All that Glitters is not Gold: Understanding the Impacts of Platform Recommendation Algorithm Changes on Complementors in the Sharing Economy

Jack Tong
Nanyang Technological University, jack.tong@ntu.edu.sg

Xiaowei Zhang
Tsinghua University, xw-zhang21@mails.tsinghua.edu.cn

Xueming Luo
Temple University, xueming.luo@temple.edu

Zhijie Lin
Tsinghua University, linzhj@sem.tsinghua.edu.cn

Jing Li
Nanjing University, jingli@nju.edu.cn

All that Glitters is not Gold: Understanding the Impacts of Platform Recommendation Algorithm Changes on Complementors in the Sharing Economy

Completed Research Paper

Siliang (Jack) Tong[*]
Nanyang Technological University
50 Nanyang Avenue, Singapore
jack.tong@ntu.edu.sg

Xiaowei (Elliott) Zhang[*]
Tsinghua University
30 Shuangqing Road, Beijing, China
xw-zhang21@mails.tsinghua.edu.cn

Xueming Luo
Temple University
1801 N. Broad Street, Philadelphia, USA
luoxm@temple.edu

Zhijie Lin
Tsinghua University
30 Shuangqing Road, Beijing, China
linzhj@sem.tsinghua.edu.cn

Jing Li
Nanjing University
22 Hankou Road, Nanjing, China
jingli@nju.edu.cn

Abstract

Sharing platforms often leverage recommendation algorithms to reduce matching costs and improve buyer satisfaction. However, the economic impacts of different recommendation algorithms on the business operations of complementors remains unclear. This study uses natural quasi-experiments and proprietary data from a home-cooked food-sharing platform with two recommendation algorithms: word-of-mouth recommendation (WMR) and botler personalization recommendation (BPR). Results show the WMR negatively affects revenue while BPR has a positive effect. The contrast revenue effects have been attributed to capacity constraints for complementors and matching frictions for consumers. WMR encourages sellers to specialize in high-quality products but limits new product development. BPR promotes innovation to suit diverse customer tastes but may reduce quality. This reflects the exploration-exploitation trade-off: WMR exploits existing competences, while BPR explores new products to satisfy personal preferences. The authors discuss implications for how to utilize recommendation algorithms and artificial intelligence for the prosperity of sharing economy platforms.

Keywords: Sharing economy, recommendation algorithm, word-of-mouth recommendation, botler personalization recommendation, specialization, new product offerings

[*] These authors contributed equally to this work.

Introduction

The sharing economy, also known as the "peer-to-peer economy" or "collaborative consumption", has penetrated into various aspects of people's online and offline lives, giving rise to many new forms of employment and stimulating new consumer demands. The disruptive growth of the sharing economy has led us to revisit classic problems in online platform research. Participation in sharing economy marketplaces by complementors and consumers is characterized by distinct and varying motivations and characteristics that distinguish them from traditional e-commerce platforms. As a result, some of the consensus on platform design in e-commerce platforms may be challenged (Filippas et al. 2020; Luo et al. 2021). This study is concerned with a classic problem of platform economy: the economic consequences of recommendation algorithms. In fact, recommendation systems are considered as infrastructure for online platforms. We believe the uniqueness of sharing economy may make the economic impacts of recommendation algorithms less obvious.

However, while a few studies have begun exploring the economic consequences of this unique platform ecology, there are still several challenges that limit our access to novel, fine-grained findings. First, while the sharing economy has changed the pattern of resource allocation in many sectors, it seems that most research has focused on services such as accommodation-sharing (e.g. Airbnb) (Zervas et al. 2017) and ride-hailing (e.g. Uber) (Babar and Burtch 2020), which largely limits our understanding of the large-scale heterogeneous complementors in different scenarios. Second, researchers often lack sufficient real-world conditions to influence platform policies, which makes it difficult to empirically test some theories in reality. Third, user privacy and platform barriers also limit current research from providing a dynamic and comprehensive assessment of the heterogeneous behavior of individual complementors.

Our paper circumvents these challenges by leveraging natural quasi-experiments and rich proprietary datasets. The setting here is a major home-cooked food-sharing platform, which matches small-scale entrepreneurs cooking meals at their own kitchens with nearby customers. Our data track the implementations of two algorithms that are pertinent for most sharing platforms: word-of-mouth recommendation (WMR) and botler personalization recommendation (BPR).[1] Essentially, WMR is the algorithm based on others' preferences in terms of consumer reviews (review rating stars). In contrast, BPR is the algorithm based on one's own preference and taste (personalization based on individual own browsing, searching, and purchasing behavior). Both of these algorithms are based on the naïve geo-fencing technology, but BPR goes further by incorporating more user information beyond electronic word-of-mouth (eWOM) to achieve personalized and precise marketing. These two algorithms are decided by the platform and **exogenous** to all sellers and all buyers, and our identification strategy is to rely on these exogenous events and natural quasi-experiments. In our setting, we have the following two research questions about the economic consequences of recommendation systems introduced by sharing platforms.

