



Abstracts Booklet

Conference on the Digital Economy by the Indian School of Business

Taj Malabar Resort & Spa, Kochi

Jan 3-5, 2023

1 Distribution Contracts for Digital Content Supply Chains in the Presence of Online Piracy
Prasenjit Mandal and Abhishek Roy

Content manufacturers often face the issue of piracy while distributing digital contents (e.g., books, music, games etc.) through online platforms such as Amazon, Apple iOS and Google Play. Existing literature on digital piracy focused on a variety of strategies such as implementing copyright protection technologies to limit piracy of information goods. In practice, distribution contracts like wholesale contract or agency contract adopted by online platforms and content manufacturers may impact the digital piracy. Moreover, manufacturers may increase the content quality that enhance the demand of genuine content. In this paper, we develop a game-theoretic model with a content manufacturer, an online platform and a continuum of consumers in which the manufacturer distributes its content to consumers through the platform. We examine the economic impact of digital piracy on the distribution contract choice problem of the platform. Further, we investigate how the platform's distribution contract choice affects its own protection level and content manufacturer's quality to deter digital piracy. We find that the wholesale contract is more effective compared to the agency contract to deter the piracy when the commission rate is low. In the presence of piracy, the manufacturer sets a higher content quality and the platform decides a higher protection level under wholesale contract than that in agency contract when the commission rate is low. Our results suggest that, in the absence of piracy, the content manufacturer and the platform are better off only with the agency contract when the commission rate is high, and a win-win situation never arises with the wholesale contract. In contrast, in the presence of piracy, all parties may benefit from the wholesale contract when the commission rate is low. Our study provides important insights to practitioners to understand the role of distribution contracts in limiting digital piracy.

An Unintended Consequence of Platform Dependence: Empirical Evidence from IT-enabled Food-Delivery

2	<p>Platforms</p> <p>Varun Karamshetty, Michael Freeman and Sameer Hasija</p> <p>IT-enabled food-delivery platforms increase restaurants' catchment area and reduce search costs for consumers. The former allows restaurants to pool their risks and decrease noise in demand, whereas the latter increases competition and noise in demand. Overall, it is unclear how these competing factors affect restaurants' ability to forecast their demand -a critical factor for profitability. In this paper, we empirically investigate the impact of IT-enabled food-delivery platforms on restaurants' demand forecast error. Using detailed transaction-level datasets, we find that a 10 percentage point increase in dependence on food-delivery platforms leads to a 2.83% increase in overall forecast error. We also find that the majority of increase in overall error is due to an increased error in forecasting intra-day demand "pattern", and a smaller portion is due to error in forecasting inter-day demand "amplitude". Based on our findings, we offer suggestions for restaurants on managing their relationship with IT-enabled food-delivery platforms.</p>
3	<p>Lemon Ads: Adverse Selection in Multi-Channel Display Advertising Market</p> <p>Francesco Balocco, Yixin Lu, Ting Li and Alok Gupta</p> <p>Many advertisers in the display advertising market feel they are paying for impressions that users rarely see. Such quality concerns have driven increasing adoption of the private marketplace (PMP) channel in parallel with the real-time bidding (RTB) channel. RTB is open to all advertisers, whereas PMP requires advertisers to stipulate an ex-ante agreement with publishers that provides quality assurance of the ad inventory. Despite being seen as a credible channel for high-quality impressions, it is unclear how the adoption and use of PMP would affect the payoffs of publishers and advertisers in RTB, and whether the co-existence of the two market channels benefits both publishers and advertisers in the long run. To address these questions, we first develop a stylized game-theoretical model to characterize the strategic interactions between advertisers and publishers in a dual-channel setting. The equilibrium analysis shows that dual-channel publishers can extract adverse selection rents from their single-channel counterparts in RTB. As such, advertisers do not benefit from the quality assurance promised by PMP. We then conduct an empirical test of our theoretical predictions using a large proprietary dataset. The results provide strong evidence of adverse selection. In particular, we find that dual-channel publishers offer significantly lower quality impressions than their single-channel counterparts in RTB. Our study provides timely insights to practitioners in the ad industry.</p>
4	<p>Status Downgrade: The Impact of Losing Status on a User Generated Content Platform</p> <p>Vandith Pamuru, Wreeto Kar and Warut Khernamnuai</p>

