nonetheless requires a span and depth of knowledge that no one person, no matter how expert, will typically have mastered. For, example, an entire development will require architectural knowledge of IoT system subsystems, knowledge of component subsystems themselves, knowledge of the application or use case and customer needs and how component technologies can be architected to address the use case, and an understanding of the economics and modes of commercialization of the system.

Low floors, or the accessibility and feasibility of making a meaningful contri-

[6]See October, 2018 Women of IoT say push diversity, dont patronise and dont discriminate downloaded at https://www.iot-now.com/2018/10/09/89071-women-iot-say-push-diversity-dont-patronise-dont-discriminate/

bution for people at varying levels and types of expertise, whatever one's field and background, is achieved through a variety of approaches. For example, the process of conceiving ideas, proposing designs, selecting most promising projects, and development and prototyping is partitioned into discrete activities and events to make any one work activity a manageable size. Breaking down the process also allows each activity to be designed with a series of sub-steps to guide learning-by-doing. (For example, all participants are able to design high-level system architectural designs in a step-by-step graphical user interface drag-and-drop software-guided instructional process). They do so in competition with other such teams, competing for rank, certificates and cash rewards.

The IoT Open Innovation Lab's program is made more accessible and scaleable by running all activities over an online platform, whether individual participants choose to collaborate in-person or virtually. While the program does pose an opportunity cost, roughly a dozen hours over several weeks for any one event, there is considerable flexibility in the times one chooses to participate and to collaborate with one's team. Individuals are not required to participate in all events or challenges, but can participate in individual activities a la carte-depending on their tastes, availability, and in certain cases according to skills required. The program has, to date, been available without any charge to participants. Therefore, despite the technical challenge and time requirements of IoT project development and opportunity cost, the program offered is a real and legitimate opportunity for those in all stages of career and career path. Participants have included experienced and award-winning entrepreneurs, retirees, business people, educators and teachers, industry experts in areas from healthcare to environment, engineers and computer science graduates, unemployed or under-employed workers, Ph.D. researchers, and students from all fields of study and countless industries. To date there are roughly 5,500 participants, from 50 U.S. states, most countries in the Americas, Europe, Australia, Asia, and seven African nations.

3.2 Research Design

This research follows analogous past lab experimental research that seeks to estimate gender-based preferences (Croson and Gneezy 2009), where comparable males and females are exposed to the same stimulus and differential responses among comparable males and females are then recorded. Here, the crux of the experiment is to compare the propensity of females and males (from the same field of study and career stage) to participate in the representative tech opportunity, described above. This section reviews key details of the research design.

3.2.1 Overview: Embedding an Experiment within the IoT Platform Launch

The research presented here was based on an experiment embedded within the initial growth campaign of the IoT Open Innovation Lab. This campaign was directed to a wide cross-section of Northeastern University's student and alumni community during the summer of 2017. This initial launch campaign was not associated with a specific event or challenge, but was instead meant to explain the goals and activities of the platform, in general, so as to establish a critical mass of participants to launch activity and to encourage future adoption. As the model of this program was to create a platform with many user interactions, where individuals join teams who then compete with one another, the goal was to bring large numbers of students and alumni from all fields on board in a short period of time.

This launch campaign provided invitations to 112,770 alumni and students of the university to participate in the program. Critical to the research design, it is not only possible to observe those who choose to participate on the platform, but also those who do not participate within the risk set and to distinguish the sex, degree or field of training, and graduation year of both. This set-up therefore provides a simple basis to evaluate the relative propensities of females and males to participate, accounting for field of training and age or cohort.

This launch campaign was carried out through electronic communications, whereby each subject was first contacted with an introductory email. From the

electronic mail, they could then click through to an online platform for further explanation. The online platform first described the Internet of Things as a technological opportunity, followed by the sorts of technology-related opportunities and activities that would become available to them by joining the IoT Open Innovation Lab's platform. The platform then provided an opportunity to affirmatively join the platform signing-in. In cases where an individual did not open the first email, one other follow-up email was delivered precisely one week after sending the initial email.

3.2.2 Collaboration and Competition Treatments and Random Assignment Procedures

The 112,770 alumni and students were randomized as to whether they were assigned to a competitive or collaborative treatment. For the experiment, we collaborated with the IoT Open Innovation Lab to build two separate campaign email systems and two accompanying separate online information and sign-up platforms, with the first emphasizing collaborative aspects of interactions among participants, and the second emphasizing competitive aspects of interactions among program participants. Both sets of descriptions were true and correct, but each emphasizing different aspects of the organization of the activity. This came down to modifying a small fraction of the text and explanation, with the bulk of detail presented in both treatments being otherwise identical. See the Appendix for copies of text contained in both the emails and the online platforms. To emphasize, in the end, just one program was run, which contained both elements of collaboration and of competition.

Each of the 112,770 subjects was also randomized according to the particular (business) day over the course of a two-month campaign that an individual would receive their invitation to join the program and platform (between August 1, 2017 until October 25, 2017). This first dimension of randomization simply attends to the empirical concern of smoothing any variation, and particularly any date-and-type-specific variation, in the propensity of reading and responding to the invitation to join over the summer and fall.

Randomization was carried out by random ordering of individuals and then alternately assigning to collaborative and competitive cells and to individual days to achieve balanced and random assignment.

3.2.3 Generalizability and Representativeness

As a field experiment at scale (Muralidharan and Niehaus 2017), it is important and useful to clarify precise dimensions and likely extent of representativeness. Regarding the nature of the work, this program is meant to capture and generalize to a wide array of existing tech work. At the same time, it should be stressed that the nature of technology is really focused on a set of information technologies (i.e., software, hardware, embedded devices, networking, cloud computing, data science, control systems). It should also be noted that the decision to participate in this tech program does not have the same consequences or opportunity costs of, say, choosing a program of study or career path or even a particular job. Choices to participate in this part-time and elective program are meant to be more accessible, but therefore less consequential. We therefore interpret estimated effects as informative of individuals' preferences at the margin.

Regarding the nature of the workers, as detailed further in the description of the data set, it is first important to note that the experimental population is highly representative in capturing a wide range of fields and ages of workers, but equally important to stress that the sampling frame focuses only on those who sort into higher education and only from the time they arrive at college and later. The sample is also entirely a U.S-trained majority, by virtue of focusing on just alumni and students of a U.S. university as a sampling frame. Further, the vast majority of individuals are Americans and living in the U.S. This is a helpful feature of the sampling frame in interpreting the leaky pipeline of workers, given that patterns in the U.S. may not entirely generalize to other countries, as was discussed in the Introduction. It should therefore be emphasized that this study should be read as a study of U.S. workers.

3.2.4 Interpretation of Experimental Preferences and Sorting (of Already Sorted Subjects)

In carrying out this experiment it should be important to bear in mind that the preferences driving sorting in the experiment should be the same preferences that have driven important career choices of the subjects within their lives outside of the experiment. Tangibly, the distribution of males and females within technical (and non-technical) fields within the experiment should already reflect the leaky pipeline of female technical workers in the United States that begins in early adolescence and is reflected in choice of fields of study in college education (Seymour 2002). Anticipating the empirical approach and later attempts to interpret results, here we illustrate how we might approach the interpretation of experimental results in relation to underlying sorting processes with several simplest examples.

3.2.4.1 Sorting into Technical Fields and Gender-Based Preferences

To illustrate possible sorting patterns and how they might be interpreted, consider a first simplest scenario in which gender Y has a higher taste or preference for tech relative to gender X.

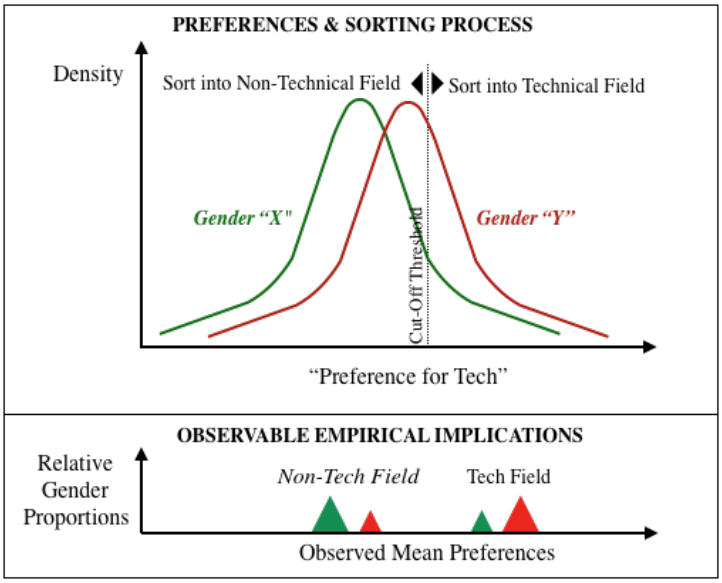
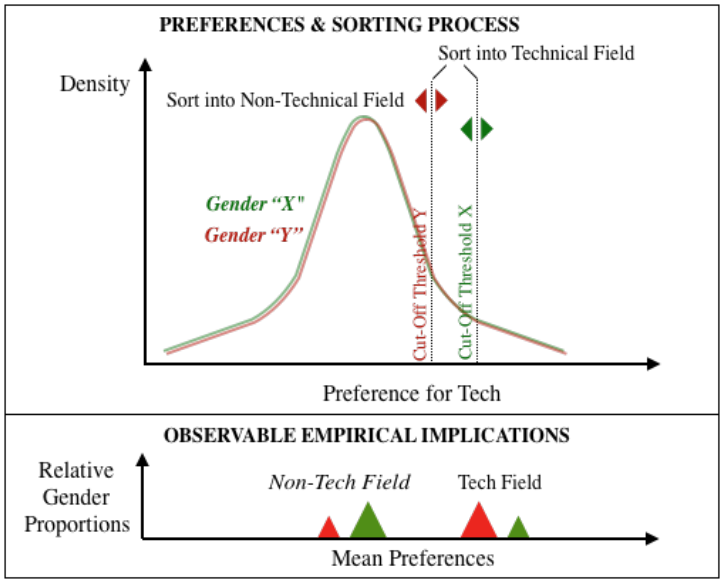


Figure 3.1: Case of General Gender-Based Preferences for tech across sexes

Figure 3.2: Case of Different Hurdles or Costs faced by the sexes



The differential interests of females and males are cited as an important factor contributing to differences in career choices (Lubinski and Benbow 2006, Ceci et al. 2009, Eccles 1994). Using the Person-Environment fit theories and People-Thing framework (Prediger 1982), studies have shown females to have higher propensity towards people and lower towards things (Woodcock et al. 2013). Therefore, under this simplest characterization, we might consider that individuals will enter into technical fields so long as their overall preference (sum of all pecuniary and non-pecuniary payoffs) exceeds the costs or hurdle or participating, as illustrated in Figure 3.1. Within such a simple interpretative framework, we might understand this scenario as one sex simply being right-shifted or left-shifted relative to the other. If this were the case, we would observe a larger proportion of gender Y in technical fields, and for gender Y to exhibit a higher mean preference for tech overall, and also for gender Y in tech fields or non-tech fields to exhibit a higher mean preference for tech in relation to gender X counterparts.

3.2.4.2 Sorting into Technical Fields and Differential Thresholds for Males and Females

In considering sorting and preferences within the experiment in relation to a population that has already sorted themselves into technical and non-technical fields, we should also bear in mind that it will be more than just supply-side preferences for tech work that affect the preferences we observe of those in tech and non-tech fields. We might summarize or conceive of all other effects shaping sorting to different fields as a threshold or differential cost for males and females with similar preferences to enter into technical fields.

A very simplest scenario in which a differential cost or threshold operates for genders X and Y, reflect that X and Y have identical preferences for tech, is illustrated in Figure 3.2. In such a scenario, although any population-wide measure would record similar preferences across X and Y, we would observe systematic differences among X and Y when compared within immediate peer groups in technical or non-technical fields. The higher cut-off for gender X, as in Figure 3.2, would lead to a higher mean preference for tech for gender X in either technical or non-technical fields.

These simplest scenarios and notions of underlying unobserved distributions of preferences and accompanying relative costs or hurdles will be referenced in interpreting the empirical analysis.

3.3 Data and Variables

As a means of better discerning population-wide patterns, the research design featured here is of a field experiment at scale in the words of Muralidharan and Niehaus (2017). We were able to securely collect and anonymize relevant information and contact details for the experiment for a total of 112,770 individuals, which includes 8,106 currently-enrolled students and 104,664 alumni. This observable population includes individuals trained in all fields [7]. Prior to anonymizing the data, we used a

[7] Northeastern has nine colleges and schools that offer bachelor's, master's, and doctoral degrees in a wide variety of academic disciplines and professional areas, levels of education (degree attainment), and career stages

Table 3.1: Variable definitions, means and standard deviations

	Description	Mean	Std. dev
Dependent Variable			
PARTICIPATION	Indicator for whether an individual registers on the platform	.0286327	.1667728
Explanatory Variables			
Female	Indicator for whether individual is a female	.4491663	.4974119
Compete	Indicator for whether in competitive condition	.4988639	.5000014
Genderedness	Percent of females in the area of study	44.85301	19.67246
HealthScience	Indicator for whether individual studied in Bouve Health Science	.1716231	.3770546
ArtsMedia	Indicator for whether individual studied in College of Arts and Media	.0662322	.2486888
ComputerScience	Indicator for whether individual studied in College of Computer Science	.037961	.1911031
Engineering	Indicator for whether individual studied in College of Engineering	.202279	.4017013
Science	Indicator for whether individual studied in College of Science	.0877592	.282946
SocialScience	Indicator for whether individual studied in College of Social Science and Humanities	.1430606	.3501365
Business	Indicator for whether individual studied in D'Amore Mc Kim School of Business	.2509442	.4335588
Law	Indicator for whether individual studied in School of Law	.0401407	.1962902
N	93739		

publicly-available resource (https://gender-api.com/) to code the probabilistic gender of each individual. This risk set of potential participants is 45 percent female. One-third of sample participants attained a graduate degree. Apart from begin US-trained, a majority of the individuals are located in the United States. These data were matched to data describing randomized assignments of individuals. We also collected detailed observational data by working with members of the IoT Open Innovation Lab to implement relatively standard mail campaign analytics (to collect statistics on the opening of electronic mail and click-throughs to the platform).

Table 3.1 lists and defines main variables in the study. The median years of experience among alumni is 12. In terms of the breakdown by career stage, our sample is comprised of 4 percent undergraduate students, 2 percent graduate students, 44 percent junior alumni (with less than 12 years of experience) and 51 percent senior alumni.

Although assignments were randomized according to the procedures described above, it is useful to examine whether, ex- post, they are similar in their characteristics across competitive and collaborative treatments, as in Table 3.2.

Certain fields of study are more skewed to one gender or the other, as in Table 3.3. The gender breakdown by field of study in our sample[8] given in Table 3.3 is broadly consistent with the overall gender distribution across fields of study in higher education in the United States (Bui 2014). Further, the proportion of females

[8]There are 8 broad areas of study offered at the University: Engineering, Computer Science, Business, Social Science, Science, Law, Arts and Media, and Health Science

Table 3.2: Descriptive Statistics, Stratified by Competition and Collaboration Treatments

	Mean(Collaborate)	Mean(Compete)	difference	p	N
Female	.4482064	.4501318	-.0019254	.5570546	92055
HealthScience	.1723077	.1709344	.0013733	.579333	92674
ArtsMedia	.0661431	.0663218	-.0001787	.9129091	92674
ComputerScience	.0377622	.038161	-.0003988	.7507836	92674
Engineering	.2006025	.2039655	-.003363	.2025626	92674
Science	.0892738	.0862356	.0030382	.1021773	92674
Humanities	.1418182	.1443105	-.0024923	.2786098	92674
Business	.2517698	.2501136	.0016561	.5609571	92674
Law	.0403228	.0399576	.0003652	.7770446	92674
Genderedness	44.92887	44.7768	.1520665	.2366825	93739
N	93739				

engaged in university participation in our sample has grown over past decades, as reflected in Table 3.4.

Our key variable of interest is the individual's decision to participate in the tech opportunity upon receiving the email. From the platform, itself, we observe the choices of individuals and whether they choose to participate or not by December 31. By the end of the 67-day campaign, a total of 3116 subjects registered to participate, or a participation rate of 2.7 percent. The overall number of females who joined was 1259 and the number of males who joined was 1839. Among those who chose to participate, there are 653 students (24 percent) and 2031 alumni (76 percent). Overall, we observe the highest participation in the areas of Engineering and Business and the lowest in Law.

3.4 Results

3.4.1 Estimating Equation

The unit of analysis is the individual within the broad risk set of potential participants, indexed by i. The outcome of interest being modeled is whether an individual chooses to join this tech program (within the two weeks following receipt of the first invitation they receive). A fully-elaborated model of this choice might, say, attempt to factor in cost and benefit functions, allowing for heterogeneity across the population. Here, however, we focus on a reduced-form formulation to summarize differences between females and males in the most direct and simplest manner.