RQ-1 (Economic Effect): How do WMR and BPR algorithm changes *directly* affect the revenue of complementors on the sharing platform?

RQ-2 (Mechanisms): How do WMR and BPR algorithm changes affect complementor operations and ultimately have an *indirect* impact on economic revenues?

These two research questions may seem to have been discussed in traditional scenarios, but we are the first to analyze them systematically in the context of sharing economy, and we have also come to some logical but unintended conclusions. First, we find that WMR, compared to naïve geo-fencing recommendation, reduces complementors' revenues after controlling for other potential mechanisms, while BPR, compared to WMR, has a positive impact. Second, we find that the negative economic impact of WMR is likely to be achieved by curbing innovation by complementors, while the positive economic impact of BPR is likely to be through promoting innovation. Third, we find that while on average WMR reduces revenues of complementors, at the same time, it improves sellers' service quality, while conversely, BPR reduces sellers' service quality. This provides implications for the multi-objective decision making of platform managers and motivates us to adopt novel machine learning approaches to capture heterogeneous causal effects.

[1] Botler in this paper means the combination of robot and butler, an artificial intelligence algorithm implemented on the focal platform. This is inspired by the first robot butler (Botler) introduced by Starwood Hotel and Resort. https://techcrunch.com/2014/08/13/starwood-introduces-robotic-butlers-at-aloft-hotel-in-palo-alto/. Our data provider—the app platform decided to implement BPR because others' preference may not match individuals' own taste (e.g., despite high ratings, a five-star shrimp seller preferred by peers may not satisfy a buyer who likes beef).

Background and Conceptual Model

The Economic Effects of Recommendation Algorithms

Recommendation algorithms function as the pillars of interactive information technology systems that assist buyers in searching, screening, and evaluating alternative products or sellers that are available (Xiao and Benbasat 2007). Recommendation algorithms have been widely applied in e-commerce contexts (e.g., Yelp.com's "the best 10 restaurants" in a city or Amazon.com's "consumers who bought this item also bought"). In the online shopping environment, a large number of products or services available may also make it difficult for buyers to evaluate the underlying quality. Recommendation algorithms help customers to reduce search costs (Xiao and Benbasat 2007), explore new items (Resnick and Varian 1997), and sort through large category sets (Häubl and Trifts 2000).

However, as the recommendation algorithm is directly oriented towards searching and screening behavior on the demand side, its economic impact on the performance of the supply side (complementors) does not seem to be as evident. Bodapati (2008) demonstrated the causal impact of recommendation systems on sales revenues and suggested that the increase in sales can be viewed as the outcome of brand awareness and customer satisfaction. Pathak et al. (2010) found that recommendation systems have a positive impact on long-tail sales. Based on a field experiment, Lee and Hosanagar (2019) also showed that collaborative filtering algorithms on e-commerce platforms increased the absolute sales of various products, but also led to a decrease in sales diversity. Much of the evidence here seems to point that recommendation algorithms implemented by online platforms (at least e-commerce platforms) can effectively benefit sellers in addition to improving the buying experience for consumers. In recent times, a few scholars have begun to focus on the boundary conditions of these positive effects. Kumar and Hosanagar (2019) highlighted that recommendation algorithms on e-commerce platforms, while increasing the exposure of focal products, also simultaneously increase the visibility of alternative products, ultimately and unexpectedly causing the cannibalization of their own sales. The analytical model from Shi and Raghu (2020) found that conditional on quality, recommending high-taste-dispersion products may increase or decrease producer profits, depending on the joint effect of profit margin and purchase probability. However, this stream of research hardly ever discusses the economic effects of recommendation algorithms in the context of sharing platforms or gig economy. The specificity of both sellers and buyers on sharing platforms motivates us that the economic impact of recommendation algorithms may be more complex and confusing than that of traditional e-commerce platforms.