	<p>Platform recognition in the form of non-financial incentives such as badges, status, and ranks are commonly employed to incentivize user contributions. We study one such incentive, "status", in the context of a user-generated content platform, specifically a third-party online restaurant-review platform. While previous literature has extensively focused on the impact of non-financial incentives on subsequent contributions from rewarded users, we focus on the temporal impact on user contribution when such incentives are withdrawn. Particularly, we explore whether, after a contributor loses status, the intrinsic quality and perceived quality of contents generated by the contributor are affected differently. We utilize natural language processing (NLP) techniques to extract these quality metrics from the textual content of online reviews in our dataset. Using a unique empirical strategy, we show that the intrinsic quality of reviews significantly decreases after a reviewer is demoted by a platform. However, these reviews, despite their lower intrinsic quality, continue to be perceived highly useful by platform users. We also find that characteristics related to temporal association with the platform or with the elevated status, do not moderate the effect of status loss on intrinsic or perceived quality of reviews written post-demotion. These results have significant implications for platform managers who design systems related to status and how such status should be shown to platform users.</p>
5	<p>Gender Bias in Remote Work During the COVID-19 Pandemic: Evidence from Healthcare Kartik K. Ganju, Jeffery McCullough, Kathy Li, Chandy Ellimoottil</p> <p>The COVID-19 pandemic resulted in a massive shift to the digital delivery of services. In this paper, we examine if the gender of the service provider moderates this transition. Specifically, we utilize data on in-person and virtual primary healthcare to study if the gender of the physician affects the shift to virtual healthcare. We find that female physicians experienced a larger reduction in the delivery of services via conventional methods during the pandemic. Although female providers delivered more of their services digitally, they suffered a net decrease in the services they provided. For female physicians, the likelihood of having a child in the household was correlated with the amount of virtual services provided. However, correlations between being a parent and digital delivery of service were absent for male physicians. Relative to their male colleagues, female healthcare providers with lower autonomy (such as nurse practitioners and physician assistants) did not use a significantly higher amount of telemedicine. Their digital service provision was also uncorrelated with being a parent. Finally, women's presence in firms increased remote service delivery – physicians of both genders in majority-female clinics provided significantly more telemedicine. These results underscore the gendered difference of the shift to remote services, implications for mothers, low autonomy workers, and institutions.</p>
6	<p>Uncovering data & algorithmic bias in digital markets Esi Adeborna and Luvai Motiwalla</p>

	<p>With the increase in the use of algorithms in decision making in different spheres of life, the need for “fairness” in these machine learning (ML) algorithms has become a progressively important concern. However, any remedy to bias must start with awareness that bias exists as it helps guide toward a solution. Therefore, this work is an empirical study that aims to uncover data and subsequent algorithmic bias in one of such ML applications. Specifically, we consider recommender systems (RS), one of the most successful applications of ML technology in practice today, and answer three main questions: (1) Does bias exist in RS input data? (2) Does the use of RS promote diversity of recommendation list or reinforces existing data bias? and (3) Does the application of RS over the years improve the recommendation diversity and reduce bias? We use Cell Phones and Accessories purchase dataset from Amazon for our empirical study. We employ the Long Tail phenomenon and quantify the shape of purchase distribution by calculating Gini coefficient from the Lorenz curve to determine the presence of bias. In the end, our findings from our experiments showed the presence of bias in both data and the RS algorithm year-over-year.</p>
7	<p>Content Moderation and AI: Impact on Marginalized Communities Jonathan Gomez Martinez and Ramnath Chellappa</p> <p>Social media platforms are increasingly adopting AI-based content moderation to handle the scale of interactions and to mitigate abusive usage and spread of misinformation. While these purported goals may sound noble, the lack of context in many AI-based natural language processing algorithms may inadvertently be harmful, at least in some instances. Our research sheds light on these unintended consequences of technology usage by an empirical investigation of a marginalized community on Twitter. Exploiting a natural experiment where Twitter increased its reliance on AI-based moderation, our results show that such usage can be insensitive to interaction amongst a community like the LGBTQ+ users and fundamentally force them to change their basis of interaction. We caution against the overzealous use of such technology without adequate understanding of its broader impact.</p>
8	<p>Gender Effects in Online Evaluations: Evidence from a Large-Scale Online Discussion-based Community Anand Gopal, David Waguespack, Tristan Botelho and Johanna Birnir</p> <p>Online communities thrive on the basis of interactions between like-minded individuals, and usually involve some form of feedback or evaluations by peers. In these contexts, there is systematic evidence of gender-based biases in evaluations, even when these evaluation processes involve low-threshold actions like “Likes” and Upvotes/Downvotes. How can such biases be attenuated? We study the efficacy of one approach taken to reducing these biases—the process of anonymization of gender information on the community. We use data from a large-scale digital discussion platform for academics, Political Science Rumors, to examine the presence of gender bias. When users on the community post a discussion message, they are randomly assigned a pseudonym in the form of a given (or first name), such as “Daniel” or “Haylee,” and each post subsequently garners positive and negative votes from</p>