Table 3.3: Proportion of Females and Male by Field of Study

Field of Study	Proportion of Males	Proportion of Females	Total
ArtsMedia	2,373 39.09%	3,698 60.91%	6,071 100%
Business	14,706 64.28%	8,173 35.72%	22,879 100%
Computer Science	2,519 73.78%	895 26.22%	3,414 100%
Engineering	15,069 82.29%	3,244 17.71%	18,313 100%
HealthScience	3,678 23.43%	12,022 76.57%	15,700 100%
Law	1,572 43.4%	2,050 56.6%	3,622 100%
Science	3,501 43.86%	4,481 56.14%	7,982 100%
SocialScience	6,711 51.52%	6,315 48.48%	13,026 100%
Total	50,129 55.08%	40,878 44.92%	91,007 100%

Table 3.4: Proportion of Females and Males (in sample) by Graduation Decade

Graduation Decade	% Male	% Female
1950	91.51	8.49
1960	84.29	15.71
1970	69.80	30.20
1980	58.49	41.51
1990	54.61	45.39
2000	47.80	52.20
2010	50.38	49.62

Within a reduced-form framework, we model the likelihood that an individual i chooses to participate in a work opportunity, in relation to characteristics of the work, characteristics of the work environment, characteristics of the individual, and a random component. Characteristics of the work and work environment are in our context, of course, held constant apart from the emphasis on either competitive or cooperative interactions with other participants. Therefore, all variation in the work environment is captured by our indicator variable corresponding to the work environment to which the individual is assigned $Competition_i$.

Individuals' characteristics that could affect preferences for this work opportunity are divided into field of training (captured by a set of fixed effects), experience, age, and life circumstances (captured by another set of fixed effects for different age cohorts from students, to junior alumni, to senior alumni); gender (captured by the dummy variable, *Female*; and other characteristics and random error ε. We are interested in estimating whether there is a differential response among otherwise similar females and males. Our main model of interest, therefore, takes the following form:

$$\begin{aligned} Prob\{Participation_i\} = \beta \cdot Female + \gamma \cdot Competition_i + \\ \delta \cdot Female_i \cdot Competition_i + FieldFE_i + AgeCohortFE_i + \varepsilon_i \end{aligned} \tag{3.1}$$

where β, γ, δ are model coefficients and ε is an error term to be estimated. The error term ε should be interpreted as the sum of any purely random error term ε and other characteristics not controlled by the model. The main coefficient of interest is that on Female, β. This measures the differential probability of females to participate, conditional on their degree area of study, the organization approach, and their age cohort. Secondarily, we are interested in coefficients γ and δ in assessing whether there is any response to work organized to be competitive and whether females experience a differential (negative) response. Crucial to this set-up, we are not only able to estimate these coefficients while controlling for field and age or cohort fixed effects; we are also able to directly estimate any heterogeneous

relationships, as the scale and breadth of the experimental population under study allows us to estimate the coefficients of interest for these field and age subsets (either by interactions or stratified model regression).

Given the randomized nature of the experiment and given the simplicity of interpreting linear probability models, we will estimate coefficients via ordinary least squares (OLS). Results are robust to nonlinear specification, as in logit and probit. We report heteroskedasticity-robust standard errors.

At this juncture of presenting the estimating model, it should be emphasized that the measurement of gender-based preferences in this research (and preceding such experiments in behavioral economics) is, of course, not a direct observation of preferences. Rather, we may observe systematic differences in choices and differences in revealed preferences. Moreover, even differences in revealed preferences here should be thought of in the broad sense of reflecting differences in expectations of benefit, net of costs, in choosing to participate in this IoT program, conditional on whatever controls are included.

3.4.2 Overall Gender-Based Differences

The baseline results presented here report average patterns across the entire experimental population. This serves to illustrate how average patterns conform to stereotypical gender-based preferences and these results are seemingly robust to a range of specifications and analyses.

3.4.2.1 Overall Gender-Based Differences in Preferences for Tech

Model (1) of Table 3.5 begins with a simplest specification, regressing the discrete decision to participate in the program, Participation, on the indicator variable, *Female*, and a constant term. This highest level description of the data shows that men in this broad and representative experimental population are more likely to participate. On average and unconditionally, 3.2 percent (s.e. = 0.1 percent) of males chose to participate (as given by the point estimate of the constant term). Females are 0.7 percent (s.e. = 0.1 percent) less likely to participate, almost one-quarter lower chance of participating relative to males.

Table 3.5: OLS estimates of overall Participation

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.0074***	-0.0074***	-0.0075***	-0.0075***	-0.0110***	-0.0037**
	(0.0011)	(0.0011)	(0.0016)	(0.0016)	(0.0016)	(0.0017)
Competition		-0.0068***	-0.0070***	-0.0069***	-0.0066***	-0.0063***
		(0.0011)	(0.0016)	(0.0016)	(0.0015)	(0.0015)
Female X Competition			0.0003	0.0002	0.0004	-0.0001
			(0.0022)	(0.0022)	(0.0022)	(0.0022)
HealthScience						-0.0047**
						(0.0023)
ArtsMedia						0.0055+
						(0.0030)
ComputerScience						0.0433***
						(0.0050)
Engineering						0.0196***
						(0.0026)
Science						0.0052*
						(0.0028)
Humanities						0.0043*
						(0.0024)
Business						0.0092***
						(0.0024)
Law						0.0000
						(.)
Constant	0.0323***	0.0357***	0.0358***	0.0301***	0.1165***	0.1011***
	(0.0008)	(0.0010)	(0.0012)	(0.0040)	(0.0087)	(0.0091)
Age Dummies	No	No	No	No	Yes	Yes
Day Dummies	No	No	No	Yes	Yes	Yes
Observations	92055	92055	92055	92055	90170	89949
R^2	0.0005	0.0009	0.0009	0.0017	0.0181	0.0210

Heteroskedasticity robust Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.4.2.2 Overall Gender-Based Differences in Preferences for Competition

Model (2) adds the indicator variable, *Competition*, to determine whether we observe any evidence this variation is in any way interacting with females' willingness to engage in this tech opportunity. As reported in model (2), we indeed find a significant negative causal effect of competitive interactions on willingness to participate, as indicated by the negative 0.7 percent coefficient (s.e. = 0.1 percent) on *Competition*. Note that, at least to one decimal place, the estimated coefficient on *Female* remains unchanged. Therefore, relative to male participation, the level of participation among females is lower by about an entire quarter or 25 percent. This drops again by another quarter under the competitive treatment.

Model (3) then adds the interaction of *Female* and *Competition*, and finds there to be no statistical interaction. That there is no interaction between Female and

Competition, is consistent with the competitive work organization in no way explaining or interacting with any relative disinclination of females to participate in this tech program. While the absolute effect of competition is the same across males and females, the relative effect for females is considerably greater given the lower baseline participation rates of females. These results are consistent with past findings in the experimental literature that females are relatively disinclined to participate under competitive conditions (Niederle and Vesterlund 2007, Croson and Gneezy 2009, Niederle and Vesterlund 2011).

Model (4) confirms that results are unaffected by adding dummies for the particular days in which individuals first received their invitation to join. Recall that individuals were randomized across which of 60 days they would receive the invitation to join the program. Given the randomization, it is unsurprising that estimates do not statistically change. However, we retain day dummies in the analysis to follow, as this appears to usefully explain variation.

Model (5) adds the age or cohort dummies to exploit variation (i.e., fixed effects for students, junior workers with fewer than 12 years experience, and senior workers with greater than 12 years experience). We find that the baseline likelihood of participating declines somewhat in older cohorts consistent with potentially higher opportunity costs and possibly greater distance from new technologies. Model (6) reports these field fixed effects along with controls for year and cohort effects. These control variables also explain a good deal of variation in adoption decisions. However, adding these controls does not perturb the estimated coefficients on *Female* and *Competition* and their interaction.

3.4.3 Gender-Based Preferences by Career Stage

Our sample covers cohorts across quite a wide range of ages. With changing social norms and increased participation of women in the labor force over the last few decades, we run our analysis on the sample stratified by career stage to see whether our main result holds across cohorts. Here, we distinguish undergraduate students, graduate students, junior professionals (graduated with less than the median of 12 years of experience) and senior professionals (greater than 12 years experience).

We re-estimate Model (2) of Table 3.5 stratified by field of study in Table 3.7. The coefficient on *Female* is negative and significant across all career stages, consistent with the full sample results. This implies that the likelihood of participation of female undergraduate students is 23 percent less compared to that of undergraduate males. Among graduate students, females are 31 percent less likely to participate compared to male graduate students. Female junior alumni are 34 percent less likely to participate compared to men, while female senior alumni are 21 percent less likely to participate than males. The coefficient on *Competition* is negative and significant across all career stages except for Undergraduate students where it is insignificant. We also report the ratio of female to male participation in our tech program in each career stage. While comparing cohorts is inherently complex, we do not find any increasing or decreasing trend in the gender gap over career stages. Therefore, at least in population-average patterns, females have an average lower propensity to join this representative tech work and learning opportunity, and have a proportionally higher aversion to competition. Moreover, the persistence of the gender difference in willingness to participate is seen across different cohorts, right from the level of undergraduate students to junior and senior alumni with several years of experience in the industry, seems to suggest that these differences, or at least a greater part of it, existed before entry into college level education.

Table 3.7: OLS estimates of Participation Stratified by Career Stage

	(1) Undergraduate	(2) Graduate	(3) Junior Alumni	(4) Senior Alumni
Female	-0.0245**	-0.0450***	-0.0161***	-0.0034***
	(0.0096)	(0.0155)	(0.0019)	(0.0011)
Competition	-0.0064	-0.0289*	-0.0092***	-0.0031***
	(0.0096)	(0.0155)	(0.0019)	(0.0011)
Constant	0.1086***	0.1494***	0.0484***	0.0164***
	(0.0086)	(0.0131)	(0.0016)	(0.0009)
Observations	3627	1671	39215	45657
R^2	0.0019	0.0073	0.0025	0.0004
Ratio of Female to Male Participation	0.77	0.69	0.66	0.79

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.4.4 Gender-Based Preferences by Field

We first examine results when field fixed effects are added. In Model (1), the estimated coefficient on *Female* significantly diminishes (compared to our previous findings) when adding field fixed effects. This is either consistent with there truly being no gender-based preferences, once controlling for which field individuals chose to sort into (i.e., prior to the experiment); or that there may be gender-based preferences, but more discerning analysis of differences across field is required. The following clarifies that the situation is the latter.

In order to investigate differences by field of study, we stratify the sample into eight different Colleges and Schools at the University[9]. We first re-estimate our preferred Model (2) of Table 3.5 – i.e., using only significant explanatory variables related to *Female*, *Competition* and a constant– but doing so separately for each field of study. Model (1) of Table 3.8 reports OLS estimates for all fields. We also report the ratio of female to male participation in the tech program across areas. The coefficient on *Female* varies widely in significance and magnitude across areas. The point estimate is negative and insignificant for non-technical areas of Arts and Media and Business, while for Humanities and Law it is positive and insignificant. In Health Sciences, females are significantly less likely to participate than males, as indicated by the negative and significant coefficient on *Female* of -0.01 (p= 0.01). This amounts to female participation being 48 percent lower than that of males. Results reverse for tech areas of Engineering and Computer Science. In Engineering, females have a 54 percent higher propensity to participate compared to similarly trained males as indicated by the positive and significant coefficient on *Female* (p=0.01). The absolute participation rate for females is even higher in Computer Science, but the ratio with men is lower at 1.05, with male participation rates also considerably higher in Computer Science. This is especially notable given that Computer Science and Engineering are the very fields most associated with tech industries and, it follows too, they are the fields most closely-relevant to IoT.

[9]In our analysis, we dropped 19,031 individuals from the College of Professional Studies which offers a variety of programs and accreditation catered to part-time students and working adults

Table 3.8: OLS estimates of Participation stratified by Field of Study

			ArtsMedia	Business	Computer Science	Engineering	Health Science	Science	Humanities	Law
Female	0.0005	0.0005	-0.0072	-0.0004	0.0051	0.0214**	-0.0103**	-0.0060	0.0001	0.0008
	(0.0013)	(0.0013)	(0.0044)	(0.0023)	(0.0106)	(0.0044)	(0.0025)	(0.0039)	(0.0026)	(0.0042)
Competition	-0.0068**	-0.0069**	-0.0125**	-0.0039+	-0.0239**	-0.0065*	-0.0034*	-0.0106**	-0.0072**	-0.0047
	(0.0011)	(0.0011)	(0.0042)	(0.0022)	(0.0092)	(0.0029)	(0.0017)	(0.0039)	(0.0026)	(0.0041)
HealthScience	-0.0674**									
	(0.0048)									
ArtsMedia	-0.0516**									
	(0.0051)									
ComputerScience	0.0000									
	(.)									
Engineering	-0.0393**									
	(0.0048)									
Science	-0.0486**									
	(0.0050)									
Humanities	-0.0575**									
	(0.0048)									
Business	-0.0505**									
	(0.0047)									
Law	-0.0635**									
	(0.0051)									
Genderedness		-0.0005**								
		(0.0000)								
Constant	0.0824**	0.0514**	0.0383**	0.0308**	0.0898**	0.0392**	0.0216**	0.0393**	0.0252**	0.0176**
	(0.0047)	(0.0015)	(0.0046)	(0.0018)	(0.0075)	(0.0022)	(0.0025)	(0.0038)	(0.0023)	(0.0039)
Observations	91007	92055	6071	22879	3414	18313	15700	7982	13026	3622
R^2	0.0070	0.0034	0.0019	0.0001	0.0020	0.0020	0.0019	0.0012	0.0006	0.0004
Ratio			0.81	0.98	1.05	1.54	0.52	0.84	1.00	1.04

Standard errors in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

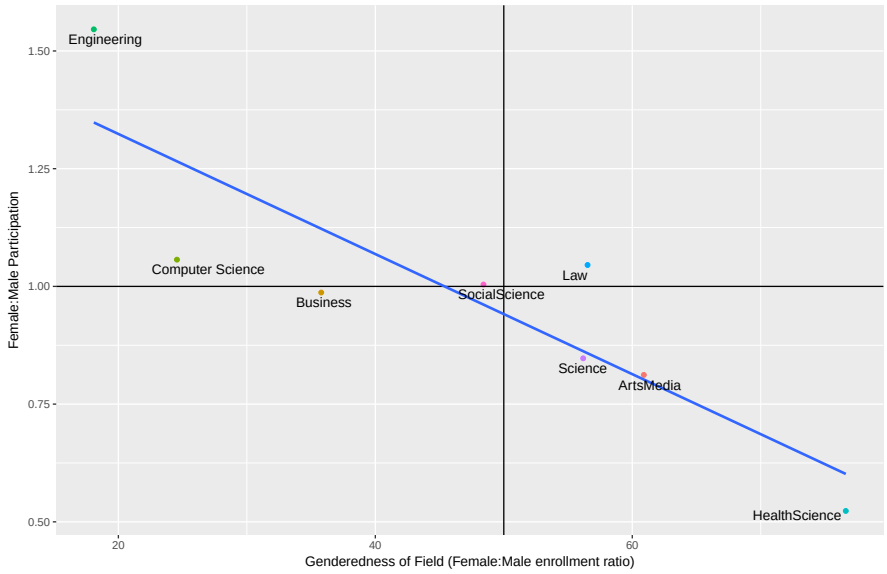


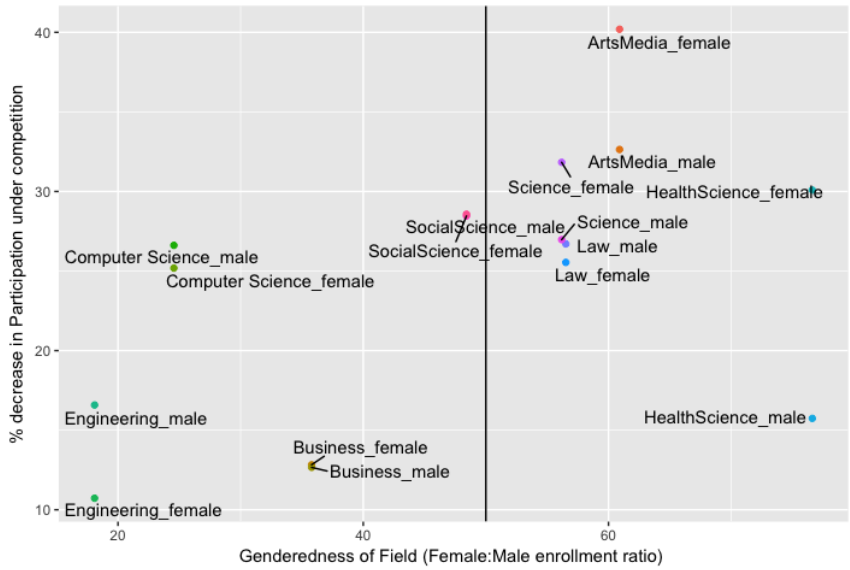
Figure 3.3: Ratio of Female to Male Participation vs Genderedness

Results are also reported graphically in Figure 3.3. Each point in Figure 3.3 denotes estimates from distinct fields (i.e., separately estimated models). The vertical axis plots the ratio of female to male participation and the horizontal axis represents the *Genderedness* of a field, or the proportion of females in a field. There is also a simple fitted line to these data to highlight that fields with the highest proportion of female enrollment have the lowest ratio of female to male participation in the program. In the technical areas of Engineering and Computer Science, (where the proportion of females is 18 percent and 26 percent respectively) the female to male ratio of participation in the tech program is greater than 1 (1.54 for Engineering and 1.05 for Computer Science). On the other hand, the ratio of female to male participation in the tech program declines as the proportion of females in a field increases, as can be seen across Humanities, Arts and Media and the Health Sciences. In the case of Health Sciences, where females constitute 77 percent of the enrollees, the ratio of female to male participation in the program is 0.52.