The Revenue Effects of WMR and BPR in the Sharing Economy

We focus on two types of recommendation algorithms on sharing platforms, namely WMR and BPR. WMR emphasizes the "wisdom of the crowd", which uses previous user-generated rating data as the basis for recommendations, is a typical type of non-personalized recommendation. Informal word-of-mouth (WOM) recommendation has been validated to have important functions in information acquisition and persuasion (Berger 2014). Therefore, designing recommendation algorithms based on integrated electronic word-of-mouth (eWOM) has become an important technical means for platforms (Verma and Yadav 2021). BPR, on the other hand, emphasizes personalized recommendations for consumers based on their historical browsing, searching, and purchasing behavior. We believe that there are non-negligible differences between sharing economy platforms and traditional online platforms, which leads us to propose two presumptions, providing a theoretical guide to our analysis of the direct economic impact of WMR and BPR.

From a complementor's perspective, we propose the "*Constrained capacity presumption*". In sharing economy markets, complementors are often non-professional individual participants who leverage their existing assets or resources to provide products or services and participate in market transactions. These supplies are usually not non-scale free (Levinthal and Wu 2010), so the complementors often face various extents of capacity constraints. Constrained capacity leads to constrained supply, which may further lead to constrained economic revenue. However, studies on the capacity constraints of complementors in the sharing economy and the potential economic impact remain conspicuously absent (Tae et al. 2020). Based on this presumption, we believe that WMR may not be as effective in promoting complementors' revenue as it has in the past, and may instead, on average, cause harm to their revenue. WMR directly promotes the exposure of highly-rated sellers, but for these highly-rated sellers, even if they are not recommended by the algorithm, they seem to have already achieved considerable sales due to their good reputation (Abrate and Viglia 2019). Increasing exposure may seem to attract a large volume of demand to these highly-rated sellers, but capacity constraints may result in them not being able to handle most of the new incoming demand, and their sales may not see a significant increase (we call it the "*demand spillover issue*"). Instead, when some consumers who pursue specific tastes turn to these highly-rated sellers due to the increase in search costs, these sellers may find it more difficult to match with consumers who have a higher willingness to pay (Adomavicius et al. 2018), which would negatively affect their average order value. For those lowly-rated sellers, especially new sellers on the platform, they already have an information disadvantage as latecomers to the market (Kerin et al. 1992), and WMR further reduces their exposure. In a single-category sharing market (such as accommodation-sharing platforms and food-sharing platforms), the lack of category information can further abate the number of recommended complementors and exacerbate the negative impact of WMR. On the other hand, BPR, which considers personalized consumer needs, will allow more complementors with average ratings but unique products to appear on some consumers' result pages. This not only disperses excessive demand for highly-rated sellers, but also increases the sales of many lowly-rated sellers. Therefore, we expect BPR to have a positive impact on the seller's revenue.

From a consumer's perspective, we propose the "*Matching friction presumption*". In addition to the uniqueness of complementors, consumers on sharing platforms also appear to differ from those on traditional e-commerce platforms. Their motivation for participating in sharing markets is attributed to their demand for home benefits, social interaction, cultural experience, etc (So et al. 2018). These highly heterogeneous consumers have richer transaction motivations, and therefore, product quality is usually not their only pursuit: for example, in the case of food-sharing platforms, if a consumer only pursues high-quality products or services, they can just dine at a nearby restaurant instead of choosing a home chef. Under this presumption, we believe that the WMR algorithm that only emphasizes product quality may overlook consumers' demand for other product/service features on sharing platforms, and if these demands are crucial to consumers, WMR may actually exacerbate the matching friction between buyers and sellers, resulting in the change in consumers' purchasing behavior (Xu 2020). In contrast, BPR can clearly provide more private and personalized recommendations for consumers, similar to the role played by recommendation algorithms on traditional e-commerce platforms. As Li et al. (2022) pointed out, the presence of personalized recommendations on E-Commerce platforms increases both the breadth and depth of consumers' consideration, resulting in increases in both absolute sales and consumers' basket value. Hence, we expect BPR to effectively reduce the search costs for consumers on sharing platforms, alleviate matching friction, and ultimately promote seller revenue. Based on the above discussion, we propose the following two hypotheses.