	<p>readers. We analyze the up votes, down votes, and net votes garnered by 1.4 million posts where names are randomly assigned to posters. We find that posts from randomly assigned “female” names receive 2.5% lower evaluation scores in the forms of cumulative votes, all else equal. Further, we find that when “female” users post emotive content with a negative tone, the posts receive disproportionately more negative votes. Names that are clearly “male” or ambiguous do not receive these penalties. While the observed effect is modest, the patterns are nonetheless surprising, given the well-known process of randomization on this forum. We discuss the implications of our work for platform design, gender bias mitigation, and role incongruity in online communities where such simple evaluation processes are extremely common.</p>
9	<p>It Takes a Village: Contextualized and Community-Engaged EdTech Interventions in Rural India Aravinda Garimella and Ananya Tiwari</p>
	<p>Experimental research on education in developing countries in the past decade has focused on one main policy question: How can we improve learning outcomes? While education inputs such as infrastructure, teachers, and materials typically account for most of the education expenditure, a critical determinant of the extent to which these inputs translate into better learning outcomes is pedagogy inside and outside the classroom. We conduct a large-scale field experiment across 31 schools with 4,500 students in the northern Indian state of Haryana in collaboration with an Indian NGO. With the support of school administrators, teachers, students, parents, and the state government, we implement a set of three stackable, technology- aided pedagogical interventions that are both contextualized and community-engaged. We contribute to the rich experimental literature in education, economics, and information systems that examines how to improve academic and well-being outcomes of students in developing countries.</p>
10	<p>Social Determinants of Health and ER Utilization: Role of Information Integration during COVID-19 Indranil Bardhan, Tianjian Guo and Anjum Khurshid</p>
	<p>Emergency room (ER) admissions are the front door for the utilization of a community’s health resources and serve as a valuable proxy for a community health system’s capacity. While recent research suggests that social determinants of health (SDOH) are important predictors of patient health outcomes, their impact on ER utilization during the COVID-19 pandemic is not well understood. Further, the role of hospital information integration in moderating the impact of SDOH on ER utilization has not received adequate attention. Utilizing longitudinal claims data from a regional health information exchange spanning six years including the COVID-19 period, we study how SDOH affects ER utilization and whether effective integration of patient medical information across hospitals can moderate its impact. Our results suggest that a patient’s economic well-being significantly reduces their risk of future ER admission. The magnitude of this relationship is more significant when patients are treated at hospitals with high information</p>

integration of health data but weakens when patients receive care at hospitals with lower levels of information integration. Instead, patients' family support can reduce ER utilization when they are treated at hospitals with low information integration. In other words, different dimensions of SDOH are important in low versus high information integration settings. Furthermore, predictive modeling shows that patient visit type and prior visit history can significantly improve the predictive accuracy of ER utilization. Our research implications support efforts to develop national standards for collection and sharing of SDOH data, and their use and interpretation for clinical decision making by healthcare providers.

11 Between a Rock and Hard Place- To grow or to maintain? Studying the Effects of Cloud-Sourcing on M&A Growth

Moksh Matta, Kiron Ravindran, Hyeokkoo Eric Kwon and Gautam Ray

M&As are commonly employed to pursue growth and compete in the marketplace. M&As are used to redeploy existing resources or to acquire capabilities from a target company. Global M&A deals hit a record high in 2021 with over 62,000 deals being announced, up from 24% in 2020 (PwC 2021). The value of these deals reached an all-time high of US\$5.1 trillion. The role of IT in M&A activity had been under recognized with CIO's of acquirers finding about their firm's M&A activities from press reports rather than being involved in identification and screening of M&A targets, and in the planning of post M&A integration. However, recent work, both academic and practitioner-oriented have started recognizing the importance of IT in M&A outcomes (e.g. Benitez et al. 2018; Deloitte 2022; Hedman and Sarker 2015; Lau et al. 2012; Tanriverdi and Uyusal 2011; Tanriverdi and Uysal 2015; Yetton et al. 2013). Tanriverdi and Uyusal (2015) find that the acquirer's IT integration capability is positively related with the market response to an M&A announcement. Chengxin et al. (2021) find that the standardization of the acquirer's enterprise systems affect the premiums paid and the market response to the M&A announcement and that the extensiveness of the enterprise systems affects long-term M&A performance. Benitez et al., (2018) find that a flexible IT infrastructure provides business flexibility to integrate the IT better and realize economic benefits from the M&A.

These studies suggest that the acquirer's IT investments not only improves post M&A performance but also helps to identify and screen M&A targets. However, in all these studies, the acquirer continues to own the technology post merger. This raises the question do firms have to own their IT infrastructure to take advantage of M&A opportunities? More specifically, following Benitez et al. (2018) argument about flexible IT infrastructure, can cloud-sourced IT infrastructure also help firms realize growth opportunities from M&A? This is the research question we pursue in this study. Answering to the question becomes especially challenging given the limited information on firms' cloud sourcing strategies.

Against this backdrop, we analyze a novel and unique cloud-sourcing data from the International Data Corporation (IDC), which has never been used in the literature to the best of our knowledge. The data is matched with M&A data from the SDC Platinum database and 200 million online job vacancy postings posted on over 40,000 distinct online job platforms in the United States. Drawing on a series of empirical analyses and robustness checks for endogeneity, long-term effects, and variable and model specifications, we

endeavor to answer the question.

We find that investing in cloud computing is positively associated with value and number of M&A deals. The non-essentiality of holding ownership and control rights (OCR) over resources while appropriating value from them is a key contribution of our work. It has widely been argued in the property rights theory (PRT) literature that OCR over resources is essential for “resource owner’s ability to create, appropriate, and sustain value from resources” (Foss et al. 2005, p. 542). OCR is expected to provide the incentives and latitude (through residual rights) needed to utilize resources efficiently for performance gains (Grossman and Hart 1986; Hart and Moore 1990). Our findings suggest that cloud-sourcing can enable firms to use and appropriate value from infrastructure resources while eschewing OCR.