We now turn to exploring gender differences across individuals in different fields of study under Competition. Consistent with the full sample results, we find the coefficient on *Competition* in Table 3.5 to be negative and significant across all fields of study with the exception of Law where it is negative but insignifi-

cant. Therefore, competition has a dissuasive effect on participation among men and women across all fields of study. Results are presented graphically in Figure 3.4. Here, the vertical axis represents the percentage decrease in participation under competition and the horizontal axis denotes the Genderedness of the field of study. Each field has two points in the graph- one for males and one for females. We find widely varying differences across areas in the proportional decrease in participation induced by competition for males and females. Consistent with the full sample results, under competition, female participation reduces proportionally more compared to males for most fields of training (this is also consistent with gender economics findings of Flory et al. (2014)), however, this trend reverses for technical areas of Engineering and Computer Science. For example, in Engineering, participation of females under competition reduces by 10 percent compared to their baseline, whereas it reduces for males by a greater amount of 16 percent relative to their baseline. This case of Engineering women runs counter to the robust lab results in gender economics (Croson and Gneezy 2009, Niederle and Vesterlund 2011) where women have been found to have a lower taste for competition compared to men. We also note widely varying gender differences in the magnitude of reduction in participation under competition across areas. For example, in the Health Sciences, male participation under competition reduces by 16 percent but that for females decreases by a much larger margin of 30 percent. It is noteworthy that the largest decline in female participation under competition (relative to males) is in a helping occupation where women have traditionally been over-represented.

Figure 3.4: Percent decrease in participation of males and females under competition across fields



3.4.4.1 By Field of Study and Majors

We also explore differences across individuals sorted into individual majors within each field of study. The most important first-order pattern that will be revealed here is simply that whereas there are important differences across fields, there are far fewer differences within a given field and across majors.

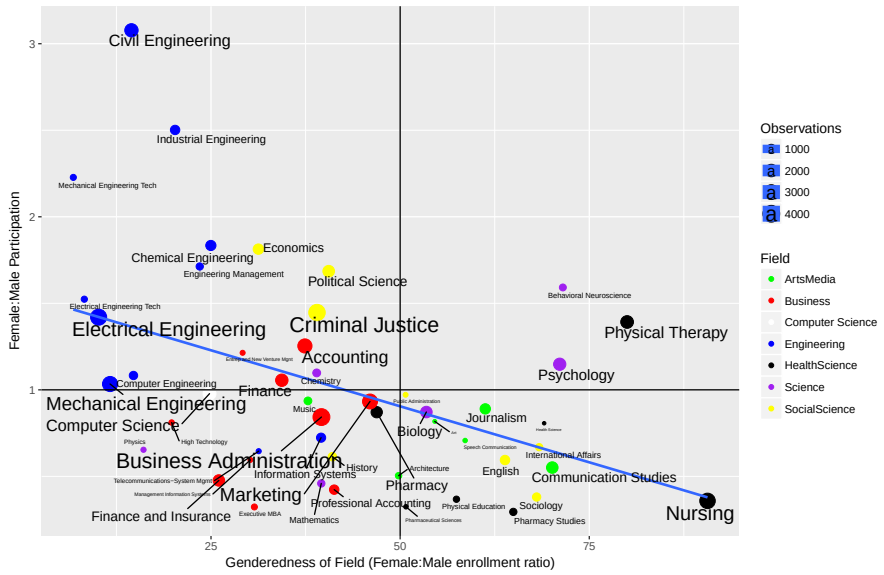
We re-estimate Model (2) of Table 3.5 separately for all majors within each field of study. For example, in Engineering, there are subfields or majors of civil, industrial, computer, electrical, biological, mechanical and so on. Likewise, there are majors or subfields affiliated with each of the schools at the university. For ease of interpretation, we combine results across all fields and all majors and report these in Figure 3.5. Each individual point represents estimates from a distinct major [10]. The size of a dot in the figure corresponds to the number of individuals in that major. The horizontal axis represents the *Genderedness* in a major whereas the vertical axis represents the ratio of female to male participation in a major. Notably, the two points corresponding to the two most gendered fields of Engineering and

[10] due to the large array of programs and majors of study, we limited our analysis to majors with more than 300 observations

Health Sciences are at the extreme ends of the fitted line.

In comparison to the large differences in the ratio of female to male participation across fields found earlier, we find smaller differences within fields (across majors). It appears that main results linking differential propensity to participate in tech and genderedness of fields is mainly driven by the field and not the subfield.

Figure 3.5: Ratio of Female to Male participation vs Genderedness of field of study



3.4.4.2 Gender-Based Sorting in the Experiment and Gender-Based Sorting into Fields

The above patterns capture important differences across males and females, who by the time of university, have already sorted themselves into different fields. To the extent that the earlier patterns reflect more systematic gender sorting and separation into these different fields, we might expect some number and degree of earlier results are explained by a measure of gender-based sorting into fields that precedes this experiment. To test this idea, we construct a measure of prior gender sorting that is simply the gender proportion or "Genderedness" by field, the proportion of females in each subject's field.

As reported in Model (2) of Table 3.8, *Genderedness* is highly predictive of choices made in the experiment. The negative coefficient on *Genderedness* indi-

cates that the higher the percentage of females in one's field of study, the lower the probability of one's participation in the tech program. Equivalently, the likelihood of participation is highest in areas Gendered in favor of men (i.e. with the lowest percentage of females) such as the technical areas of Engineering and Computer Science. Most stark, this external measure of sorting in the wider economy, outside of the experiment, even explains away the coefficient on the female dummy within the experiment, ascertaining that our experimental results and sorting patterns are explained by differences in proportions of female to male enrolment in fields of study. These results strongly affirm that the design of this program in regards to the experimental design successfully relates to sorting in the wider economy and is externally valid.

Lastly, probit estimates are reported in Table 3.9.

3.5 Discussion and Conclusion

This paper presented estimated gender preferences for participating in tech-related work in an unusually large and widely-representative field experiment that included individuals across all fields and career stages. Females, on average, have a lower inclination to participate in the tech-related opportunity in the experiment and are about 25 percent less willing to participate than males. However, this reflects female preferences relative to males of those within (the more heavily weighted) non-technical fields. Within technical fields, females are at least as willing to participate. Thus, the especially wide and representative sample across fields proves important to discerning these patterns. Results are robust across different cohorts at different ages and stages of education and career.

The evidence here indicates the results reflect preferences for the nature of tech work, and are not explained by preference for or anticipation of the tech work environment. Although we find evidence of gender-based preferences for the work environment here, it in no way interacts with or explains participation patterns. Further, tastes for competition here are found to be closely-related to preferences for tech work. Analogous to the above results, females were relatively less will-

ing to participate under the competitive treatment, on average. However, again, this reflects females preferences relative to males of those within (the more heavily weighted) non-technical fields. Within technical fields, females are at least as willing to participate under the competitive treatment.

The results contrast with simplest possible interpretations of sorting processes or leaky pipeline of females in technical occupations, as discussed in Section 3.4 and as summarized below in Figure 3.6. The exhibited results are inconsistent with general and prevailing gender-based preferences for tech (or competition), in which case we would expect female preferences for tech to be generally lower in all fields, relative to male counterparts. The observed patterns are also inconsistent with simplest characterizations of females having similar preferences to males but having to meet a higher threshold or hurdle to enter into technical fields. Under such an interpretation, not only would we expect higher preferences for tech among those in technical fields, but also within non-technical fields.

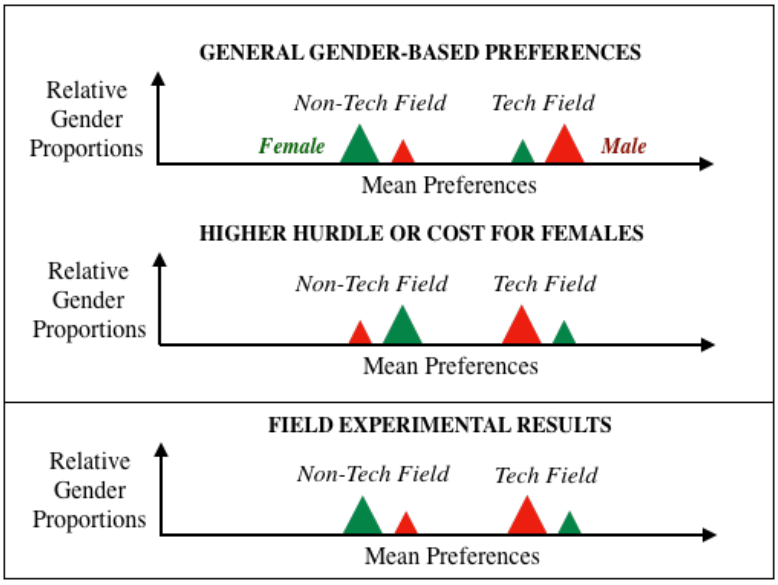


Figure 3.6: Summary of Scenarios

Instead, the findings presented here appear to be consistent with either a combination of gender-based preferences (in which the distribution of preferences tends to be lower among females), but where females also face a higher threshold for entering into technical fields. The results are also consistent with more complex dis-

tributions of preferences whereby females selecting into technical fields possess or are socialized to develop more extreme preferences for tech. For example, women who seek a career in male dominated areas are socialized into those cultures through education, and actively seek that culture later on (Bennett et. al, 1999). Females in Engineering have been reported to do gender in ways to gain acceptance and cope with the Engineering culture by emulating males in their program, accepting gender discrimination, achieving a reputation of capable engineers (Powell et al. 2009). It is possible that women in Engineering develop more extreme preferences as a result of socialization.

Our findings should be interpreted with caution. The limitations of our research design can pave the path for further studies to validate our findings in a lab environment to control for aspects such as personality traits which is often not achievable in a field setting. To the extent that it was possible, we aimed to communicate the nature of our synthetic experimental object (the activity that individuals will participate in) as an activity involving the design and prototyping of Internet-of-Things related ideas. This could be interpreted by potential participants in several ways depending on whether it is seen as an extra-curricular activity or a form of part-time work. While it is not a "learning" experience per se as it does not involve instruction delivery such as in a typical online class, it is neither a "work" activity and is not intended to invoke usual connotations of full time employment. However, certain dimensions of the activity that our synthetic experimental object represents relate to technology related work, such as software development, systems design, networking, computer hardware engineering etc.

Next, while we do not unpack "preferences", this could encompass a number of factors such as opportunity costs, confidence levels, social preferences, risk preferences etc (Croson and Gneezy 2009). For example, the differences between men and women can be attributed to differences in opportunity costs of time. Our design allows us to examine differences in groups of comparable men and women across different cohorts including those in University who have similar opportunity costs of time (i.e. similar family obligations, involvement in student clubs, course work

etc).

Beyond increasing the supply of female participants in tech-based crowdsourcing work, the findings have broader implications for policy in attempting to expand the supply of technical workers, and particularly to remedy the leaky pipeline of female technical workers. It appears that supply-side preferences for tech are important in determining which females enter into technical fields. The persistence of gender difference in willingness for participation across different cohorts (including undergraduate students) suggests that much of these preferences may have been established early in life prior to university. The results are consistent with the leaky pipeline beginning at an early age, well before higher education, resulting in the separation and by the time they enter into university and appear within the sampling frame. This is also consistent with much of sorting and preferences in the experiment being statistically explained by correlation with the proportion of women sorting to different fields. Whether the extreme preferences of females in technical field are the result of facing high hurdles or accentuated socialization, these are potentially both addressable by productive interventions.

Girls and young women find themselves in a stereotype threat situation when performing with technology in general (Cooper and Weaver 2003). Such a stereotype may signal to women that their group does not "belong" and may see this as a barrier or deterrent, leading them to question their identity as a female in STEM. Reminders of identity compatibility (that is, being a female in STEM areas) can also undermine academic performance (Cheryan et al. 2009). Furthermore, the lack of role models in STEM areas and professions reinforces the identity incompatibility and gender-task stereotype which has been shown to affect women's self-selection into competition Kamas and Preston (2012). Our findings may also have implications for tech bootcamps for school girls aimed at encouraging long-term participation in the STEM areas by providing a more friendly environment to females through introducing women role models and mentors at an early stage that could help contradict stereotypes of gender incompatibility. Moreover, to the extent that preferences are socially constructed over the formative years during school, inter-

ventions at an early stage may go a long way in changing perceptions of technology related work.

Appendix

Figure 3.7: Email for Competitive condition

Dear *|FNAME|*,

I'm reaching out to invite you to Northeastern's new **IoT Open Innovation platform**, linking our students, alumni, staff, faculty and affiliated companies.

This is a two-sided **competitive crowdsourcing** platform to ideate, innovate, and prototype new Internet of Things (IoT) applications using hardware, software, networking, data, and knowledge of use-cases.

On one side of the platform, **companies** seek solutions to IoT innovation challenges. On the other side of the platform, **you** can **COMPETE** to contribute to solutions for cash and other benefits.

|CONTENT|

|NETWORK|

Click HERE to learn more and sign up to the platform.

(This invitation is not transferable or to be forwarded.)

Sincerely,

Figure 3.8: Email for Collaborative condition

Dear *|FNAME|*,

I'm reaching out to invite you to Northeastern's new **IoT Open Innovation platform**, linking our students, alumni, staff, faculty and affiliated companies.

This is a two-sided **collaborative crowdsourcing** platform to ideate, innovate, and prototype new Internet of Things (IoT) applications using hardware, software, networking, data, and knowledge of use-cases.

On one side of the platform, **companies** seek solutions to IoT innovation challenges. On the other side of the platform, **you** can **COLLABORATE** to contribute to solutions for cash and other benefits.

|CONTENT|

|NETWORK|

Click HERE to learn more and sign up to the platform.

(This invitation is not transferable or to be forwarded.)

Sincerely,

Figure 3.9: Main platform web page

Figure 3.10: Competitive Page

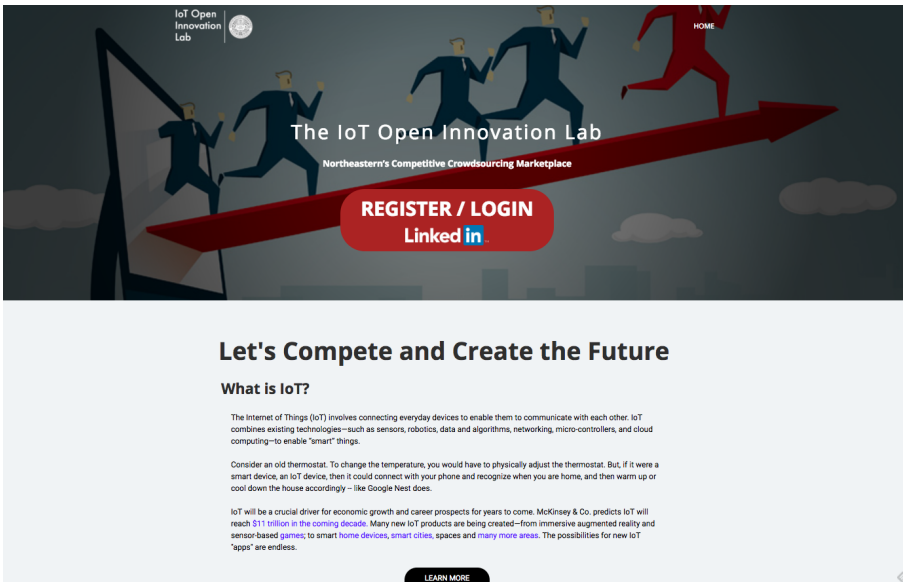


Figure 3.11: Collaborative Page

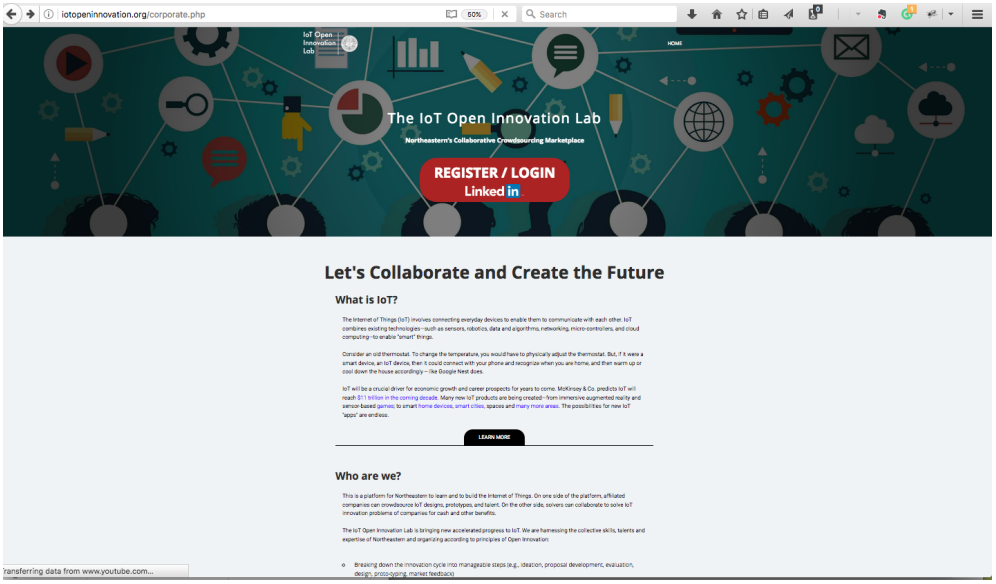


Table 3.9: Probit estimates of overall Participation

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.1135**	-0.1136**	-0.1054**	-0.1057**	-0.1638**	-0.0555*
	(0.0171)	(0.0171)	(0.0231)	(0.0231)	(0.0246)	(0.0264)
Competition		-0.1037**	-0.0965**	-0.0962**	-0.0963**	-0.0931**
		(0.0168)	(0.0218)	(0.0218)	(0.0229)	(0.0231)
Female X Competition			-0.0179	-0.0196	-0.0204	-0.0304
			(0.0343)	(0.0343)	(0.0361)	(0.0366)
HealthScience						-0.1244*
						(0.0615)
ArtsMedia						0.1433*
						(0.0639)
ComputerScience						0.4948**
						(0.0633)
Engineering						0.3363**
						(0.0568)
Science						0.1366*
						(0.0613)
Humanities						0.1045+
						(0.0594)
Business						0.1965**
						(0.0565)
Law						0.0000
						(.)
Constant	-1.8480**	-1.7988**	-1.8021**	-1.8963**	-1.1823**	-1.4530**
	(0.0109)	(0.0133)	(0.0148)	(0.0709)	(0.0838)	(0.0996)
Age Dummies	No	No	No	No	Yes	Yes
Day Dummies	No	No	No	Yes	Yes	Yes
Observations	92055	92055	92055	92055	90170	89949
Pseudo R^2	0.001	0.003	0.003	0.0067	0.05	0.0679