Hypothesis 1 (H1). *Compared to naïve geo-fencing recommendation, the implementation of WMR has a negative impact on the revenue of complementors on the sharing platform.*

Hypothesis 2 (H2). *Compared to WMR, the implementation of BPR has a positive impact on the revenue of complementors on the sharing platform.*

The Pathways for the WMR and BPR Effects

In the previous subsection, we have discussed the *direct* impacts of algorithms. These impacts are a result of algorithms directly influencing the matching process, and thus they do not depend on sellers' responses. However, the impacts of algorithms may be more complex. In this subsection, we further discuss the *indirect* effect of the algorithm changes on sellers' revenues. We believe that sellers' strategic responses can be important intermediaries in the economic impacts of the algorithm changes. An analytical model from Zhou and Zou (2022) has shown that recommendation systems not only directly affect consumers' utility, but also stimulate sellers to strategically adjust prices to compete for platform recommendations. Unlike the two presumptions we proposed earlier, changes in seller revenue at this point are essentially the result of their adaptive strategies in response to algorithm changes. Specifically, prior research has identified two key pathways to the success of businesses: *specialization* with high quality of current products (Chevalier and Mayzlin 2006) and *innovation* with new products (Ernst et al. 2010; Lee and Hosanagar 2019). We believe that they may serve as partial mediation variables between algorithm changes and sellers' revenues.

The positive impact of specialization on sales revenue has been repeatedly validated in the literature, and there is no reason to deny this positive effect on sharing platforms, although the heterogeneity of consumers and complementors may make it a bit more complicated (Rossi 2023). High review ratings usually signal high quality (Chevalier and Mayzlin 2006). Thus, it is feasible that as the WMR algorithm helps a buyer to find sellers with high review ratings more easily by prioritizing them on the top of the recommendation list, sellers could be incentivized to adopt a specialization focus on improving the review quality of current products to drive the demand and sales revenues. We expect that, despite the overall negative impact of WMR on revenue, it will incentivize complementor to achieve quality improvements, thereby mitigating the negative revenue impact. On the other hand, innovation (new product development) is also an important strategy for sellers to attract potential demand and achieve revenue growth (Ernst et al. 2010). Customers often prefer personalized recommendations and appreciate products matching their own preferences (Bodapati 2008; Häubl and Trifts 2000). Hence, it is likely that as BPR enables the buyer to find sellers with more customized cuisines by prioritizing them on the top of the recommendation list, sellers may respond by adopting an innovation focus on introducing more new products to account for the diverse customer tastes on the platform. We expect that BPR will incentivize sellers to increase their sales diversity, thereby further boosting their revenue. Hence, we propose the following two hypotheses.

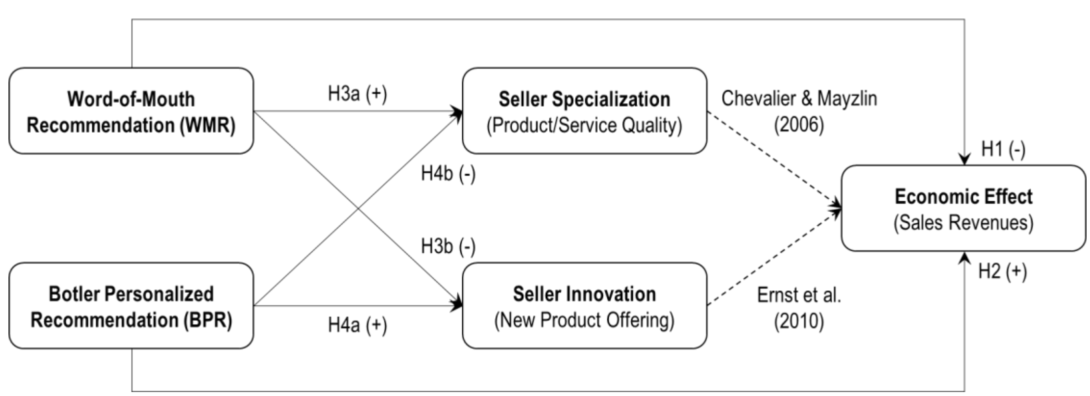
Hypothesis 3a (H3a). *Compared to naïve geo-fencing recommendation, the implementation of WMR has a positive impact on the specialization of complementors on the sharing platform.*

Hypothesis 4a (H4a). *Compared to WMR, the implementation of BPR has a positive impact on the innovation of complementors on the sharing platform.*

Whereas we have predicted the positive impacts that WMR and BPR may have on sellers' strategies, some potential unintended effects are also worth noting. On the one hand, it is possible that when sellers are incentivized by WMR to become more specialized with high quality of current products, they might be less willing to develop new products. This is because new dishes are uncertain *ex ante* and may not always get good review feedback from buyers (which might discourage sellers to introduce new dishes). Models proposed by Fleder and Hosanagar (2009) also suggested that a recommendation system based on historical ratings can cause a decline in sales diversity. Especially for consumers looking to discover new products, this negative impact can be alarming (Skiti et al. 2022). On the other hand, it is plausible that when sellers are incentivized by BPR to become more innovative, their rating quality might suffer. This is because being innovative is not without costs and sometimes leads to failure and poor performance (Henderson and Clark 1990), and the newly developed dishes may not turn out to be popular enough and thus drag down the sellers' quality reputation. In addition, the non-specialist nature and capacity constraints of sharing complementors make it more difficult for them to enhance innovation while balancing product quality well, which undoubtedly adds to their woes. Hence, we propose the following two hypotheses. To sum up, our complete conceptual model is compiled in Figure 1.