Our analysis also reveals that the adoption of a public cloud-sourcing strategy positively contributes towards firms’ ability to sense and seize M&A opportunities while private cloudsourcing strategy turns out to be ineffective. This finding suggests that cloud-sourcing, by mitigating infrastructure maintenance concerns, can enable firms to more efficiently utilize their information resources for identifying and investing in M&A targets. In other words, cloudsourcing helps firms to resolve the trade-off between growth and maintenance. While prior work offers insights into the type of IT infrastructure that is conducive for M&A growth (e.g. Benitez et al. 2018; Lau et al. 2012) or the importance of IT integration challenges in M&A deals (e.g. Baker and Niederman 2014; Henningsson and Carlsson 2011; Hedman and Sarker 2015; Tanriverdi and Uysal 2011), our work highlights how the effective management of the extent of IT infrastructure OCR plays a crucial role in achieving M&A growth.

Our analysis further indicates that an exploitative cloud orientation is more conducive for appropriating the benefits of public cloud-sourcing for M&A growth. This finding suggests that firms’ understanding and ability in leveraging cloud technologies for efficiency improvements plays a crucial role in realizing gains from cloud-sourcing. In studying the effects of outsourcing strategies on firm-level outcomes, several moderating factors have been explored by prior works – cost structure (Loh and Venkatraman 1995), IT capabilities (Wang et al. 2008), absorptive capacity (Rothaermel and Alexandre 2009), IT intensity (Chang and Gurbaxani 2012), IT labor, hardware, software (Han and Mithas 2013), credible commitments and control rights (Susarla and Mukhopadhyay 2019). Our work contributes to this research by highlighting the moderating role played by technological orientation. By examining the combined growth effects of technological orientation and outsourcing strategies, our work also ties into research on the organizational implications of technological orientation (e.g. Subramani 2004; Im and Rai 2008; Lee et al. 2015).

Our analysis also reveals the informate industry environments are more conducive for appropriating the benefits of public cloud-sourcing for M&A growth. The emphasis on improving information flows in these industries, this finding suggests, plays a crucial role in ensuring that information flows supporting M&A decision-making benefit from cloud-based efficiencies. Prior research suggests that technological resources can be used for three roles based on industry type automate, informate, and transform (Zuboff 1985; Armstrong and Sambamurthy 1999). Studies have examined the moderating influence of these industry types in different contexts such as IT leadership (Chatterjee et al. 2001), IT announcements (Dehning et al. 2003), and Y2K spending (Anderson et al. 2006). Our work ties into this research by demonstrating the moderating influence of these industry types in the IT outsourcing context. Our work also ties into PRT research on OCR strategies and their outcomes (e.g. Leiponen et al. 2008; Magelssen et al. 2020; He et al. 2021) by introducing the role played by technological resources as an appropriability condition. The theoretical mechanism of the cloud-sourcing impact, i.e., cloud-sourcing de-couples the use of infrastructure for growth from its

	<p>maintenance thus balancing growth with maintenance concerns, is further supported by the additional analyses regarding the impact on related and unrelated M&As and qualitative interviews with experts from Microsoft, Wipro, Altair and Google.</p>
12	<p>Tethered Durable Goods and Installed Base Degradation via Software Updates: Implications for Product Policy</p> <p>Ramesh Shankar</p> <p>An installed base of products that are technologically tethered, whether digital products such as smartphones or predominantly non-digital products such as automobiles, can be now degraded in performance through software updates. In this paper we study a monopolist technology vendor's decision to degrade their installed base when they release their newer version. This raises new possibilities and temptations for sellers, and raises new questions not hitherto addressed in the literature on durable goods and innovation, on the optimal product policy for durable goods monopolists. In a two-period setting featuring a monopolist selling a durable good to a unit mass of consumers with uniformly distributed valuation for the good, in the absence of a secondary market, we first establish that installed base degradation can indeed emerge in equilibrium, with or without innovation. It is no longer necessary for the seller to produce an improved version, in order to obsolete the previous version. We establish three novel and surprising results: (i) When consumers in period 1 expect product degradation in period 2, the seller is sometimes better off adopting leasing in period 1, despite having the ability to limit product durability. This reverses the findings from the traditional durable goods literature, where leasing remedies the seller's inability to limit product durability. (ii) The seller engages in "unplanned obsolescence" due to an inability to commit: the seller degrades their version 1 in period 2 (i.e. engages in obsolescence) even though it is profit-improving to avoid degrading. (iii) When consumers in period 1 do not expect and are unaware of the upcoming product degradation in period 2, under some conditions the seller's commitment problem causes them to offer excessive durability. This reverses the findings from the traditional durable goods literature, where the seller's commitment problem causes them to offer insufficient durability. The seller does not face the problem of unplanned obsolescence when consumers are unaware of the upcoming product degradation.</p>
13	<p>Senior Executives' Digital Agency and Abnormal Market Returns</p> <p>Sudhir Voleti and Abhishek Kathuria</p> <p>Despite the importance of digital technologies (DT), investors have a limited ability to appraise a firm's investments in these technologies. An emerging stream of literature seeks to uncover means to reduce information asymmetry between external investors and internal stakeholders regarding DT initiatives. We argue that these prior studies fail to find short-term returns to DT initiatives</p>