Standard errors in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Chapter 4

Evaluation of innovative crowdsourcing ideas

4.1 Introduction

In recent times, organizations have turned to open innovation (Chesbrough et al. 2003) and crowdsourcing which offer the potential to access distant and diverse knowledge via engagement with external crowds. While crowdsourcing can take many forms, one form involves broadcast search where firms broadcast problems to a community of solvers to gather solutions. A common issue organizations engaged in crowdsourcing efforts encounter is that of processing and evaluating the large number of ideas or solutions generated. For example, since Dell launched its ideation platform, the IdeaStorm[1], it has generated more than 28,000 ideas and over 100,000 comments since its inception in 2007. Out of these, only 550+ user contributed ideas have actually been implemented. Similarly, British Petroleum (BP) received over 100,000 suggestions to its open call to the general public after the Deepwater Horizon accident (Mascarelli 2010). How do experts go about evaluating ideas and solutions generated through these new channels and which kinds of ideas are successfully evaluated by the gatekeepers responsible for evaluation?

Institutional gatekeepers play an important role in assessing, filtering and enforcing selection criteria. For example, patent offices act as institutional gatekeep-

[1] `https://dellideas.secure.force.com/`

ers in determining the inventiveness and novelty of a patent application compared to prior art. In the case of scientific publications, journals act as gatekeepers for publication quality. In mediated markets, seeking coverage and recognition from analysts that follow categories impacts firm survival. Evaluation by institutional gatekeepers can make the difference between proceeding to a further round of funding, such as in venture capital funding, or progressing to further rounds of proposal development such as for scientific proposals sought by funding agencies. Such staged evaluation reduces the risk for organizations in awarding or investing in under-developed ideas that have have yet to prove their market potential. A growing body of literature highlights the role of institutional gatekeepers in evaluating novelty and recombinant knowledge breadth in various settings (Ferguson and Carnabuci 2017, Porter and Rossini 1985, Boudreau et al. 2016, Rindova and Petkova 2007).

Through the empirical context of a crowdsourcing platform that brings together a community of solvers to design solutions for complex social problems, we explore which types of ideas are successfully evaluated by institutional gatekeepers or domain experts. In our setting, knowledge gets reconfigured using 'fresh" approaches from a variety of domains such as public policy, health care, computing etc.

The existing innovation literature has largely generated insights on the antecedents of successful innovations (through distant search/search scope) using patent data and is premised upon pre-defined patent or technology classes or journals. In this work, we use a natural language processing technique to decompose text-based descriptions of proposed solutions into constituent themes of latent topics to operationalize measures of proposal's recombinant breadth, atypicality or novelty of the proposal and the degree of distant search (or content distance), and study whether these measures are associated with the likelihood of successful evaluation. We extend this to analyze which patterns of combinations are more or less likely to succeed.

In prior innovation studies, an invention's success is measured in terms of the number of forward citations and economic value it generates over time, and inventiveness has been the key criteria for patent examiners in evaluating patents. In con-

trast, in the setting we study, the success of contributor proposed ideas/solutions in crowdsourcing is judged by the host organization or institutional gatekeepers who are more interested in immediate solutions and feasibility of proposed solutions rather than inventiveness. Secondly, while breakthrough inventions in the innovation literature are known to be associated with wide combinations of knowledge and distant knowledge, we have limited understanding of whether broad or narrow combinations of knowledge, or distant knowledge in developing ideas and solutions are associated with crowdsourcing success.

One key related study reports on 922 organizations that respond to 100,000+ crowdsourced suggestions from external contributors Piezunka and Dahlander (2015). It is shown that when faced with a large number of ideas, organizations pay attention to familiar ideas rather than ideas that embody distant knowledge (they operationalize content distance in terms of how the content of a suggestion diverges from the content of suggestions to which an organization has previously attended to), and that this gets more pronounced in the event of crowding. Therefore, despite getting access to distant knowledge, in the filtering stage institutional gatekeepers appear to struggle to make use of the knowledge they gained through crowdsourcing. A few distinctions and differences that set our context apart is the solicitation of proposals to a problem rather than open-ended suggestions. Moreover, there are differences in the quantity of ideas received, our study has significantly lower volume of proposals received by the evaluators. Furthermore, conceptually, Piezunka and Dahlander (2015) test whether distant suggestions receive attention. Here we specifically develop and test whether proposals building upon distant knowledge, novel or atypical combinantions, and diverse or broad recombinations, as conceptualized in the literature, are associated with successful evaluation.

On the one hand, one would expect successful proposals to break away from cognitive frames built upon existing knowledge as this may filter out knowledge from peripheral areas (Kaplan and Tripsas 2008, Tripsas and Gavetti 2000). Therefore, drawing from distant and diverse knowledge can lead to a useful solution. Jeppesen and Lakhani (2010) studied crowdsourced scientific problem-solving on

the Innocentive platform (an intermediary between firms and solvers) and found that successful solutions were more likely to be generated by solvers who were technically marginal in terms of expertise to the focal field of the problem. However, solutions drawing from diverse domains may suffer at the hands of institutional gatekeepers and market audiences who could face cognitive hurdles in categorizing (Kovács and Hannan 2010) and evaluating the value of such offerings and the practical usefulness of the idea or feasibility of implementation.

On the other hand, a deep understanding of the problem, the existing approaches used to solve the problem and their fundamental limitations is also important. According to the foundational perspective of creativity (Weisberg 1999), deep knowledge and immersion in a domain are required to produce novelty (Csikszentmihalyi 1997). Narrower recombinations of knowledge enable the identification of anomalies that can lead to breakthroughs by exposing challenges in the existing ways of approaching problems. Kaplan and Vakili (2015) refer to this as "cognitive novelty" and find cognitive novelty to be associated with patents that make local search, while economic value to the product of wide recombinations as well as novelty. This perspective of creativity is also consistent with the view that with an increasing burden of knowledge, specialization is needed to achieve breakthroughs (Jones 2009, Agrawal et al. 2013). Therefore, according to this foundational view, narrow recombinations or local search combining familiar knowledge could lead to useful solutions.

Our work draws from the literature on search, crowdsourcing, and knowledge recombination and makes two main contributions. First, we test the aforementioned effects of distant and recombinant search and combinatorial novelty on the successful evaluation of proposed solutions in a novel crowdsourcing context. Second, methodologically, we use natural language processing to operationalize measures of distant and recombinant search, and atypicality.

The rest of the chapter is organized as follows: Section 4.1.1 and 4.1.2 reviews the empirical literature on crowdsourcing and knowledge recombination. Section 4.2 follows with an overview of the empirical context and the data. In Section 4.3,

we describe the results of the analyses. Section 4.4 concludes with a discussion.

4.1.1 Crowdsourcing for Distant Search

In the search for innovative ideas and solutions, organizations turn to external crowds, engaging in "distant" search (Afuah and Tucci 2012, Chesbrough et al. 2003), commonly now known as crowdsourcing. The use of crowdsourcing to find solutions to problems dates back in history to famous examples such as the problem of finding longitude at sea for which the British Government set the Longitude Prize in 1714. Recent advances in Information Technology have facilitated engagement with external crowds through dedicated digital crowdsourcing platforms[2]. A comprehensive taxonomy of crowdsourcing platforms can be found in Afuah (2018).

In the recent times, crowdsourcing has been leveraged in a number of different areas such as innovation management (Poetz and Schreier 2012), software development (such as TopCoder), and humanitarian aid (Rogstadius et al. 2013). Crowdsourcing has also been leveraged to gather solutions to complex social problems through online, distributed problem-solving. As our empirical context, we use data from an open innovation platform which uses design thinking to address complex social problems that require unique approaches and ideas from multiple domains combined with analogical problem solving.

Previous work on crowdsourcing has primarily focused on the potential of crowdsourcing (such as access to knowledge, talent and expertise that are distant relative to the organization's current knowledge base) and ways to solicit solutions from crowds (Afuah and Tucci 2012, Jeppesen and Lakhani 2010, Boudreau et al. 2011, Boudreau and Lakhani 2013, Dahlander and Piezunka 2014). Organizations using crowdsourcing also benefit from access to large numbers of solutions, which increases the likelihood of finding a successful solution (Kornish and Ulrich 2011, Girotra et al. 2010). Contributor characteristics and the effect of past participation on successful idea generation have also been studied in crowdsourcing contexts. For example, Bayus (2013) finds that serial contributors are more likely to generate

[2]Popular contemporary crowdsourcing platforms that have been studied in the literature include Innocentive for scientific problem solving and TopCoder for competitive software programming etc.

ideas that get implemented by organizations. Huang et al. (2014) find that contributors with greater participation learn to develop promising ideas. Rhyn et al. (2017) examine the effect of network relationships among crowds on generating useful crowdsourced contributions. Chan and Schunn (2015) show that iteration is necessary to convert distant combinations into creative concepts.

Distant search through crowdsourcing offers the potential for externalizing problem solving and idea generation to crowds in an effective manner. The literature on creativity and innovation suggests that distant combinations are likely to lead to creative solutions than near ones. However, organizations are faced with the challenge of sifting through and filtering large volumes of information generated by crowds. For example, since Dell launched its ideation platform, the IdeaStorm [3], it has generated more than 28,000 ideas and over 100,000 comments since its inception in 2007. Out of these, 550+ user contributed ideas have been implemented. British Petroleum (BP) received over 100,000 suggestions to its open call to the general public asking after the Deepwater Horizon accident (Mascarelli 2010). Piezunka and Dahlander (2015) report on how 922 organizations responded to 100,000+ crowdsourced suggestions from external contributors and show that when faced with a large number of ideas, organizations face crowding and pay attention to familiar ideas rather than ideas that capture distant knowledge. Therefore, despite getting access to distant knowledge, in the filtering stage institutional gatekeepers appear to struggle to make use of the knowledge they gained through crowdsourcing. Organizations can also be path-dependent in making decisions and may be tuned to building on familiar knowledge (Cohen and Levinthal 1990).

In the absence of crowding, do organizations positively evaluate proposed solutions building upon distant knowledge? This is particularly relevant for broadcast search problems that require fresh perspectives unfamiliar to the organization seeking solutions. We examine this question in our study.

4.1.2 Knowledge Recombination, Novelty and its evaluation

Nelson and Winter (1982) state "the creation of any sort of novelty in art, science of

[3] `https://dellideas.secure.force.com/`

practical life- consists to a substantial extent of a recombination of conceptual and physical materials that were previously in existence". The predominant view in the innovation literature is that broad search and knowledge recombination forms the basis for innovative breakthroughs (which to a large extent also generate economic value) (Rosenkopf and Nerkar 2001, Fleming and Sorenson 2004, Fleming 2001, Kaplan and Vakili 2015, Huang et al. 2009, Gruber et al. 2013) as it bridges distance and diverse domains, thereby preventing knowledge lock-in and "habit forming" (Weisberg 1999).

Highly cited patents are found to more likely be combinations of diverse knowledge domains (Hall et al. 2001). Moreover, recombination with distant knowledge is more likely to lead to higher citation rates as this involves exploratory or long jump search (March 1991, Gavetti and Levinthal 2000). In the innovation literature these perspectives of knowledge recombination have largely developed in the empirical context of patents and scientific publications, focusing on technological innovation and scientific invention.

The recombination of ideas and knowledge forms the basis for innovation (Fleming 2001) and impactful inventions (Rosenkopf and Nerkar 2001, Fleming and Sorenson 2004, Sutton and Hargadon 1996). Novel research involves an explorative search process (March 1991) that recombines existing knowledge in new ways. This combinatorial view of novelty has largely found empirical support based on patent studies. A parallel body of work examining academic research (Uzzi et al. 2013) finds that atypical combinations are characteristic of extraordinary publications and that papers which combine high novelty with conventionality are more likely to receive higher citations.

In the innovation literature, the positive relationship between knowledge recombination is vis-a-vis the typicality of combinations of knowledge components (Fleming 2001, Uzzi et al. 2013, Mukherjee et al. 2016) (conceived in terms of mean and variance effects) and the combinatorial breadth of combinations (Fleming 2001, Fleming and Sorenson 2004, Rosenkopf and Nerkar 2001) conceived in terms of mean effects (Lerner 1994) and independent of combination typicality. An

established finding in innovation scholarship is that inventions that combine knowledge across diverse knowledge domains are more impactful (Rosenkopf and Nerkar 2001, Fleming and Sorenson 2004, Fleming 2001, Kaplan and Vakili 2015, Huang et al. 2009, Gruber et al. 2013).

A growing stream of literature highlights the role of institutional gatekeepers in evaluating novelty and recombinant knowledge breadth in various settings. Institutional gatekeepers play an important role in assessing, filtering and enforcing selection criteria. For example, patent offices act as institutional gatekeepers in determining the inventiveness and novelty of a patent application compared to prior art. In the case of scientific publications, journals act as gatekeepers for publication quality. For a product to compete in the market, it has to be viewed as a player and conform to the product categories in which it competes. Zuckerman (1999) shows that gaining this recognition has a bearing on a firms performance in mediated markets such as the financial markets.

While gatekeeping institutions strive to maintain objectivity in their evaluation process, they can be influenced by categorical structures in their decision making. For example, Ferguson and Carnabuci (2017) show that patent applications that span more domains are less likely to be approved by examiners due in part to organizational and cognitive hurdles involved in the appraisal of such applications.

Funding agencies, in their evaluation of scientific proposals, are less likely to positively evaluate proposals spanning disciplinary boundaries that are well established (Porter and Rossini 1985). Likewise, Boudreau et al. (2016) provide evidence from a randomized controlled trial highlighting that novel scientific research proposals that depart from existing literature are scored lower by evaluators.

In R&D projects, managers and stakeholders are found to face the challenge of evaluating novelty. While some degree of novelty is warranted in developing new products and services (Knudsen and Levinthal 2007), projects involving too much novelty can be difficult for evaluators to categorize and interpret (Rindova and Petkova 2007).

In the creativity literature, high levels of creativity are found to emerge from

the combination of different schematas and cognitive structures (Mumford and Gustafson 1988). Moreover, perspectives from distinct fields upon combination have the potential to generate novel and deviant ideas (Perry-Smith and Shalley 2003). On the other hand, lab experiments in creativity show a bias against creativity by those who espouse it, and more so under uncertainty (Mueller et al. 2012).

Taken together, the extant findings seem to suggest that while broad knowledge recombinations have the potential to generate novel solutions, too much spanning across fields increases the likelihood of creating cognitive difficulties for evaluators. Furthermore, while some novelty is needed to break away from the past (as shown with products and services by Rosenkopf and Nerkar (2001)), too much novelty can be difficult to process, and therefore a blend of conventionality and atypicality may be optimal (Uzzi et al. 2013). Therefore, we may expect to find a positive relationship between recombinant breadth and likelihood of success, and likewise for atypicality by the same logic.

In our empirical setting, solutions are solicited for existing complex social problems. A successful proposal ought to be unique to the context in which it gets proposed, i.e. a fresh approach to approaching the challenge (novel in that sense). Moreover, the proposed solution has to be implementable and not merely some futuristic idea or vision. Here, we test whether solutions spanning diverse knowledge domains or solutions building upon deep domain knowledge are more likely to be positively evaluated.