Hypothesis 3b (H3b). *Compared to naïve geo-fencing recommendation, the implementation of WMR has a negative impact on the innovation of complementors on the sharing platform.*

Hypothesis 4b (H4b). *Compared to WMR, the implementation of BPR has a negative impact on the specialization of complementors on the sharing platform.*



Notes. We do not hypothesize the dotted lines, which relate the direct effects of service quality and product innovation on sales revenues, because these are relationships that have been studies previously. The effect of WMR and BPR are incremental to the geo-fencing baseline recommendation.

Figure 1. Conceptual Model

Sharing Platform Setting and Data

Platform and Recommendation Algorithm Changes

We obtained a rich proprietary dataset from a major food-sharing platform in Asia. The app-based platform was launched in September 2014, and is the largest platform for home-cooked food-sharing business in the local market, with over 1 million active sellers and 10 million active buyers each month. It is a two-sided platform, matching supply-side complementors who cook dishes at their home kitchens with the demand-side nearby buyers who order meals. Different from UberEats and Grubhub whose sellers are professional restaurants, the food-sharing platform hosts amateur individuals, or small-scale entrepreneurs. These individuals exploit under-utilized kitchen resources and home cooking skills for business exchanges, a key feature of sharing economy (Zervas et al. 2017). This platform is similar to other sharing economy platforms such as Airbnb, Uber, TaskRabbit and Skillshare. Specifically, they are all matchmakers with both high platform intermediation and high consociality (Perren and Kozinets 2018). The platform also has several unique and interesting features. For example, the platform enables sellers to provide dishes with diverse flavours, prices, quality levels and inventory availability, and independently formulate their operation strategies. The dish services are displayed at the street level, and most buyers order food more frequently.

The food-sharing platform implemented two algorithm changes (WMR and BPR) on top of the baseline geo-fencing matching algorithm. In the baseline, the platform uses geo-fencing mobile technology and links sellers and buyers based on their location proximity.[2] As the platform collected more demand-side data on customers' review ratings or peer preferences of historical purchases over time, it decided to launch a rating-based new algorithm of WMR on July 1st, 2016. The implementation of WMR on top of geo-fencing allows highly-rated nearby sellers to be displayed on the app main screen by default. As WMR can recommend sellers based on review ratings or "wisdom of the crowd," it was intended to improve customer satisfaction by prioritizing highly-rated nearby sellers to the buyer. At the same time, WMR also takes into account the real-time location of consumers and the geographic distribution of merchants. When multiple sellers with the same rating appear within the proximity range of the consumer, WMR will recommend the seller that is closest in geographic distance. In other words, when we compare the performance of sellers during the implementation of naïve geo-fencing and WMR algorithms, we are actually focusing on the incremental impact that arises from the introduction of eWOM.

Three months later, on October 13th, 2016, the platform implemented a personalization algorithm of BPR based on individual buyers' past browsing, searching, and purchasing behaviors on the app. This algorithm involved personal data on each individual buyer's food searching behavior, preference of cuisine, dish price, order time of the day, and purchase recency, frequency, and quantity. Thus, it offers personalized recommendations for each buyer. After extensively testing and adjusting the IT systems, the app platform confirmed that BPR indeed worked and can recommend personalized dishes according to each individual's personal food preference. BPR algorithm was launched on top of WMR and geo-fencing baseline to assure that the best-fit nearby sellers with high ratings who can satisfy their personal preference are displayed on the app main screen by default. BPR was deemed necessary because others' preference may not match individuals' own taste (e.g., despite high ratings, a five-star shrimp seller preferred by peers may not satisfy a buyer who likes beef). In other words, when we compare the performance of sellers during the implementation of WMR and BPR algorithms, we are actually focusing on the incremental impact that arises from the introduction of additional user information beyond eWOM.