	<p>as they do not accurately capture senior executives' agency. We suggest that senior executives' communications with external investors reflect agency, thus reduce information asymmetry, and thereby should form the basis for an investigation into the short-term abnormal returns accruing from DT. We use theoretical insights from the vast literature on agency and structure to define DT Agency as the ability of Senior Executives to apply DT Schemas to existing resource arrays, where DT Schemas are senior executives' individual and shared knowledge structures concerning the use of DT to mobilize and remobilize a firm's resources. We hypothesise that senior executives' DT Agency as reflected through their DT-related communications with external investors is associated with positive short- and long-term and abnormal stock market returns, with past-oriented communications realising greatest short-term returns, and future-oriented communications realising greatest long-term returns. A battery of econometric analyses of earnings conference call transcripts from all S&P1500 firms over 15 years (2005-2019), constituting 52,918 usable observations, demonstrate broad support for our model.</p>
14	<p>Economic Impact of Category-Expansion-Oriented Recommendations: Evidence from Randomized Field Experiments</p> <p>Ravi Bapna, Meizi Zhou, Gedas Adomavicius and Jonathan Hershaff</p>
	<p>Successful new category introduction boosts business revenue, increases customer satisfaction and engagement, and helps to enhance life-time customer value. Machine-learning-based recommendations is one commonly used class of techniques for new category introduction, often used to facilitate the diversity of product sales and consumption. While traditional recommender systems based on collaborative filtering algorithms are known to decrease sales diversity (Lee and Hosanagar, 2019), specific ML-based techniques have been proposed, e.g., using different item ranking strategies, to improve recommendation diversity and avoid monotony in the products shown to consumers (Adomavicius and Kwon, 2011; Zhang and Hurley, 2008). However, the evidence for the effectiveness of such techniques – in particular, the in-vivo causal economic impact of recommending new categories – is largely missing in the literature. This paper addresses this gap by combing machine learning (ML) and randomized controlled trials (RCTs) in collaboration with a leading e-commerce company. Specifically, we conduct randomized experiments and answer three questions: (1) How does recommending a new category to an existing consumer influence her purchase behavior (recommendation effect)? (2) How does a consumer's purchase behavior change when she is recommended multiple non-purchased categories (choice effect)? (3) How does a consumer's purchase behavior change when she is recommended personalized non-purchased categories (personalization effect)?</p>

15	<p data-bbox="271 212 902 240">Impact of AI on Reviews and Outcomes</p> <p data-bbox="271 261 869 290">Rachit Kamdar and Siva Viswanathan</p> <p data-bbox="271 328 2042 488">Online reviews play an important role in disseminating information and reducing information asymmetry, which makes reviews an important aspect of online platforms. Despite years of research and study on online reviews, platforms still face the problem of under provisioning of reviews. Researchers and practitioners have tried motivating review generation using different means such as financial incentives, non-monetary rewards, and social norm-based interventions. Recently, platforms have started leveraging Artificial Intelligence tools to help reviewers write reviews. This study tries to answer three important research questions in this context:</p> <ol data-bbox="271 509 1252 603" style="list-style-type: none"> <li data-bbox="271 509 1043 537">1. Who uses AI generated tags for textual review writing? <li data-bbox="271 541 1151 569">2. How do users use AI generated tags for textual review writing? <li data-bbox="271 572 1252 603">3. What are the impacts of using AI generated tags on review outcomes? <p data-bbox="271 624 2042 748">Using data from a large Food-tech review platform in India, where we can observe the usage of AI generated tags, we find that users writing more negative reviews are more likely to adopt AI generated tags. Furthermore, we find a significant bias in usage of negative AI generated tags. Finally, we observe that the usage of AI generated tags leads to reduction in review length, increase in novelty, however, the helpfulness votes of the reviews decrease.</p>
16	<p data-bbox="271 850 2042 911">Collaborative success in online communities: Explaining peer production performance in the 2022 R/place event</p> <p data-bbox="271 932 1167 960">Shizhen Chen, Tian Heong Chan and Anandhi Bharadwaj</p> <p data-bbox="271 999 2042 1123">Online communities are open collectives of individuals whose members are often anonymous and communicate in an environment constrained by technological mediation. The loose-knit nature of online communities coupled with communication dominated by asynchronous texts, create significant difficulties in engaged collaborative work. Yet, individuals in online communities have shown a tremendous capacity for successful collaboration.</p> <p data-bbox="271 1144 2042 1339">In this paper, we empirically investigate the collaborative mechanisms in Reddit's r/place 2022 event. R/place was a "social experiment" held over four days from April 1, 2022, to April 4, 2022, that involved a large white online canvas wherein any registered user could edit the canvas by changing the color of a single pixel and replacing it with a pixel of a different color. Figure 1 shows a snapshot of the canvas hours after the event started. By having a cooldown timer that prevents any single user from continually placing pixels on the canvas, the event emphasizes collaborative effort, offering a unique opportunity to study how online communities collaborate in a highly novel setting.</p>

Figure 1: A snapshot of the 2022 r/place canvas hours after the event began.