4.2 Empirical Context

Our empirical context is a Web-based crowdsourcing platform called OpenIDEO which brings together a world-wide community of contributors to develop solutions for a range of complex social problems requiring multi-disciplinary and analogical approaches through design-thinking. Till date, over 60 social challenges elicited by organizations have been tackled through this platform's world-wide community of over 100,000 users. We obtained the data through a third-party platform. Problems for which solutions are sought from the community range from managing electronic

waste to addressing sanitation issues and children's education in third-world countries. Our sample consists of the entire set of problems and proposals posted on the platform from its launch in 2010 until April 2017.

Figure 4.1 showcases a problem hosted by Oxfam and Nokia seeking solutions for improving maternal health with mobile technologies in low-income countries. Figures 4.2 and 4.3 showcase proposals for this problem and one from a problem on managing e-waste.

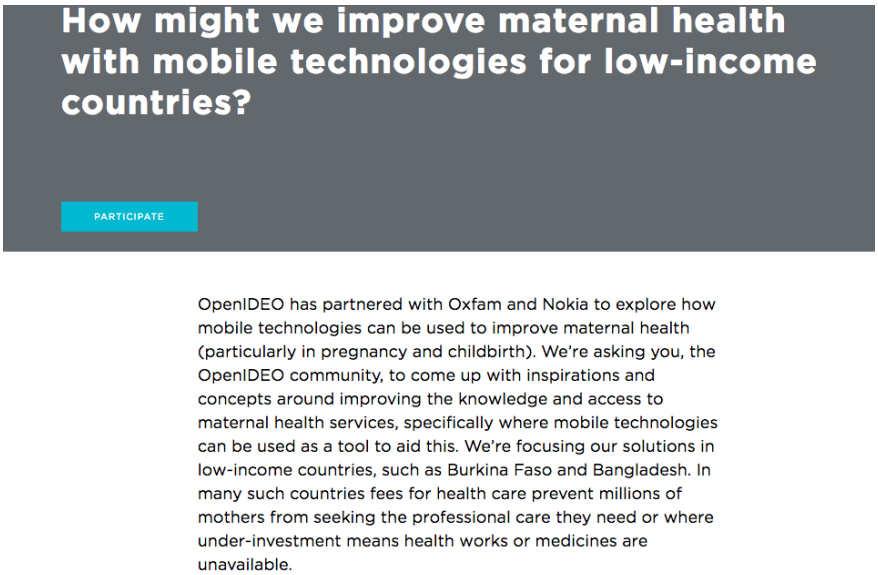


Figure 4.1: Example of a problem hosted by Oxfam and Nokia

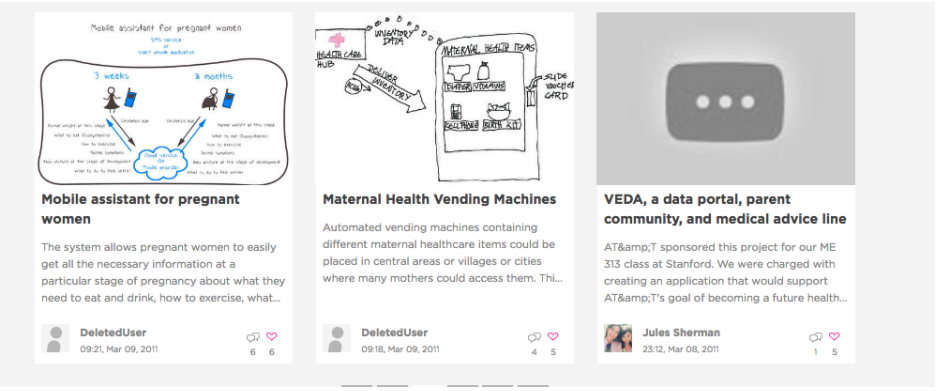


Figure 4.2: Proposals for the maternal health-mobile tech problem

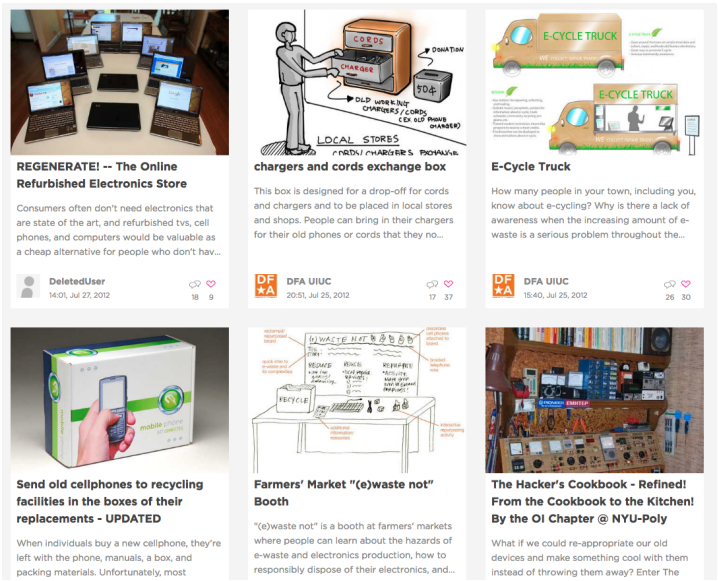


Figure 4.3: Proposals for an e-waste management problem

The platform follows a staged design process for solution generation. A host organization posts the problem in the form of a text description and defines its scope. Following this, an *Inspiration* phase which lasts anywhere from 2 to 4 weeks involves a brainstorming session wherein members share with the community analogous problems and anecdotes from previous experience, case studies, interviews etc. In the *Concepting* phase that follows, members submit text-based proposals to the posted problem. This is followed by the *Shortlist* phase in which an expert panel (consisting of experts from the platform and the stakeholders) shortlist a subset of proposals for further consideration. Shortlisted proposals are further refined with community feedback and a subset of these is chosen for implementation in the field. We focus on the shortlisting phase. In evaluating proposals, the criteria is to seek proposals that use a "fresh" approach to tackling the problem and ones that are unique to the context. The evaluation panel also considers whether the proposals are implementable and not merely some futuristic idea or vision for solving the problem.

4.2.1 Topic Modeling

We take a computational approach to measure the recombinant breadth, content distance and atypicality of proposals using the unstructured text in the proposals. We leverage Latent Dirichlet Allocation (LDA) (Blei et al. 2003), also known as topic modeling, which is an unsupervised machine-learning algorithm to learn the latent "topics" or themes in the unstructured text. Topic modeling is premised on the idea that documents are mixtures of topics, where a topic is a probability distribution over words. Topic modeling is increasingly gaining popularity in several contexts including Management (see Aksuyek et al. (2013), Kaplan and Vakili (2015), Huang et al. (2017), Jung and Lee (2016)) as a means to explore and analyze large collections of unstructured text documents. Here, the key idea is to learn the semantic space of all text-based proposals and problem statements in the dataset, and construct measures of interest through this space.

Drawing parallels from the classification of patents into patent classes, topics generated through LDA are akin to patent classes created in patent offices since both are generated from empirical data (Aharonson and Schilling 2016). While patent classes are created by patent offices, topics are generated by a learning algorithm. Fundamentally, in both cases, the idea is to capture latent themes in the underlying data. Therefore, applied to proposals in our context, topic modeling captures the language used to describe concepts in different knowledge domains. We built the topic model using Mallet (McCallum 2002), a software package developed in Java using 200 and 400 topics.

4.2.2 Data and Variables

The dataset consists of 10,675 proposals submitted to 42 problems posted on the platform by a variety of organizations and institutions. The mean number of submissions per problem is 241, out of which on an average 33 ideas are shortlisted by the evaluators to progress to the next phase for further refinement. Finally, per problem only 8 ideas on average get implemented in the field.

Dependent Variable

Our dependent variable is a binary variable, *Success* indicating whether the pro-

Table 4.1: Variable definitions

	Description
Dependent Variable	
Success	Indicator for whether a proposal gets shortlisted or not
Explanatory Variables	
RecombinantBreadth	Measure of the breadth of the topics in the proposal
Content Distance	Measure of how distant the proposal is from the problem domain
Atypicality	Measure of the rareness of the topic combinations occurring in a proposal
Control Variables	
ContributorExperience	Number of past proposals submitted by the contributor
SuccessRatio	Ratio of past shortlisted proposals to total proposals submitted by contributor
NumCitations	Number of inspirations the proposal cites

posed idea is successfully shortlisted through evaluation (i.e. it is judged to be useful and has the potential to get implemented).

Control Measures: We account for potential factors that may influence a proposal's success. These include contributor level controls such as the contributor's experience in participating in past problems, *ContributorExperience*, and the ratio of past proposal success to number of proposals submitted, given by *SuccessRatio*. We also control for the number of inspirations cited by the contributor, *NumCitations* as drawing upon these may improve the quality of the idea.

Independent Variables

Variables definitions are given in Table 4.1. The key independent variables are operationalized below.

Content Distance: This is a measure of the extent to which the proposal is distant from the problem domain (reflecting distant search). The problem itself is a text-based description characterizing the problem domain. Each proposal and each challenge is represented as a vector in a K-dimensional space (where K is the number of topics in the entire topic-space) where the weight of each dimension is the topic weight.

We use a Cosine similarity measure between the vector representations of the topic weights of the proposal and the problem to compute the semantic content similarity between the proposal and the problem domain. We obtain the cosine distance by subtracting the cosine similarity from 1.

Recombinant Breadth: We operationalize the recombinant breadth of the

proposal as the variety of the proposal's constituent topics. Recall that each proposal is essentially a vector of topic weights. We use an entropy based measure of diversity (Harrison and Klein 2007) given in Equation 4.1 to calculate *Recombinant Breadth*.

$$\text{RecombinantBreadth} = -\sum_{i=1}^{K} p_i \times \ln(p_i) \tag{4.1}$$

where p_i is the weight of topic i and K is the total number of clusters or topics. Proposals in which topics are spread evenly across a greater number of topics will have a higher value of *Recombinant Breadth* [4].

Atypicality of Combinations: This is a measure of the novelty of the proposal in terms of its constituent knowledge components. There are several ways in which atypicality can be operationalized. Similar to the approach used in Ferguson and Carnabuci (2017), (where combinations of four digit IPC patent classes are used to derive atypicality) we measure atypicality of the topic combinations of a proposal using the Jaccard index, a widely used measure of similarity between sets. For each proposal, we compute rthe Jaccard index for each pairwise topic combination of its top-5 highest weighted topics, as given in Equation 4.2. Here, X is a set of topic combinations across all proposals that have topic X and Y is a set of topic combinations across all proposals for a given problem that have topic Y. We then calculate the minimum Jaccard for the proposal from among all the topic combinations. A lower minimum Jaccard coefficient is indicative of the atypicality of topic combinations.

$$\text{Jaccard}(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \tag{4.2}$$

where X is a set of combinations with topic X and Y is a set of combinations with topic Y.

[4] We run the the topic model with a range of topics and find K=200 topics to yield meaningful and granular topics. Our results are robust across a range of topics above K = 200, we run our analysis with K = 300 and K= 400 topics

4.2.3 Estimation approach

In the data we find contributors to have submitted multiple proposals across problems. In other words, contributions are nested within contributors and problems. This violates the independence assumption as proposals submitted by the same contributor or those proposals submitted within a particular problem may be correlated. Since failing to account for this non-independence can lead to biased estimates, we address this issue by incorporating a generalized linear mixed model (GLMM) (Verbeke and Molenberghs 2009). A mixed effects model decomposes covariates into cross-sectional variation and longitudinal variation, and is used to resolve non-independencies in the data. The key advantage of using a mixed effects model over a fixed effects or random effects is that it can help reduce the variability in the random error. The random effects essentially give structure to the error term, representing idiosyncratic variation due to contributor differences and differences across problems. We add the fixed effects for our controls and our econometric specification focuses on identifying the effect of recombinant breadth, content distance and atypicality on the likelihood of proposal success. One key assumption of the mixed effects model is that the effects associated with the random variable groups be uncorrelated with the means of the fixed effect from the random variable groups. Other assumptions are similar to the underlying linear models (normally distributed errors, linearity and so on). We use the lme4 package (Bates et al. 2014) in R with maxmium likelihood estimation to fit our mixed model.

4.3 Results

Descriptive Statistics are summarized in Table 4.2. The correlations matrix is given in Table 4.3.

In terms of modeling strategy, we begin by fitting a simplified model with random intercepts for the contributor and the problem and a set of controls. We fit a series of generalized linear mixed models, adding fixed effects one by one. For model fit, we choose the model that reduces the deviance and has a lower value for Akaike Information Criterion (AIC) compared to the previous model. In R, the

Table 4.2: Descriptive Statistics

Variable	N	Mean	St. Dev.	Min	Max
RecombinantBreadth	10,675	4.562	0.574	0.392	6.789
ContentDistance	10,675	0.695	0.204	0.050	0.999
Atypicality	10,675	0.007	0.003	0.003	0.091
SuccessRatio	10,675	0.020	0.076	0	1
ContributorExperience	10,675	5.749	13.829	1	109
NumCitations	10,675	1.197	5.489	0	295

Table 4.3: Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Success	1	0.011	-0.062	0.049	0.081	0.003	0.078
RecombinantBreadth	0.011	1	0.171	0.061	-0.012	-0.038	-0.020
ContentDistance	-0.062	0.171	1	0.290	-0.036	-0.044	-0.063
Atypicality	0.049	0.061	0.290	1	0.055	0.049	0.028
SuccessRatio	0.081	-0.012	-0.036	0.055	1	0.141	0.125
ContributorExperience	0.003	-0.038	-0.044	0.049	0.141	1	0.190
NumCitations	0.078	-0.020	-0.063	0.028	0.125	0.190	1

likelihood ratio test is carried through the anova command.

Table 4.4 presents model estimates and fitness statistics in Column 1 for a baseline model with control variables only and random effects for the problem and contributor. All our variables have been re-scaled and standardized.

Next, we examine the relationship between a proposal's Recombinant breadth and its likelihood of being shortlisted by adding a fixed slope for proposal in Column 2. We observe a significant negative effect of *RecombinantBreadth* on the likelihood of successful evaluation (β= -0.135, p< 0.01). We also find a reduction in deviance compared to the baseline model, $\chi^2_1 = 11.5, p = 0.00$[5]. In terms of odds ratios, a one standard deviation increase in *RecombinantBreadth* results in a roughly 10 to 12 percent decrease in the odds of proposal success. Studies involving gatekeeping settings have considered breadth and typicality together (Kovács and

[5]We also add a random slope model to model potential effects of problem on *Recombinant-Breadth*. For example, some problems may be more or less disposed to diverse ideas. Although this random-slope model in terms of deviance is superior to the model we estimate in Column 2, on adding the other independent variables, the models with the random slope fail to converge. Therefore, we proceed with random intercept models for the rest of the analysis

Hannan 2010), therefore, we check whether the proposed effects of recombinant breadth hold after controlling for atypicality. In Column 4, to examine the effect of proposal atypicality or novelty, we add a term for *Atypicality* and find no significant effect of atypicality on successful proposal evaluation.

In Column 3 we present results with *ContentDistance*. We find a significant effect of content distance on the likelihood of proposal shortlisting success (β=-0.265, p<0.01). A one-standard deviation increase in *ContentDistance* is associated with a 23 percent decrease in the odds of proposal success.

In Column 5, we examine any potential moderating effect of proposal *Atypicality* on the relationship between *RecombinantBreadth* and success. It may be that proposals that draw upon diverse knowledge domains and are novel are rewarded differentially higher, or that proposals invoking diverse knowledge domains but are more typical are rewarded more. To test this, we add an interaction term in Column 5. However, we do not find a significant interaction between *Atypicality* and *RecombinantBreadth*.

To test the sensitivity of our results to greater number of topics (which would make topics more granular), for robustness we report results with K= 400 topics in Table 4.5. We largely find our earlier results with 200 topics to hold with 400 topics. That is, we find a negative and significant effect of recombinant breadth and content distance on the probability of proposal getting shortlisted and no effect of proposal atypicality.