These two recommendation algorithms of WMR and BPR could be generalizable to other popular e-commerce platforms (Insider 2016; Wired 2017). For example, on Amazon.com, products are recommended by what other customers purchased ("Trending Deals") based on collaborative filtering techniques. Netflix recommends videos based on both nuanced threads from viewers' past watching behaviors and popular videos watched by others, and the algorithm influences 80% of Netflix video watching (Wired 2017). Recently, driven by artificial intelligence and cloud computing technologies, personalization algorithms have been developed to provide effective matching to satisfy customer needs

[2] The food-sharing platform started the business with the location proximity recommendation algorithm as the baseline for several reasons. First, it aims to provide convenience to buyers who order take-out food, and location proximity will save waiting time for buyers (and better tastes of the timely delivered food). Second, due to the small size of mobile phone screens, it can only recommend a handful sellers from a large number of options. Third, it is challenging for the platform to implement more advanced algorithms in the nascent stage due to data deficiency on consumer reviews and purchase records (cold start problem).

based on individuals' personal preference and taste. These two algorithms are important for sharing platforms, which host heterogeneous groups of buyers and sellers. Without WMR, buyers will face higher uncertainty and spend much time to explore sellers with diverse reputation. Without BPR, buyers might find a nearby top-rating seller who still cannot suit their own tastes.

Because these two algorithms were engineered and launched solely by the platform, there was no pre-announcement. The platform confirmed that it was impossible for sellers or buyers to know the new algorithm *ex ante* and to perform any strategic behaviors that may confound our results of the algorithm impact. In this sense, the changes of WMR and BPR were **exogenous** shocks to both sellers and buyers. Once implemented *ex post*, the nature of the algorithms was highlighted on the top of landing pages and could be easily identified by both buyers and sellers. The platform executives also affirmed that there were no other major functional changes on the platform during the launches of these two algorithms.

Data Descriptives

The food-sharing platform provided us a large dataset on a randomly selected sample of sellers over a 9-month period from April 2016 to December 2016. The data is at the seller-daily level, and all our subsequent analyses keep the panel size and buyer composition the same so that our results would not be affected by the exit or entry of sellers on the platform (Sun et al. 2017). In other words, the market structure in our data is stable. Sellers are randomly sampled from the food-sharing platform's census of sellers. We use a random number generator via the *RANUNI* function and sort the random numbers in sequence, from which a sample is extracted (Andrews et al. 2016). Our panel of sellers has been on the platform for the whole time period covering the WMR and BPR platform algorithm changes. As a result, the individual daily level panel data includes the same 7,883 sellers with 1,478,840 observations for WMR and 1,450,413 observations for BPR. Figure 2 summarizes the timeline of the platform algorithm changes and our observation windows.

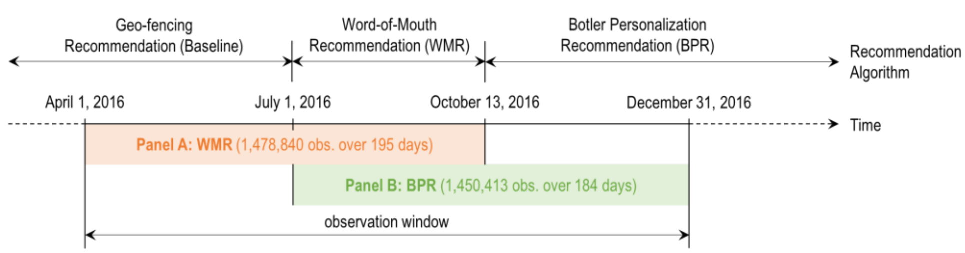


Figure 2. Timeline of the Algorithm Changes and Sample Selection

Table 1 presents the descriptive summary of the data. On average, each seller receives 38.73 RMB and 33.05 RMB daily sales revenue, 3.26 and 3.28 stars of review rating in Panel A and Panel B respectively. A strong majority (more than 70%) of sellers on the platform are female, with an average age of 41 years old. These statistics are consistent with the nature of the food-sharing platform, where most cooks are middle-aged housewives who spend their spare time preparing and selling home-cooked dishes for additional incomes.