Our theoretical and empirical investigation focuses on how two features of online communities affect their ability to coordinate. The first feature is based on what the community members talk about. We theorize that communities with members that talk about similar topics—an indication of shared interests—should perform relatively better compared to communities with members that talk about divergent topics. On the other hand, the second feature is based on who does the talking. Here, we theorize that communities with a hierarchical communication structure (that is, the presence of a single person that commands the attention of the community) should perform relatively better compared to communities with a flat communication structure (where attention is spread over more people).

We empirically investigate our claims using a novel dataset collected and merged from various online data sources. Reddit published the r/place pixel placement data after the conclusion of the event. Importantly, the volunteering spirit of these online users afterwards helped generate the place atlas and r/thefinalclean, two projects that helped identify (1) community ownership of each region on the canvas, and (2) the intended artwork of each community. We combined these datasets to create a time-varying dependent variable, capturing the rate at which a community moves towards its goal—the number of pixels placed by a community that matches its intended artwork in 15-minute time intervals. Our final dataset composes a panel of 261,961 observations that tracks the activities at 15-minute intervals over 1,377 subreddits.

The primary dataset that we used to construct our independent variables comes from community posts over all Reddit communities that participated in the event, from October 31st, 2021, to March 28th, 2022 (a period of five months immediately preceding the r/place event).

To measure the degree to which the members in a community talk about similar topics, we first constructed a word embedding model trained on the entire five months' worth of post across all Reddit communities in the sample. This results in every word used in the online conversations being placed in a semantic vector space, whereby more similar words are nearer to each other. For each community, we then calculate the generalized variance in the distribution of words in that embedding (intuitively, the measure captures the volume of the word cloud generated by the community, with higher volume indicating lower shared interests). To measure the degree to which communication in a community is hierarchically structured, we first calculate the number of comments in response to posts to each person within the community. We then examine the distribution of this number across individuals in the community—we measure hierarchical structure by the variance of this distribution (intuitively, high variance indicates the presence of outlier individual(s) that dominate the community's attention).

Using an empirical model that controls for many measurable features of online communities and individual subreddit-level unobserved effects, we show that communities with higher levels of shared interests and following a more hierarchical communication structure do perform relatively better on r/place.

Importantly, in the middle of the event, Reddit unexpectedly announced two expansions, each of which doubled the size of the canvas. The expansions resulted in sharp exogenous drops in the competition intensity between online communities for space, allowing us to investigate if the two features we investigate yielded benefits as competition intensity changes. We further show that the performance advantage of communities with higher levels of shared interests performs equally well in situations of high as well

	<p>as low competition intensity. However, the same is not true for hierarchical communication structure—online communities with hierarchical communication structures perform well under low competition intensity, but their performance degrades significantly under high competition intensity, leading to them underperforming relative to communities with flatter communication structures. We discuss theoretical and managerial implications.</p>
17	<p>Will my startup get funding?: A Machine Learning Model to predict startup success Anupam Purwar and Bhagwan Chowdhry</p> <p>We develop a machine learning model that predicts whether a startup that has secured no or some initial (seed/angel) funding will get the next round of investment. Previous work on this topic has primarily attempted mining databases of the startup ecosystem consisting of investors, incubators and startups. Also, use of web-based open sources to derive signals for the startup success prediction task has been investigated, but signals picked were limited to direct mentions in the form of web links. We gathered data of 5000+ startups over a period of 4 years. We used structured data about the startup founders, team size, product offering, industry sector, funding raised so far, as well as textual information about their company. Using these data fields, we trained a machine learning model with multiple features using random forest technique. We show that utilizing companies' textual data on their product offering/business plan provides a substantial performance boost in comparison to only using structured data about the startup ecosystem. We provide an in-depth analysis of the developed machine learning (ML) model as well as different ML techniques experimented to provide a quantitative comparison among them. This has helped us obtain insights into easily discoverable quantitative signals as well as market factors influencing the startup funding.</p>
18	<p>Imagine! Never write, type or punch your password: A novel Zero-Knowledge Proof protocol for passwords Bhagwan Chowdhry and Vasundhara Sharma</p> <p>Password hacks have become an increasingly challenging security concern for online engagements. The use of passwords for authentication continues to be popular despite their weaknesses and potential for hacking. We propose a novel protocol that is (i) easy to implement (ii) easy for customers to understand and adopt and (iii) relatively robust against hackers and eavesdroppers.</p>
19	<p>Unveiling the complementarities among digital innovations using semi-supervised learning</p>