Furthermore, we construct an alternate measure of atypicality as in Wang et al. (2017) who measure novelty in terms of whether a published academic paper makes first-time ever combinations of referenced journals. Applying the same to our data, we operationalize atypicality in terms of rare topic combinations occurring in a proposal. For each proposal, we pair up the top-5 most weighted topics (e.g. if the top-most weighted topics in a proposal are T_1, T_2, T_{10}, T_{20} and T_{29}, the combinations would be T_1-T_2, T_1-T_{10}, T_1-T_{20}, T_1-T_{29}, T_2-T_{10}...and so on)

For each topic pair over all proposals submitted for a problem, we compare the co-occurrences of the topic pair by forming a co-occurrence matrix for all the

Table 4.4: Mixed Model estimates with logistic regressions (with 200 topics)

	Dependent variable:				
	Probability of getting shortlisted				
	(1)	(2)	(3)	(4)	(5)
Fixed Effects					
RecombinantBreadth		-0.138***	-0.112***	-0.113***	-0.109***
		(0.039)	(0.040)	(0.040)	(0.040)
ContentDistance			-0.269***	-0.278***	-0.280***
			(0.038)	(0.040)	(0.040)
Atypicality				0.028	0.031
				(0.037)	(0.037)
NumCitations	0.243***	0.243***	0.229***	0.229***	0.229***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
RecombinantBreadthXAtypicality					-0.016
					(0.025)
γ_{00} intercept	-2.130***	-2.132***	-2.146***	-2.157***	-2.157***
	(0.117)	(0.120)	(0.123)	(0.123)	(0.122)
Random Effects					
Contributor	0.4256	0.4187	0.3970	0.3978	0.3977
Problem	0.4721	0.5033	0.5409	0.5246	0.5210
Observations	10,675	10,675	10,675	10,675	10,675
Log Likelihood	-3582.5	-3576.7	-3552.7	-3552.4	-3552.2
Akaike Inf. Crit.	7173.1	7163.3	7117.4	7118.8	7120.4
Bayesian Inf. Crit.	7202.2	7199.7	7161.1	7169.8	7178.6
Note:					*p<0.1; **p<0.05; ***p<0.01

proposals in the problem. The i,jjth element in this co-occurrence matrix is the number of times topic i and topic j co-occur. The ease of combining topic i and topic j is given by the cosine similarity between the co-occurrence profiles T_i and T_j:

$$\text{Cosine}_{\text{T}_\text{i},\text{T}_\text{j}} = \frac{\text{T}_\text{i} \cdot \text{T}_\text{j}}{||\text{T}_\text{i}||||\text{T}_\text{j}||} \tag{4.3}$$

where T_i and T_j are row vectors of the co-occurrence matrix. For each proposal, the novelty of the proposal is then given as the sum of the cosine distance of the combinations:

$$\text{Novelty} = \sum_{\text{i,j}} 1 - \text{Cosine}(\text{T}_\text{i}, \text{T}_\text{j}) \tag{4.4}$$

Table 4.5: Mixed Model estimates with logistic regressions (with 400 topics)

	Dependent variable:					
	Probability of getting shortlisted					
	(1)	(2)	(3)	(4)	(5)	(6)
Fixed Effects						
RecombinantBreadth		-0.324***	-0.302***	-0.301***	-0.304***	-0.296***
		(0.041)	(0.042)	(0.042)	(0.043)	(0.043)
ContentDistance			-0.265***	-0.255***	-0.255***	-0.259***
			(0.037)	(0.040)	(0.040)	(0.038)
Atypicality				-0.027	-0.027	
				(0.040)	(0.040)	
NumCitations	0.243***	0.239***	0.227***	0.227***	0.227***	0.229***
	(0.042)	(0.041)	(0.041)	(0.041)	(0.041)	(0.042)
RecombinantBreadthXAtypicality					0.008	
					(0.028)	
AtypicalityAlternate						0.059
						(0.045)
γ_{00} intercept	-2.130***	-2.151***	-2.172***	-2.163***	-2.164***	-2.159***
	(0.117)	(0.126)	(0.130)	(0.132)	(0.132)	(0.133)
Random Effects						
Contributor	0.4256	0.4222	0.3982	0.3972	0.3972	0.3952
Problem	0.4721	0.5689	0.6131	0.6274	0.6285	0.6359
Observations	10,675	10,675	10,675	10,675	10,675	10,675
Log Likelihood	-3582.5	-3553.3	-3529.7	-3529.5	-3529.4	-3528.9
Akaike Inf. Crit.	7173.1	7116.5	7071.4	7072.9	7074.8	7071.7
Bayesian Inf. Crit.	7202.2	7152.9	7115.0	7123.9	7133.1	7122.7
Note:				*p<0.1; **p<0.05; ***p<0.01		

We report results with this alternate measure, *AtypicalityAlternate*, in Table 4.5, Column 6. We find the estimate of this alternate measure to also be insignificant, consistent with our estimate for *Atypicality*.

Our data has repeated cross sections with multiple observations per participant. Since we do not observe participants through all the challenges that are hosted, the structure of our data is akin to an unbalanced panel. Given the data structure and for robustness, we also use an alternate estimation approach to account for unobserved heterogeneity due to contestant characteristics by incorporating a fixed effects logit model (Kennedy 2003, Maddala and Lahiri 1992) in Table 4.6. Here, the controls include contributor attributes and challenge-level attributes. A caveat, however, is that using contributor-level fixed effects effectively means each contributor appears

in the dataset with at least one successful and one unsuccessful proposal. However, since a very small percentage of contributors have successful proposals, the dataset essentially gets reduced to 404 contributors across 2136 challenge-contributor observations (for the Mixed effects model we were able to use all 10675 observations across 6662 proposal contributors). In Column 2, we do not find significance for any of the measures of interest. This may largely be attributed to the curtailed sample.

We also test our models with a random effects model and report the results in Table 4.6, Columns 3-4. In Column 4, we find significance for *ContentDistance* and *RecombinantBreadth* and do not find significance for *Atypicality*, consistent with the earlier results.

Table 4.6: Fixed and random effects logistic regressions (with 200 topics)

	Dependent variable:			
	Probability of getting shortlisted			
	(1)	(2)	(3)	(4)
TotalSubmissionsInChallenge	-0.5349***	-0.5211***	-0.4047***	-0.3839***
	(0.0960)	(0.1019)	(0.0379)	(0.0418)
SuccessRatio	-0.4999***	-0.4974***	0.0174	0.0286
	(0.0475)	(0.0476)	(0.0433)	(0.0427)
NumCitations	0.1291***	0.1248***	0.1739***	0.1628***
	(0.0437)	(0.0438)	(0.0405)	(0.0403)
RecombinantBreadth		0.1343		-0.0737**
		(0.0890)		(0.0352)
Atypicality		0.0216		0.0327
		(0.0576)		(0.0359)
ContentDistance		-0.1048		-0.2039***
		(0.0739)		(0.0364)
ContributorExperience			-0.0974	-0.1086
			(0.0789)	(0.0750)
Observations	2136	2136	10675	10675
Pseudo R-squared	0.17	0.17		
Note:				*p<0.1; **p<0.05; ***p<0.01

4.4 Discussion and Conclusion

Our results help inform our understanding of factors that drive successful evaluation of ideas by gatekeepers in crowdsourcing. We find that proposals that use greater recombinant breadth are less likely to succeed and we also find a negative relationship between proposal distance (building upon distant knowledge relative to the problem domain) and probability of success. We find no support for the role of atypical combinations in predicting proposal success.

Contrary to the tension view in creativity which finds that distant and diverse recombinations are associated with breakthroughs, our findings suggest that a narrow focus with deep knowledge is beneficial for the successful evaluation of proposals to complex, social problems that require narrow recombinations of deeper knowledge. Our findings are consistent with (Kaplan and Vakili 2015) who find that narrow and local search lead to "cognitive novelty". A caveat here is that though we are not claiming this leads to better solutions per se, as we are looking at successful evaluation in the first round of a staged evaluation process, what our findings show is that domain experts prefer solutions with leverage narrow/deep knowledge rather than distant and diverse recombinations.

It is unclear whether the evaluators' preferences are due to cognitive hurdles encountered in the evaluation of proposals involving distant and atypical combinations or whether it is the case that proposals involving distant and atypical combinations are in general infeasible and hold little practical value. Future research could explore which of these possibilities is indeed at work. If the former is the case then our findings pave a way for managers to effectively assess the potential of distant and atypical combinations could be useful to fully harness the power of crowdsourcing. If it is the case that distant and atypical combination proposals are seen as impractical and infeasible for certain categories of problems, then perhaps reorienting the crowdsourcing efforts to garner solutions that involve local and narrow search may be helpful to managers

Though the findings need to be interpreted with some caution, our analysis shows the benefits of narrow and focused combinations of knowledge in the evaluation of crowdsourced ideas. Our contrasting results may be attributed to differences in evaluation criteria and the kinds of knowledge required to solve particular kinds of problems. While inventiveness is the key criteria for patent examiners, in contexts such as ours, where evaluators may care less about novelty of combinations per se and more about feasibility, proposals drawing from diverse knowledge domains may face greater challenges during evaluation in relation to their categorization and assessment of potential value, as described in previous studies.

Since proposals build upon domain knowledge from a variety of domains such as education, health care, social sciences, HCI, public policy etc., our results may generalize to settings such as design, and other crowdsourcing settings.

One limitation of this work is that here is that we measure success in terms of successful first round evaluation, and do not focus on the ideas that make it all the way till the end. Such staged evaluation is also seen in entrepreneurial venture capital funding (also seen in crowdfunding these days) where the goal is to attain a level of funding and then proceed for the next level. What is key is that while these staged rounds are crucial, a lot of focus is on whether ideas actually achieve market success. In our case, further research to understand which ideas get implemented in the end (and not just which ones make it through the first round of filtering in evaluation) and are successful could pave the way for more research. This could involve closely working with platform providers and challenge hosts to get access to this sort of data.

4.4.1 Future work

Our results show that field experts prefer solutions which are built upon narrower recombinations of knowledge and local search. However, understanding whether such proposals truly lead to better solutions could be a fruitful line of future research. There are several contexts in which field experts play an important role as gatekeepers. Exploring the degree to which their biases and preferences play an role in filtering in/out ideas or proposals has important implications.

Moreover, while we did not study this specifically, it would be useful to understand which kinds of problems lend themselves to solutions garnering deep knowledge along with narrow recombinations of peripheral knowledge and which problems are better suited for broad recombinations across knowledge domains. Another practical question pertinent to managing collaborative crowdsourcing is how to identify and involve deep domain knowledge experts (specialists), those who can bring broad peripheral knowledge (generalists), and finally contributors who are better suited at recombining disparate knowledge.

Our analysis accounts for variation in problem and the effect of contributor

variation, thus overcoming the non-independence problem that may undermine estimation with a random or fixed effects model. However, we are unable to make strong causal claims in the relationships between our key independent variables and the likelihood of proposal success. For future work, research designs that capture causal mechanisms may be leveraged. Moreover, we focus on two characterizations of atypicality or combinatorial novelty. There are several dimensions of this construct. Exploring and capturing other dimensions of atypicality may also be useful.

External crowds are increasingly gaining prevalence in making decisions with respect to whether or not products, ideas etc are to be funded. This shifting of power from the experts to crowds in areas as diverse as arts, gaming, technology etc. raises new managerial questions. Mollick and Nanda (2015) find no significant difference between funding decisions of experts vs crowds on a crowdfunding platform. However, little is known about whether or not crowds and experts in these new settings differ in their evaluation of products or ideas that draw from diverse domains. This work can be extended using novel experimental designs to study such questions.

Organizations and crowdsourcing intermediaries are faced with the formidable challenge of filtering ideas and solutions generated by crowds. We have constructed topic-modeling based measures of a proposal's recombinant breadth and search distance to study their effect on proposal's successful evaluation by field experts. A future line of enquiry could involve using more advanced computational techniques such as machine learning algorithms to learn which types of ideas organizations could pay greater attention to beyond the ones that evaluators would be inclined towards (i.e. the solutions that are missed due to inherent biases in the evaluation process). Lastly, while topic modeling is a promising method to deal with large amounts of unstructured data, for future analysis, topics may be validated with the help of experts.

Chapter 5

Concluding Remarks

In this thesis, I provide an empirical basis for understanding the operational drivers of innovation management in digital platforms. In Chapter 2, I study how firms successfully manage innovation in mobile app development. In Chapter 3, I turned the focus to the important question of user participation in crowdsourcing, and through a unique randomized controlled trial on a tech-based crowdsourcing platform, I study gender-based preferences for tech work, showing our results to be consistent with the leaky pipeline of supply of female workers. In Chapter 4, I investigate whether successful contributions in crowdsourcing are associated with established conceptualizations of knowledge recombination.

Each chapter motivates a range of future research questions. On a methodological level, my thesis highlights a means for quantifying and characterizing innovation as it takes place through product evolution in mobile apps, and knowledge recombination in crowdsourcing initiatives using text mining and information retrieval. Recent advances in Machine Learning and an abundance of data offer exciting opportunities to employ such techniques to study various facets of the innovation process in settings affected by digital transformation.

A meaningful extension of the work in Chapter 2 would involve developing a deeper understanding of the role of customers (based on customer feedback and reviews) in influencing search strategies and how incorporating customer feedback impacts firm performance in such settings. Here again, computational methods applied to text-based customer reviews can be leveraged to study this topic. Further-

more, while our study focuses on the evolution of a single product, we know little about the circumstances under which firms diversify their product portfolios and launch other products and how their decisions affect their survival in these hyper-competitive settings.

Devising effective interventions to increase participation of marginalized users in crowdsourcing endeavours would be an natural extension of Chapter 3 which also has practical implications. Closely related to this is the topic of diversity in team-based crowdsourcing work.

Building on Chapter 4, a relevant path for exploration for future research could involve understanding how crowdsourcing communities form and evolve, and how to organize actors along the different stages of the innovation process (from ideation to commercialization) in open and distributed innovation. Moreover, in crowdsourcing initiatives, it would also be worthwhile to analyze which types of ideas are useful to organizations and whether computational methods to filter and sift through large number of crowdsourced ideas can be deployed to identify useful ideas. Redundancy is a recurring problem in crowdsourced idea generation. Most of the crowdsourcing idea generation platforms today focus on idea generation rather than recombination. How can crowdsourcing be organized to effectively capture and represent collectively generated knowledge to improve innovation outcomes? Overall, I hope that this thesis provides the ground for fruitful future research in the areas discussed.

Bibliography

Acemoglu, Daron, Simon Johnson, James A Robinson, Pierre Yared. 2008. Income and democracy. *American Economic Review* **98**(3) 808–42.

Afifi, A., V. Clark. 1999. *Computer-Aided Multivariate Analysis*. 2nd ed. Chapman & Hall, Ltd., London, UK.

Afuah, Allan. 2018. Creating and capturing value through crowdsourcing .

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

Agrawal, Ajay, Avi Goldfarb, Florenta Teodoridis. 2013. Does knowledge accumulation increase the returns to collaboration? Tech. rep., National Bureau of Economic Research.

Aharonson, Barak S, Melissa A Schilling. 2016. Mapping the technological landscape: Measuring technology distance, technological footprints, and technology evolution. *Research Policy* **45**(1) 81–96.

Ahuja, G., R. Katila. 2004. Where do resources come from? the role of idiosyncratic situations. *Strategic Management Journal* **25**(8-9) 887–907.

Ahuja, G., C.M. Lampert. 2001. Entrepreneurship in the large corporation: a longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal* **22**(6-7) 521–543.

Aigner, Dennis J, Glen G Cain. 1977. Statistical theories of discrimination in labor markets. *ILR Review* **30**(2) 175–187.

Akerlof, George A, Rachel E Kranton. 2002. Identity and schooling: Some lessons for the economics of education. *Journal of economic literature* **40**(4) 1167–1201.

Akerlof, George A, Rachel E Kranton. 2005. Identity and the economics of organizations. *Journal of Economic perspectives* **19**(1) 9–32.

Aksuyek, Efe, Diego Ardila, Boris Battistini, Fredrik Hacklin. 2013. Applying topic modeling to management science: empirical results and implications. *Academy of Management Proceedings*, vol. 2013. Academy of Management Briarcliff Manor, NY 10510, 16105.

Allan, J., C. Wade, A. Bolivar. 2003. Retrieval and novelty detection at the sentence level. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval*. 314–321.

Anderson, T. W., C. Hsiao. 1981. Estimation of dynamic models with error components. *Journal of the American Statistical Association* **76**(375) 598–606.

Anker, Richard. 1997. Theories of occupational segregation by sex: An overview. *Int'l Lab. Rev.* **136** 315.

Anker, Richard. 1998. *Gender and jobs: Sex segregation of occupations in the world*. International Labour Organization.

Apple. 2016. Appstore. *Available at* `https://itunes.apple.com/gb/genre/ios/id36?mt=8`.

Armstrong, Gary, Stewart Adam, Sara Denize, Philip Kotler. 2014. *Principles of marketing*. Pearson Australia.

Arrignton, Michael. 2010. Too few women in tech? stop blaming the men. *Available at* `https://techcrunch.com/2010/08/28/women-in-tech-stop-blaming-me/`.

Ashcraft, Catherine, Brad McLain, Elizabeth Eger. 2016. *Women in tech: The facts*.

Averett, Susan L., Laura M. Argys, Saul D. Hoffman, Joyce P. Jacobsen. 2017. Women and the labor market: A feminist perspective.

Azmat, Ghazala, Barbara Petrongolo. 2014. Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics* **30** 32–40.

Bacharach, Verne R, Alfred A Baumeister, R Michael Furr. 2003. Racial and gender science achievement gaps in secondary education. *The Journal of genetic psychology* **164**(1) 115–126.

Bandura, Albert, Claudio Barbaranelli, Gian Vittorio Caprara, Concetta Pastorelli. 2003. Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development* **72**(1) 187–206.

Bates, Douglas, Martin Mächler, Ben Bolker, Steve Walker. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823* .

Baysal, O., A. J. Malton. 2007. Correlating social interactions to release history during software evolution. *Fourth International Workshop on Mining Software Repositories* 7–7.

Bayus, Barry L. 2013. Crowdsourcing new product ideas over time: An analysis of the dell ideastorm community. *Management science* **59**(1) 226–244.