Variables	**Description**	**Panel A: WMR** (195 days; 7,883 sellers; 1,478,840 obs.)				**Panel B: BPR** (184 days; 7,883 sellers; 1,450,413 obs.)			
		Mean	**Std. Dev.**	**Min**	**Max**	**Mean**	**Std. Dev.**	**Min**	**Max**
WMR	Word-of-mouth recommendation	0.55	0.50	0.00	1.00	-	-	-	-
BPR	Botler personalization recommendation	-	-	-	-	0.44	0.50	0.00	1.00
S_REV	Daily sales revenue	38.73	163.11	0.00	25800.00	33.05	160.52	0.00	10960.00
S_RVW_CNT	Number of reviews	151.75	336.70	0.00	5005.00	172.76	394.99	0.00	5981.00
S_RVW_STR	Number of review stars	3.26	2.28	0.00	5.00	3.28	2.26	0.00	5.00
S_RVW_STR_SD	Standard deviation of review stars	0.12	0.43	0.00	3.54	0.13	0.40	0.00	3.54
S_DSH_NEW_CNT	Number of new dishes	0.03	0.28	0.00	34.00	0.01	0.19	0.00	25.00
S_DSH_CNT	Number of dishes	18.30	19.60	1.00	206.00	19.41	21.38	1.00	281.00
S_DSH_PRC	Average dish price	22.67	12.75	1.00	600.00	22.86	12.99	1.00	600.00
S_CPN	Coupon value offered	5.61	30.83	0.00	1504.00	4.54	29.53	0.00	2145.00
S_DEL_RDS	Delivery radius	2129.67	1297.51	500.00	10000.00	2152.71	1315.47	500.00	10000.00
S_DEL_FEE	Delivery fee	2.71	0.75	0.00	5.00	2.68	0.75	0.00	5.00
S_FTR_CNT	Number of kitchen features	1.76	1.06	0.00	7.00	1.74	1.06	0.00	7.00
S_MAL	Seller gender: Male	0.28	0.45	0.00	1.00	0.29	0.45	0.00	1.00
S_AGE	Seller age	41.58	12.09	20.00	60.00	41.44	12.11	20.00	60.00
S_COM_CNT	Number of competitors	2253.40	1394.49	0.00	4090.00	2495.54	1549.30	0.00	4519.00
B_HAS_GEN	Prop. of buyers who disclosed gender	0.14	0.00	0.13	0.15	0.14	0.00	0.13	0.14
B_HAS_AGE	Prop. of buyers who disclosed age	0.27	0.05	0.16	0.33	0.31	0.02	0.27	0.33
P_PST_CNT	Number of platform WeChat promo	0.74	0.82	0.00	3.00	0.63	0.76	0.00	3.00

Table 1. Descriptive Statistics

Preliminary Analyses

Before formally introducing our parametric models, we provide an overview of the theoretical framework that serves as the basis of our identification strategy. Given that our algorithm changes are market-wide and that all sellers are subject to our natural experimental interventions, we rely on the regression discontinuity in time (RDiT) framework for trustworthy identification.

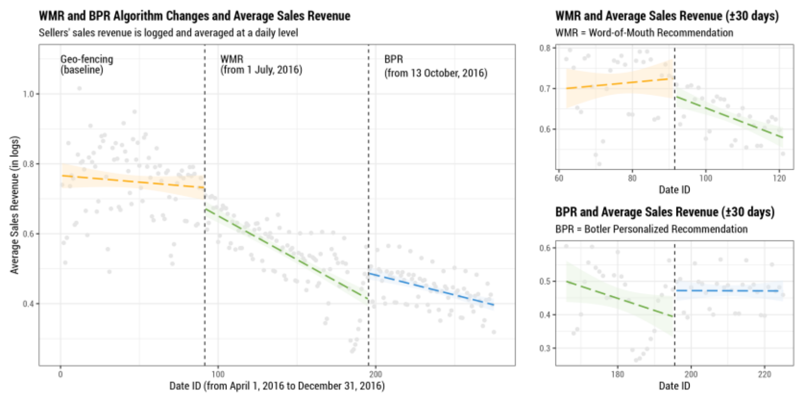


Figure 3. WMR and BPR Algorithm Change and Average Sales Revenue

RDiT is a special type of the regression discontinuity (RD) framework where time serves as the running variable, and treatment begins at a particular threshold in time (Davis 2008; Hausman and Rapson 2018). In essence, RDiT believes that during a narrow time period surrounding the algorithm change, unobserved factors would remain constant. In our design, the time trend captures three common underlying evolving trends before algorithm change, and the estimated post-change trend would mimic well the counterfactual. Thus, the data observations before an algorithm change can act as a valid control group for comparison. On the other hand, RDiT can also be seen as an extension of the RD method in interrupted time series analysis (ITSA). In previous research in medical and public health fields (Bernal et al. 2017), researchers typically used a naïve ITSA design to study the causal effects of policy shocks on individual time series at a particular time point. In our empirical strategy, a panel-level RDiT-based quasi-natural experimental design is used to integratedly analyze the average treatment effects of the algorithm changes on a large number of sellers. Our design also eliminated other potential mechanisms through techniques such as covariate adjustment and panel regression methods to make our causal analysis more credible and robust.