	<p>Balaganesh Chandran, R K Amit and R P Sundarraaj</p> <p>Recombining the existing technology helps the firm to foster high-impact innovations. In the era of digital ferment, choosing the right combination of technology to innovate is a significant issue in any manufacturing firm. The selection of complementary technology for recombining can resolve this issue. On the other hand, firms are facing a dilemma in selecting the path of technological diversity or proximity. Still, this is a grey area in the recombinant literature. Hence, we examine the existence of complementarities within and across digital innovations to find the best method of search in recombination. We consider three major digital innovations in manufacturing: the Internet of Things (IoT), Artificial learning (AI) and 3D printing. We gathered patent data to determine three-way complementarities using mixed methodology that builds on BERT (Bidirectional Encoder Representations from Transformers), Positive Unlabelled (PU)-learning-based classifier and econometric analysis. Forward citation is used as the proxy to measure the performance of the innovation. The developed integrated PU-learning-based framework can be used to identify the technological elements of the invention and facilitates the firms in patent landscaping and trend analysis. Knowing the complementary benefits of this study aids the industries and research centres in innovation portfolio management. Further, the overall framework can be used to identify the complementarities among any invention, irrespective of the field.</p>
20	<p>ARIMAX Model for Forecasting Maintenance Work (AMFM): A Multi-Stage Seasonal ARIMAX Model for Workorder Time Series Forecasting</p> <p>Anupam Purwar and Matthew Reimherr</p> <p>E-commerce business depends on smooth day to day functioning of it's warehouses/facilities. The functioning of these facilities depend on the health of material handling equipment. To keep these equipment healthy, these facilities employ the help of maintenance engineers who perform predictive/breakdown maintenance work. To ensure an effective maintenance operation necessitates efficient planning of maintenance work. For efficient planning of future maintenance work, one needs to have good estimates of the future maintenance work. Hence, we created time series of maintenance work (breakdown and miscellaneous) in terms of demand hours for every day/week. Next, we built several forecasting models and evaluated these models on the basis of forecasting accuracy metrics viz. Mean Absolute percent error (MAPE) and Root mean squared error (RMSE) to determine which modelling technique is most suitable. Seasonal ARIMA with exogenous variable (SARIMAX) was found to be the most suitable approach with additional hyper-parameters like training dataset length and training data window start/end. This paper discusses the details of this SARIMAX approach and the procedure used to identify the best facility specific SARIMAX model. The proposed solution provides forecasts using SARIMAX framework with an out of sample MAPE less than 30 percent and RMSPE less than 20.</p>

21	<p data-bbox="264 212 1500 240">A Design Science based Method for Sign Prediction of Edges in Networks</p> <p data-bbox="264 260 1061 288">Mukul Gupta, Samrat Gupta and Giri Kumar Tayi</p> <p data-bbox="264 328 2047 683">The relationships between objects in many real-world networks such as social networks, voter networks, and gamer networks have polarity. These networks are called signed networks due to the polarity of relationships. In signed networks, the relationships have positive/negative signs representing the friendship/enmity between objects. The sign information of relationships would be useful for various mining tasks to get rich insights and to support the decision-making process. However, several real-world networks are either scarcely labelled or have implicitly signed edges. To get the advantage of signed information, the signs of the unlabelled edges have to be predicted which is a challenging task. In this work, the problem of sign prediction for unlabelled edges in undirected signed networks is considered and addressed using a design science approach. The proposed method is based on graph augmentation followed by a graph regularized framework for information diffusion. The applicability of the proposed method for predicting trust/distrust in a network of reviewers created from IMDb platform establishes its utility. In addition, real-world signed networks from diverse domains are used for experiments to validate the effectiveness of the proposed method. The performance of the proposed method is compared with state-of-the-art methods thus establishing its superiority for sign prediction in networks.</p>
22	<p data-bbox="264 786 2047 847">Strawberry Or Vanilla This Week? How To Optimize Tailored Assortments For Variety-Seeking/Avoiding Consumers</p> <p data-bbox="264 866 1447 895">Sumit Kunnumkal, Dorothee Honhon, Ismail Kirci and Sridhar Seshadri</p> <p data-bbox="264 935 2047 1031">We consider the problem of a retail personalizing an assortment to a consumer who is variety-seeking or variety-avoiding, that is, less or more likely to buy the same product as in the previous period. We characterize the structure of the optimal assortment in single- and multi-period settings.</p>
23	<p data-bbox="264 1137 1886 1166">Impact of Sponsorship-Based Funding on Contribution Behaviour in Knowledge-Sharing Platforms</p> <p data-bbox="264 1185 1115 1214">Poonacha K. Medappa, Murat M. Tunc and Xitong Li</p> <p data-bbox="264 1254 2047 1377">This paper examines how the introduction of sponsorship-based funding for open-source software contributors alters their contributions in the knowledge-sharing ecosystems. While sponsorship support to the contributors is expected to increase their contributions to the host platform, it is unclear (a) how it would lead to reallocation of their efforts between core and non-core contributing activities on the host platform, and (b) whether the effort reallocation between core and non-core contributing activities</p>