Becker, Gary S. 2010. *The economics of discrimination*. University of Chicago press.

Benner, Katie. 2017. Females in tech speak frankly on culture of harassment. *Available at* `https://www.nytimes.com/2017/06/30/technology/women-entrepreneurs-speak-out-genderual-harassment.html?mcubz=1`.

Bennett, K. 1996. Software evolution: past, present and future. *Information and Software Technology* **38**(11) 673 – 680.

Bertrand, Marianne, Sendhil Mullainathan. 2004. Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review* **94**(4) 991–1013.

Bessen, J., E. Maskin. 2009. Sequential innovation, patents, and imitation. *RAND Journal of Economics* **40**(4) 611–635.

Beyer, Sylvia, Kristina Rynes, Susan Haller. 2004. Deterrents to women taking computer science courses. *IEEE technology and society magazine* **23**(1) 21–28.

Bharadwaj, A., O.A. El Sawy, P.A. Pavlou, N. Venkatraman. 2013. Digital Business Strategy: Toward a Next Generation of Insights. *MIS Quarterly* **37**(2) 471–482.

Bhargava, H.K., A.N. Mishra. 2014. Electronic medical records and physician productivity: Evidence from panel data analysis. *Management Science* **60**(10) 2543–2562.

Blackburn, Robert M, Bradley Brooks, Jennifer Jarman. 2001. The vertical dimension of occupational segregation. *Work, Employment and Society* **15**(3) 511–538.

Blau, Francine D, Lawrence M Kahn. 2007. The gender pay gap: Have women gone as far as they can? *Academy of Management Perspectives* **21**(1) 7–23.

Blau, Francine D, Lawrence M Kahn. 2017. The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature* **55**(3) 789–865.

Bleeker, Martha M, Janis E Jacobs. 2004. Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology* **96**(1) 97.

Blei, D.M., A.Y. Ng, M.I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* **3** 993–1022.

Boaler, Jo. 2015a. *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. John Wiley & Sons.

Boaler, Jo. 2015b. *What's math got to do with it?: How teachers and parents can transform mathematics learning and inspire success*. Penguin.

Bohn, R.E., R. Jaikumar. 1992. *The Structure of Technological Knowledge in Manufacturing*. Working paper (Harvard University. Graduate School of Business Administration. Division of Research), Harvard Business School, Division of Research.

Bond, S.R. 2002. Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal* **1**(2) 141–162.

Boudreau, Kevin J. 2012a. Let a thousand flowers bloom? an early look at large numbers of software app developers and patterns of innovation. *Organization Science* **23**(5) 1409–1427.

Boudreau, Kevin J, Eva C Guinan, Karim R Lakhani, Christoph Riedl. 2016. Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science* **62**(10) 2765–2783.

Boudreau, Kevin J, Lars B Jeppesen. 2015. Unpaid crowd complementors: The platform network effect mirage. *Strategic Management Journal* **36**(12) 1761–1777.

Boudreau, Kevin J, Nicola Lacetera, Karim R Lakhani. 2011. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management science* **57**(5) 843–863.

Boudreau, Kevin J, Karim R Lakhani. 2013. Using the crowd as an innovation partner. *Harvard business review* **91**(4) 60–9.

Boudreau, K.J. 2012b. Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation. *Organization Science* **23**(5) 1409–1427.

Bowles, Nellie. 2017. Push for Gender Equality in Tech? Some Men Say Its Gone Too Far.

Available at `https://www.nytimes.com/2017/09/23/technology/silicon-valley-men-backlash-gender-scandals.html`.

Bradley, Harriet. 1989. *Men's work, women's work: A sociological history of the sexual division of labour in employment*. University of Minnesota Press.

Brown, S.L., K.M. Eisenhardt. 1995. Product development: Past research, present findings, and future directions. *The Academy of Management Review* **20**(2) 343–378.

Browne, Irene, Joya Misra. 2003. The intersection of gender and race in the labor market. *Annual review of sociology* **29**(1) 487–513.

Brynjolfsson, E., Y. Hu, M.D. Smith. 2003. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science* **49**(11) 1580–1596.

Bui, Quoctrung. 2014. Who studies what? men, women and college majors. *Available at* `https://www.npr.org/sections/money/2014/10/28/359419934/who-studies-what-men-women-and-college-majors`.

Buser, Thomas, Muriel Niederle, Hessel Oosterbeek. 2014. Gender, competitiveness, and career choices. *The Quarterly Journal of Economics* **129**(3) 1409–1447.

Castao, Raquel, Mita Sujan, Manish Kacker, Harish Sujan. 2008. Managing consumer uncertainty in the adoption of new products: Temporal distance and mental simulation. *Journal of Marketing Research* **45**(3) 320–336.

Ceci, Stephen J. 2018. Women in academic science: Experimental findings from hiring studies. *Educational Psychologist* **53**(1) 22–41.

Ceci, Stephen J, Wendy M Williams, Susan M Barnett. 2009. Women's underrepresentation in science: sociocultural and biological considerations. *Psychological bulletin* **135**(2) 218.

Chan, Joel, Christian D Schunn. 2015. The importance of iteration in creative conceptual combination. *Cognition* **145** 104–115.

Chang, Emily. 2018. *Brotopia: Breaking Up the Boys' Club of Silicon Valley*. Portfolio.

Cheryan, Sapna, Victoria C Plaut, Paul G Davies, Claude M Steele. 2009. Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of personality and social psychology* **97**(6) 1045.

Chesbrough, Henry, et al. 2003. Open innovation.

Choudary, S.P., M.W. Van Alstyne, G. Parker. 2016. *Platform Revolution: How Networked Markets Are Transforming the Economy–And How to Make Them Work for You*. 1st ed. W. W. Norton & Company.

Chung, D.J. 2015. How much is a win worth? an application to intercollegiate athletics. *Management Science* 1–18.

Clark, K. B., T. Fujimoto. 1999. *Product Development Performance: Strategy, Organization, and Management in the World Auto Industry*. Harvard Business School Press.

Cohen, J.E., M.A. Lemley. 2001. Patent scope and innovation in the software industry. *California Law Review* **89** 1–57.

Cohen, Wesley M, Daniel A Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative science quarterly* 128–152.

Cohoon, J McGrath, William Aspray. 2006. *Women and information technology: Research on underrepresentation*, vol. 1. The MIT Press.

Cooper, J., K.D. Weaver. 2003. *Gender and Computers: Understanding the Digital Divide*. Taylor & Francis. URL `https://books.google.co.in/books?id=SBLpfoM-6NsC`.

Criscuolo, P., A. Salter, T. Sheehan. 2007. Making knowledge visible: Using expert yellow pages to map capabilities in professional services firms. *Research Policy* **36**(10) 1603 – 1619.

Croson, Rachel, Uri Gneezy. 2009. Gender differences in preferences. *Journal of Economic literature* **47**(2) 448–74.

Csikszentmihalyi, Mihaly. 1997. Flow and the psychology of discovery and invention. *HarperPerennial, New York* **39**.

Cusumano, M.A., K. Nobeoka. 1992. Strategy, structure and performance in product development: Observations from the auto industry. *Research Policy* **21**(3) 265 – 293.

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

Datta, D., S. Kajanan. 2013. Do App Launch Times Impact Their Subsequent Commercial Success? An Analytical Approach. *Proceedings of the 2013 International Conference on Cloud Computing and Big Data* 205–210.

Davis, J.P., P. Yin. 2014. Experimentation Strategies and Entrepreneurial Innovation : Inherited Market Differences in the iPhone Ecosystem. INSEAD Working paper.

de Reuver, Mark, Carsten Sørensen, Rahul C Basole. 2018. The digital platform: a research agenda. *Journal of Information Technology* **33**(2) 124–135.

Doering, Laura, Sarah Thébaud. 2017. The effects of gendered occupational roles on men's and women's workplace authority: Evidence from microfinance. *American Sociological Review* **82**(3) 542–567.

Eagly, Alice H. 1983. Gender and social influence: A social psychological analysis. *American Psychologist* **38**(9) 971.

Ebertz, Joy. 2017. The meritocracy myth of the tech industry in america. *Available at* `https://code.likeagirl.io/the-meritocracy-myth-of-the-tech-industry-in-america-e5c2229c8e40`.

Eccles, Jacquelynne S. 1994. Understanding women's educational and occupational choices: Applying the eccles et al. model of achievement-related choices. *Psychology of women quarterly* **18**(4) 585–609.

Eisenhardt, K.M., J.A. Martin. 2000. Dynamic capabilities: what are they? *Strategic Management Journal* **21**(10-11) 1105–1121.

Eisenhardt, K.M., C.B. Schoonhoven. 1990. Organizational Growth: Linking Founding Team, Strategy, Environment, and Growth Among U.S. Semiconductor Ventures, 1978-1988. *Administrative Science Quarterly* **35**(3) 504–529.

England, Paula. 2010. The gender revolution: Uneven and stalled. *Gender & society* **24**(2) 149–166.

Ethiraj, Sendil K., David H. Zhu. 2008. Performance effects of imitative entry. *Strategic Management Journal* **29**(8) 797–817.

Fan, W., L. Wallace, S. Rich, Z. Zhang. 2006. Tapping the power of text mining. *Communications of the ACM* **49**(9) 76–82.

Ferguson, John-Paul, Gianluca Carnabuci. 2017. Risky recombinations: Institutional gatekeeping in the innovation process. *Organization Science* **28**(1) 133–151.

Fernandez, Roberto M, M Lourdes Sosa. 2005. Gendering the job: Networks and recruitment at a call center. *American Journal of Sociology* **111**(3) 859–904.

Findlay, Patricia, Jeanette Findlay, Robert Stewart. 2009. The consequences of caring:

skills, regulation and reward among early years workers. *Work, employment and society* **23**(3) 422–441.

Fleming, L., O. Sorenson. 2004. Science as a map in technological search. *Strategic Management Journal* **25**(8-9) 909–928.

Fleming, Lee. 2001. Recombinant uncertainty in technological search. *Management Science* **47**(1) 117–132.

Flory, Jeffrey A, Andreas Leibbrandt, John A List. 2014. Do competitive workplaces deter female workers? a large-scale natural field experiment on job entry decisions. *The Review of Economic Studies* **82**(1) 122–155.

Gaimon, C., J. Bailey. 2013. Knowledge management for the entrepreneurial venture. *Production and Operations Management* **22**(6) 1429–1438.

Garg, R., R. Telang. 2013. Inferring app demand from publicly available data. *MIS Quarterly* **37**(4) 1253–1264.

Garratt, Rodney J, Catherine Weinberger, Nick Johnson. 2013. The state street mile: Age and gender differences in competition aversion in the field. *Economic Inquiry* **51**(1) 806–815.

Gavetti, Giovanni, Daniel Levinthal. 2000. Looking forward and looking backward: Cognitive and experiential search. *Administrative science quarterly* **45**(1) 113–137.

Gawer, Annabelle. 2009. Platform dynamics and strategies: from products to services. *Platforms, markets and innovation* **45** 57.

Ghose, A., S.P. Han. 2014. Estimating demand for mobile applications in the new economy. *Management Science* **60**(6) 1470–1488.

Giachetti, Claudio, Joseph Lampel, Stefano Li Pira. 2017. Red queen competitive imitation in the uk mobile phone industry. *Academy of Management Journal* **60**(5) 1882–1914.

Girotra, Karan, Christian Terwiesch, Karl T Ulrich. 2010. Idea generation and the quality of the best idea. *Management science* **56**(4) 591–605.

Glass, Jennifer L, Sharon Sassler, Yael Levitte, Katherine M Michelmore. 2013. What's so special about STEM? A comparison of women's retention in STEM and professional occupations. *Social forces* **92**(2) 723–756.

Glick, Peter, Cari Zion, Cynthia Nelson. 1988. What mediates sex discrimination in hiring decisions? *Journal of Personality and Social Psychology* **55**(2) 178.

Goering, Patricia A. 1985. Effects of product trial on consumer expectations, demand, and prices. *Journal of Consumer Research* 74–82.

Goldin, Claudia. 2014. A grand gender convergence: Its last chapter. *American Economic Review* **104**(4) 1091–1119.

Goolsbee, A., J.A. Chevalier. 2003. Measuring Prices and Price Competition Online: Amazon and Barnes and Noble. *Quantitative Marketing and Economics* **1**(2) 203–222.

Grant, R.M. 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal* **17**(S2) 109–122.

Green, J.R., S. Scotchmer. 1995. On the division of profit in sequential innovation. *RAND Journal of Economics* **26**(1) 20–33.

Gruber, Marc, Dietmar Harhoff, Karin Hoisl. 2013. Knowledge recombination across technological boundaries: Scientists vs. engineers. *Management Science* **59**(4) 837–851.

Haas, M.R., P. Criscuolo. 2015. Which problems to solve? Online knowledge sharing and attention allocation in organizations. *Academy of Management Journal* **58**(3) 680–711.

Hakim, Catherine. 1979. Occupational segregation: a comparative study of britain, the united states and other countries .

Hakim, Catherine. 2002. Lifestyle preferences as determinants of women's differentiated labor market careers. *Work and Occupations* **29**(4) 428–459.

Hall, Bronwyn H, Adam B Jaffe, Manuel Trajtenberg. 2001. The nber patent citation data file: Lessons, insights and methodological tools. Tech. rep., National Bureau of Economic Research.

Harrison, D.A., K.J. Klein. 2007. Whats the difference? diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review* **32**(4) 1199–1228.

Hartmann, Heidi I. 1979. The unhappy marriage of marxism and feminism: Towards a more progressive union. *Capital & Class* **3**(2) 1–33.

Haselhuhn, Michael P, Devin G Pope, Maurice E Schweitzer, Peter Fishman. 2012. The impact of personal experience on behavior: Evidence from video-rental fines. *Management Science* **58**(1) 52–61.

Heath, A. 2015. The 15 best productivity apps in the world. *Available at* `http://www.techinsider.io/best-productivity-apps-2015-10`.

Heiman, Amir, Eitan Muller. 1996. Using demonstration to increase new product acceptance: Controlling demonstration time. *Journal of Marketing Research* 422–430.

Henderson, R.M., K.B. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* **35**(1) 9–30.

Howe, Jeff. 2008. *Crowdsourcing: How the power of the crowd is driving the future of business*. Random House.

Huang, A. 2008. Similarity measures for text document clustering. *Proceedings of the 6th New Zealand Computer Science Research Student Conference* .

Huang, Allen H, Reuven Lehavy, Amy Y Zang, Rong Zheng. 2017. Analyst information discovery and interpretation roles: A topic modeling approach. *Management Science* .

Huang, Peng, Nicholas H Lurie, Sabyasachi Mitra. 2009. Searching for experience on the web: an empirical examination of consumer behavior for search and experience goods. *Journal of marketing* **73**(2) 55–69.

Huang, Yan, Param Vir Singh, Kannan Srinivasan. 2014. Crowdsourcing new product ideas under consumer learning. *Management science* **60**(9) 2138–2159.

Huet, Ellen. 2017. Some Tech Companies Are Trying Affirmative Action Hiring -But Don't Call It That. *Available at* `https://www.bloomberg.com/news/articles/2017-01-10/some-tech-companies-are-trying-affirmative-action-hiring-but-don-t-cal`

Hunt, Jennifer. 2016. Why do women leave science and engineering? *ILR Review* **69**(1) 199–226.

Hyde, Janet S, Sara M Lindberg, Marcia C Linn, Amy B Ellis, Caroline C Williams. 2008. Gender similarities characterize math performance. *Science* **321**(5888) 494–495.

Iansiti, M. 1998. *Technology Integration: Making Critical Choices in a Dynamic World*. Harvard Business School Press, Boston, MA, USA.

Jeppesen, Lars Bo, Karim R Lakhani. 2010. Marginality and problem-solving effectiveness in broadcast search. *Organization Science* **21**(5) 1016–1033.

Jones, Benjamin F. 2009. The burden of knowledge and the death of the renaissance man: Is innovation getting harder? *The Review of Economic Studies* **76**(1) 283–317.

Jung, Hyun Ju, Jeongsik Jay Lee. 2016. The quest for originality: A new typology of knowledge search and breakthrough inventions. *Academy of Management Journal* **59**(5) 1725–1753.

Kahn, Shulamit, Donna Ginther. 2017. Women and stem. Tech. rep., National Bureau of Economic Research.

Kamas, Linda, Anne Preston. 2012. The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization* **83**(1) 82–97.

Kaplan, Sarah, Mary Tripsas. 2008. Thinking about technology: Applying a cognitive lens to technical change. *Research Policy* **37**(5) 790–805.

Kaplan, Sarah, Keyvan Vakili. 2015. The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal* **36**(10) 1435–1457.

Katila, R. 2002. New product search over time: past ideas in their prime? *Academy of Management Journal* **45**(5) 995–1010.

Katila, R., G. Ahuja. 2002. Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of Management Journal* **45**(6) 1183–1194.

Katila, R., E.L. Chen. 2008. Effects of search timing on innovation: The value of not being in sync with rivals. *Administrative Science Quarterly* **53**(4) 593–625.

Katz, R., T.J. Allen. 1982. Investigating the Not Invented Here (NIH) syndrome: A look at the performance, tenure, and communication patterns of 50 R & D Project Groups. *R&D Management* **12**(1) 7–20.

Kavadias, S., K. Ladas. 2015. Six degrees of innovation: Predicting transformative business models. *Available at* `https://www.business.att.com/content/article/cambridge-six-degrees-of-innovation.pdf`.

Kennedy, Peter. 2003. *A guide to econometrics*. MIT press.

Kim, D.J., B. Kogut. 1996. Technological platforms and diversification. *Organization Science* **7**(3) 283–301.

Kimura, H. 2014. The Complete iOS 8 Subcategories

List. *Available at* `https://sensortower.com/blog/the-complete-ios-8-subcategories-list`.

Klein, Lisa R. 1998. Evaluating the potential of interactive media through a new lens: Search versus experience goods. *Journal of business research* **41**(3) 195–203.

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

Köhler, C., W. Sofka, C. Grimpe. 2012. Selective search , sectoral patterns , and the impact on product innovation performance. *Research Policy* **41**(8) 1344–1356.

Kornish, L. J., K. T. Ulrich. 2011. Opportunity Spaces in Innovation: Empirical Analysis of Large Samples of Ideas. *Management Science* **57**(1) 107–128.

Kornish, L.J. 2001. Pricing for a durable-goods monopolist under rapid sequential innovation. *Management Science* **47**(11) 1552–1561.

Kotler, Philip, Gary Armstrong. 2010. *Principles of marketing*. Pearson education.

Kovács, Balázs, Michael T Hannan. 2010. The consequences of category spanning depend on contrast. *Categories in markets: Origins and evolution*. Emerald Group Publishing Limited, 175–201.

Krishnan, V., K. Ramachandran. 2011. Integrated product architecture and pricing for managing sequential innovation. *Management Science* **57**(11) 2040–2053.

Krishnan, V., K.T. Ulrich. 2001. Product development decisions: A review of the literature. *Management science* **47**(1) 1–21.

Kuhn, Peter, Marie Claire Villeval. 2014. Are women more attracted to co-operation than men? *The Economic Journal* **125**(582) 115–140.

Laursen, K., A. Salter. 2006. Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms. *Strategic Management Journal* **27**(2) 131–150.

Lee, Sun Young, Selin Kesebir, Madan M Pillutla. 2016. Gender differences in response to competition with same-gender coworkers: A relational perspective. *Journal of Personality and Social Psychology* **110**(6) 869.

Leiponen, A., C.E. Helfat. 2010. Innovation objectives, knowledge sources, and the benefits of breadth. *Strategic Management Journal* **31**(2) 224–236.

Lerner, Joshua. 1994. The importance of patent scope: an empirical analysis. *The RAND Journal of Economics* 319–333.

Leuwer, Regina. 2018. How Do Users Search for Apps in the App Store? *Available at* `https://asostack.com/how-do-users-search-for-apps-5a0d3257803f`.

Lieberman, Marvin B, Shigeru Asaba. 2006. Why do firms imitate each other? *Academy of management review* **31**(2) 366–385.

Lubinski, David, Camilla Persson Benbow. 2006. Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on psychological science* **1**(4) 316–345.

Lynch Jr, John G, Dan Ariely. 2000. Wine online: Search costs affect competition on price, quality, and distribution. *Marketing science* **19**(1) 83–103.

MacCormack, A., R. Verganti, M. Iansiti. 2001. Developing Products on "Internet Time": The Anatomy of a Flexible Development Process. *Management Science* **47**(1) 133–150.

Maddala, Gangadharrao S, Kajal Lahiri. 1992. *Introduction to econometrics*, vol. 2. Macmillan New York.

Malhotra, Arvind, Ann Majchrzak. 2014. Managing crowds in innovation challenges. *California Management Review* **56**(4) 103–123.

March, J. G. 1991. Exploration and Exploitation in Organizational Learning. *Organization Science* **2**(1) 71–87.

Marianne, Bertrand. 2011. New perspectives on gender. *Handbook of labor economics*, vol. 4. Elsevier, 1543–1590.

Martin, X., W. Mitchell. 1998. The influence of local search and performance heuristics on new design introduction in a new product market. *Research Policy* **26**(7-8) 753–771.

Mascarelli, Amanda. 2010. Tapping the crowd for technologies. *Available at* `https://www.nature.com/news/2010/100701/full/news.2010.331.html`.

McCallum, Andrew Kachites. 2002. Mallet: A machine learning for language toolkit.

McDonough, W., M. Braungart. 2002. *Cradle to Cradle: Remaking the Way We Make Things*. North Point Press, New York.

McMahon, Tamsin. 2017. What's behind the tech industry's toxic masculinity problem

Inside the Valley of the Bros. *Available at* `https://www.theglobeandmail.com/technology/toxic-masculinity-in-silicon-valley/article35759481/`.

Molla, Rani. 2017. The app economy will be worth $6 trillion in five years thanks to mobile commerce. *Available at* `https://www.recode.net/2017/6/27/15876294/app-economy-6-trillion-mobile-commerce-amazon-alibaba`.

Mollick, Ethan, Ramana Nanda. 2015. Wisdom or madness? comparing crowds with expert evaluation in funding the arts. *Management Science* **62**(6) 1533–1553.

Mueller, Jennifer S, Shimul Melwani, Jack A Goncalo. 2012. The bias against creativity: Why people desire but reject creative ideas. *Psychological science* **23**(1) 13–17.

Mukherjee, Satyam, Brian Uzzi, Ben Jones, Michael Stringer. 2016. A new method for identifying recombinations of existing knowledge associated with high-impact innovation. *Journal of Product Innovation Management* **33**(2) 224–236.

Mumford, Michael D, Sigrid B Gustafson. 1988. Creativity syndrome: Integration, application, and innovation. *Psychological bulletin* **103**(1) 27.

Mundy, Liz. 2017. Why is silicon valley so awful to females? *Available at* `https://www.theatlantic.com/magazine/archive/2017/04/why-is-siliconvalley-so-awful-to-women/517788/`.

Muralidharan, Karthik, Paul Niehaus. 2017. Experimentation at scale. *Journal of Economic Perspectives* **31**(4) 103–24.

Murciano-Goroff, Raviv. 2018. Missing women in tech: The labor market for highly skilled software engineers. Working paper.

Nambisan, Satish, Kalle Lyytinen, Ann Majchrzak, Michael Song. 2017. Digital innovation management: Reinventing innovation management research in a digital world. *MIS Quarterly* **41**(1).

Narayan, Vishal, Vrinda Kadiyali. 2015. Repeated interactions and improved outcomes: an empirical analysis of movie production in the united states. *Management Science* **62**(2) 591–607.

National Student Clearinghouse Research, Center. 2017. Snapshot report. Avail-

able at https://nscresearchcenter.org/wp-content/uploads/SnapshotReport27r.pdf (2017).

Nelson, Richard R, S.G. Winter. 1982. *An evolutionary theory of economic change*. harvard university press.

Nelson, R.R., S.G. Winter. 2002. Evolutionary theorizing in economics. *Journal of Economic Perspectives* **16**(2) 23–46.

Nickell, S. 1981. Biases in dynamic models with fixed effects. *Econometrica* **49**(6) 1417–26.

Niederle, Muriel, Lise Vesterlund. 2007. Do women shy away from competition? do men compete too much? *The quarterly journal of economics* **122**(3) 1067–1101.

Niederle, Muriel, Lise Vesterlund. 2011. Gender and competition. *Annu. Rev. Econ.* **3**(1) 601–630.

Panzarino, Matthew. 2016. Apple's App Store at the end of the app era. *Available at https://techcrunch.com/2016/06/12/apples-app-store-at-the-end-of-the-app-era/*.

Parker, G., M.W. Van Alstyne. 2017. Innovation, openness, and platform control. *Management Science* doi: doi.org/10.1287/mnsc.2017.2757. Articles in Advance.

Perez-Felkner, Lara, Sarah-Kathryn McDonald, Barbara Schneider, Erin Grogan. 2012. Female and male adolescents' subjective orientations to mathematics and the influence of those orientations on postsecondary majors. *Developmental Psychology* **48**(6) 1658.

Perry-Smith, Jill E, Christina E Shalley. 2003. The social side of creativity: A static and dynamic social network perspective. *Academy of management review* **28**(1) 89–106.

Pich, Michael T, Christoph H Loch, Arnoud de Meyer. 2002. On uncertainty, ambiguity, and complexity in project management. *Management Science* **48**(8) 1008–1023.

Picker, Randal C. 2005. Cybersecurity: of heterogeneity and autarky. *The Law and Economics of Cybersecurity* 115.

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

Poetz, Marion K, Martin Schreier. 2012. The value of crowdsourcing: can users really

compete with professionals in generating new product ideas? *Journal of product innovation management* **29**(2) 245–256.

Polachek, Solomon W. 1976. Occupational segregation-alternative hypothesis. *Journal of Contemporary Business* **5**(1) 1–12.

Porter, Alan L, Frederick A Rossini. 1985. Peer review of interdisciplinary research proposals. *Science, technology, & human values* **10**(3) 33–38.

Powell, Abigail, Barbara Bagilhole, Andrew Dainty. 2009. How women engineers do and undo gender: Consequences for gender equality. *Gender, Work & Organization* **16**(4) 411–428.

Prediger, Dale J. 1982. Dimensions underlying holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior* **21**(3) 259–287.

Puliafito, Antonio, Nathalie Mitton, Symeon Papavassiliou, Kishor S. Trivedi. 2012. Editorial: Special issue on internet of things: convergence of sensing, networking, and web technologies. *EURASIP Journal on Wireless Communications and Networking* **2012**(1) 212.

Ramachandran, K., V. Krishnan. 2008. Design architecture and introduction timing for rapidly improving industrial products. *Manufacturing & Service Operations Management* **10**(1) 149–171.

Rao, H., R. Drazin. 2002. Overcoming resource constraints on product innovation by recruiting talent from rivals: A study of the mutual fund industry, 1986-94. *The Academy of Management Journal* **45**(3) 491–507.

Reskin, Barbara F, Michelle L Maroto. 2011. What trends? whose choices? comment on england. *Gender & Society* **25**(1) 81–87.

Reskin, Barbara F, Patricia A Roos. 2009. *Job queues, gender queues: Explaining women's inroads into male occupations*. Temple University Press.

Rhyn, Marcel, Ivo Blohm, Jan Marco Leimeister. 2017. Understanding the emergence and recombination of distant knowledge on crowdsourcing platforms. *Proceedings of the International Conference on Information Systems*.

Ridgeway, Cecilia L. 2011. *Framed by gender: How gender inequality persists in the modern world*. Oxford University Press.

Rindova, Violina P, Antoaneta P Petkova. 2007. When is a new thing a good thing? techno-

logical change, product form design, and perceptions of value for product innovations. *Organization Science* **18**(2) 217–232.

Roberts, John H, James M Lattin. 1991. Development and testing of a model of consideration set composition. *Journal of Marketing Research* 429–440.

Rochet, J., J. Tirole. 2003. Platform competition in two-sided markets. *Journal of the European Economic Association* **1**(4) 990–1029.

Rogstadius, Jakob, Maja Vukovic, CA Teixeira, Vassilis Kostakos, Evangelos Karapanos, Jim Alain Laredo. 2013. Crisistracker: Crowdsourced social media curation for disaster awareness. *IBM Journal of Research and Development* **57**(5) 4–1.

Roodman, D. 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal* **9**(1) 86–136.

Rosenkopf, L., A. Nerkar. 2001. Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal* **22**(4) 287–306.

Ross, J.M., D. Sharapov. 2015. When the leader follows: Avoiding dethronement through imitation. *Academy of Management Journal* **58** 658–679.

Sankaranarayanan, Ramesh. 2007. Innovation and the durable goods monopolist: The optimality of frequent new-version releases. *Marketing Science* **26**(6) 774–791.

Scotchmer, S. 1996. Protecting early innovators: Should second-generation products be patentable? *RAND Journal of Economics* **27**(2) 322–331.

Seetharaman, Deepa. 2017. Facebook's Female Engineers Claim Gender Bias. *Available at* `https://www.wsj.com/articles/facebooks-female-engineers-claimgender-bias-1493737116`.

Senot, Claire, Aravind Chandrasekaran, Peter T Ward, Anita L Tucker, Susan D Moffatt-Bruce. 2015. The impact of combining conformance and experiential quality on hospitals readmissions and cost performance. *Management Science* **62**(3) 829–848.

Seymour, Elaine. 2002. Tracking the processes of change in us undergraduate education in science, mathematics, engineering, and technology. *Science Education* **86**(1) 79–105.

Shimbun, Asahi. 2013. The Most Expensive Product Recalls. *Available at* `http://mallet.cs.umass.edu/`.

Shurchkov, Olga, Catherine C Eckel. 2017. Gender differences in behavioral traits and labor market outcomes. *The Oxford Handbook of Women and the Economy*.

Sidhu, J.S., H.R. Commandeur, H.W. Volberda. 2007. The multifaceted nature of exploration and exploitation: Value of supply, demand, and spatial search for innovation. *Organization Science* **18**(1) 20–38.

Sommer, Svenja C, Christoph H Loch. 2004. Selectionism and learning in projects with complexity and unforeseeable uncertainty. *Management science* **50**(10) 1334–1347.

Sommer, Svenja C, Christoph H Loch, Jing Dong. 2009. Managing complexity and unforeseeable uncertainty in startup companies: An empirical study. *Organization Science* **20**(1) 118–133.

Spelke, Elizabeth S. 2005. Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *American Psychologist* **60**(9) 950.

Stoet, Gijsbert, David C Geary. 2018. The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological science* **29**(4) 581–593.

Suarez, Fernando F, Michael A Cusumano, Steven J Kahl. 2013. Services and the business models of product firms: An empirical analysis of the software industry. *Management Science* **59**(2) 420–435.

Sutton, R.I., A. Hargadon. 1996. Brainstorming groups in context: Effectiveness in a product design firm. *Administrative Science Quarterly* **41**(4) 685–718.

Terrell, Josh, Andrew Kofink, Justin Middleton, Clarissa Rainear, Emerson Murphy-Hill, Chris Parnin, Jon Stallings. 2017. Gender differences and bias in open source: Pull request acceptance of women versus men. *PeerJ Computer Science* **3** e111.

Thomke, S.H. 2003. *Experimentation Matters: Unlocking the Potential of New Technologies for Innovation*. Innovation / Harvard Business School, Harvard Business School Press.

Tripsas, Mary, Giovanni Gavetti. 2000. Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic management journal* **21**(10-11) 1147–1161.

Turk, Victoria. 2017. What men in tech can do to address terrible gender inequality. *Available at* `https://www.wired.co.uk/article/heforshe-metoo-men-in-tech-need-to-step-up-for-gender-equality`.

Utterback, James M. 1996. *Mastering the Dynamics of Innovation*. Harvard Business School Press, Boston, MA, USA.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, Ben Jones. 2013. Atypical combinations and scientific impact. *Science* **342**(6157) 468–472.

Vaidos, A. 2017. Google Play Store vs Apples App Store - A Comparison. *Available at* `http://news.softpedia.com/news/google-play-store-vs-apple-s-app-store-a-comparison-512601.shtml`.

Verbeke, Geert, Geert Molenberghs. 2009. *Linear mixed models for longitudinal data*. Springer Science and Business Media.

von Hippel, E., M.J. Tyre. 1995. How learning by doing is done: problem identification in novel process equipment. *Research Policy* **24**(1) 1–12.

Walby, Sylvia. 1989. Theorising patriarchy. *Sociology* **23**(2) 213–234.

Walker, Henry A, Mary L Fennell. 1986. Gender differences in role differentiation and organizational task performance. *Annual Review of Sociology* **12**(1) 255–275.

Wang, Jian, Reinhilde Veugelers, Paula Stephan. 2017. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy* **46**(8) 1416–1436.

Weisberg, Robert W. 1999. Creativity and knowledge: A challenge to theories. *Handbook of creativity* **226**.

Wolf, Michael, Dalton Terrell. 2016a. The high-tech industry, what is it and why it matters to our economic future. *Beyond the Numbers: Employment & Unemployment* **5**(8).

Wolf, Michael, Dalton Terrell. 2016b. The high-tech industry, what is it and why it matters to our economic future .

Woodcock, Anna, William G Graziano, Sara E Branch, Meara M Habashi, Ida Ngambeki, Demetra Evangelou. 2013. Person and thing orientations: Psychological correlates and predictive utility. *Social Psychological and Personality Science* **4**(1) 116–123.

Wooldridge, J. 2001. *Econometric Analysis of Cross Section and Panel Data*, *MIT Press Books*, vol. 1. The MIT Press.

Yoo, Byungjoon, Kwansoo Kim. 2012. Does popularity decide rankings or do rankings decide popularity? An investigation of ranking mechanism design. *Electronic Commerce Research and Applications* **11**(2) 180–191.

Yu, Lingo. 2009. Mining change logs and release notes to understand software maintenance and evolution. *CLEI Electronic Journal* **12**(2).

Zuckerman, Ezra W. 1999. The categorical imperative: Securities analysts and the illegitimacy discount. *American journal of sociology* **104**(5) 1398–1438.