	<p>spills-over to other related knowledge-sharing platforms where the contributors receive no funding. In this paper, we seek to answer the two research questions by leveraging a sponsorship based funding feature that was introduced by GitHub in May 2019. We collect contributor-level data from both the host platform (GitHub, or GH) and a related knowledge-sharing platform (StackOverflow, or SO). Our empirical results show that a contributor's sponsorship listing on GH has no effect on the core, development-related contributions, but only increases non-core contributions (maintenance related tasks). We attribute the effort reallocation between core and noncore contributing activities to the different (intrinsic vs. extrinsic) motivations that drive the core and non-core contributing activities. Furthermore, we find evidence of effort distortion that leads to a negative spillover effect on contributions that occur on SO. Interestingly, we find that this negative spillover effect on SO is restricted to core contributing activities (i.e. answers), resulting in a similar effort reallocation pattern where the sponsored contributors shift their efforts relatively towards non-core contributing activities. We discuss the important implications of these findings to platform owners and contributors in the online communities.</p>
24	<p>From Seeker to Provider: Role Switching in Online Support Communities</p> <p>Michael Lee, Xunyi Wang, Ruochen Liao and Rajiv Kishore</p> <p>Online health communities offer a platform for patients to seek support from other users with similar conditions. As those support seekers' situation improves, they may ultimately want to give back to the community and provide support to others. This study explores how changes in seekers' ability, operationalized using their cognitive content and social network structure, can influence the seekers' propensity to switch roles and become support providers. We leverage the Dual Perspective Model of Agency and Communion and Social Networks literature to theorize a seeker's ability to provide support. We apply survival analysis to study seekers' first role switch from seeker to provider as well as these participants' continued providing behavior in the community. We find that while seekers who have higher positive agency and closeness centrality are more likely to switch roles to a provider, those having higher positive communality, degree centrality, and eigenvector centrality are less likely to switch.</p>
25	<p>Does Employer-Paid, Job-Protected Maternity Leave Help or Hurt Female IT Workers? Evidence from Millions of Job Applications</p> <p>Sofia Bapna and Russell Funk</p> <p>Paid maternity leave is an important pathway to increasing women's participation in information technology (IT) because it helps address work and family balance concerns, which are pronounced in the field because of long hours, travel, and the need to keep up with technology. Many governments require that employers pay for maternity leave. Using micro-level data on millions of applications from an online job portal, our study examines whether a policy change in India that expanded employer-paid, job-protected maternity</p>

	<p>leave from short-term to medium-term (12 to 26 weeks) changes companies' likelihood of considering female applicants to IT jobs and women's likelihood of applying to such jobs. We find that, post-legislation, there is a 59% decrease in the probability of women being invited for interviews. Additionally, we find a 12% increase in the likelihood of an female applicant being female. If we were to apply these effect sizes in the period before the law, the net effect, after taking into account the increase in female applicants, would be that 34% fewer women would be invited to an interview. Our study identifies an important source of leakage in the pipeline for female IT workers and may inform policies aimed at reducing gender gaps in women's representation in IT.</p>
26	<p>The Impact of Enforcing COPPA Privacy Policy on the Presence of Commercial Content on YouTube Kids Sumeet Kumar, Ashique Khudabukhsh and Deepa Mani</p> <p>YouTube Kids (YTK) was launched in 2015 as a safer alternative for children to watch videos. The videos on YTK are a subset of videos on YouTube and are expected to be free of sexual, nudity, and violent content. While among the most popular applications for children, the YTK platform has stayed in the news for many reasons, including privacy violations, hidden product promotions, and poor-quality content. In this research, we study the impact of enforcing a privacy protection policy (COPPA) on content quality, particularly the commercial intent in videos. While the COPPA decreased personal data collection, we find that enforcing the policy was followed by an increase in brand presence in videos and links to commercial URLs and a decrease in promotion disclosures. Enforcing COPPA might have reduced targeted advertisements to kids; however, the increase in the commercial intent and the reduction in promotion disclosure are determinantal for children's well-being.</p>
27	<p>A Platform's Dilemma in Controlling Marketplace Transactions Sumanta Singha, Rajib Saha and Abhijeet Ghoshal</p> <p>We study a two-sided marketplace with network effects, where buyers face congestion as they compete for the same resources the platform provides. To ease this congestion, the platform can exercise greater control on the marketplace or use pricing as a lever to limit the market size, thus containing the negative effect of congestion. However, this decision is not straightforward, as control is costly, and pricing can have unintended consequences. Using a game-theoretic model, we find a market in our setting can be characterized into two distinct categories based on the relative strength of buyers' same-side negative externality and cross-side positive network effect. This categorization helps platforms design tailor-made strategies based on the type of market it operates. Our study reveals several interesting theoretical findings. First, when buyers face relatively weak same-side negative externality, the platform can exert more effort to improve their experience even when effort becomes costlier. Second, it can decrease its efforts despite an increase in buyers' same-side negative externality. In either case, the platform profit is not necessarily hurt. We further</p>

	find that the platform can also lower its transaction fees even when the strength of the cross-side positive network effect increases.
28	Gender disparities in patent review outcomes are more pronounced for more novel patent applications Nikhil Madan, Deepa Mani and Rohin Nandkumar
	We examine how the magnitude of disparities against female inventors in patenting outcomes varies based on novelty of the patent application. An analysis of approximately 2.8 million applications reveals that gender disparities in terms of the likelihood that an application is granted a patent, the intensity of examination it is subject to, and, conditional on it being granted, the reduction in its scope, are systematically larger for more novel applications. In line with the possibility that these disparities are driven by differential treatment of applications authored by female inventors, we find that (a) female inventors of more novel applications are more likely to file requests for continued examination following rejections, and (b) the magnitude of these disparities and their variation along the novelty continuum is larger for applications authored by inventor teams with common (vs. rare) names. These findings suggest that the social and private costs of differential recognition of the work of female and male inventors are higher than previously believed.