To provide a preliminary and intuitive demonstration of the average treatment effect before presenting our parameterized model, we show the daily-level RDiT diagrams in Figure 3. We observe a typical time trend in seller revenue in both panels: the average revenue of platform sellers shows a significant downward trend over time, which may be due to a competitive crowding effect (Rietveld and Eggers 2018). Both the two algorithm changes have a significant impact on this trend: the implementation of WMR decreases seller revenue and accelerates the decrease in seller revenue over time; conversely, the implementation of BPR increases the platform revenue of sellers and slows the decrease in seller revenue over time. Consider deploying the augmented local linear methodology to increase the power of the local linear specification, the right-hand time series subplots of Figure 3 present sellers' average daily sales revenue before (-30 days) and after (+30 days) the algorithms implementations. Consistent with the global linear specification, there is an apparent jump (discontinuity) in sales revenue in the post-implementation period compared with the before-implementation period for both WMR and BPR implementation. These graphs provide preliminary evidence for the revenue effects of WMR and BPR.

Model and Results

The Revenue Effects of WMR and BPR

We next model how the algorithm changes influence seller revenues on the platform, after accounting for a host of alternative explanations at the sellers, kitchens, dishes, buyers, time trend, platform, and competition levels with panel data. Specifically, we employ the panel fixed effects modeling approach to estimate the effect of WMR and BPR on seller's revenue (in logs) as specified in Equations (1) and (2).

$$\ln(1 + S\_REV_{i,t}) = \alpha_{o,i} + \alpha_1 WMR_t + \gamma_1 T_t + \phi_1 Z_{i,t} + \mu_{1,i,t} \quad (1)$$

$$\ln(1 + S\_REV_{j,t}) = \beta_{o,j} + \beta_1 BPR_t + \gamma_2 T_t + \phi_2 Z_{j,t} + \mu_{2,j,t} \quad (2)$$

Let subscript i and j denote each individual seller in our dataset, and subscript t denote each day in the data period of WMR or BPR. The dependent variable, $S\_REV_{i,t}$, indicates the seller i's sales revenue on day t. The independent variables, WMR_t and BPR_t are binary indicators for the WMR and BPR algorithm pre-implementation period and post-implementation periods, respectively. That is,

$$WMR_t = \begin{cases} 0 & \text{from Apr. 1 to Jun. 30, 2016} \\ 1 & \text{from Jul. 1 to Oct. 12, 2016} \end{cases}, \qquad BPR_t = \begin{cases} 0 & \text{from Jul. 1 to Oct. 12, 2016} \\ 1 & \text{from Oct. 13 to Dec. 31, 2016} \end{cases}$$

Thus, the effect of WMR is incremental to the naïve geo-fencing algorithm and the effect of BPR is incremental to WMR. The coefficients α_1 and β_1 capture the effect of WMR and BPR algorithm changes on individual sellers' sales revenue. $\alpha_{o,i}$ and $\beta_{o,i}$ capture unobserved seller-specific effects, and $\mu_{1,i,t}$ and $\mu_{2,i,t}$ indicate the residual random error term. The error term here can be considered the grand statistical noise ($\mu_{1,i,t}$ or $\mu_{2,j,t}$) plus the individual level random intercept ($\alpha_{o,i}$ or $\beta_{o,j}$). To rule out alternative explanations, we have a comprehensive set of factors as the control variables. $Z_{i,t}$ and $Z_{j,t}$ are vectors containing time-variant variables: sellers' rating level (from 0 to 5 stars), volume, and dispersion (standard deviation), kitchen features, food delivery distance, delivery fee, number of new dishes, number of dishes, dish prices, coupon value, competition defined as the number of rival sellers in seller i's city district by zip code, and the number of social media (WeChat) postings generated by the platform. T_t contains the time trends, which account for platform market growth and saturation that are related to the platform evolution and time dynamics, and day level dummies to control for the effects of seasonality and holidays. We strive to derive more reliable estimates with this comprehensive set of control variables. Again, in our model, the platform algorithm change is **exogenous** to (unexpected by) both sellers and users, so the estimates here are not endogenous but rather possible effects with causal inference (Davis 2008).

Variable	(1) WMR FE	(2) WMR RE	(3) WMR Tobit	(4) WMR Panel Tobit	(5) BPR FE	(6) BPR RE	(7) BPR Tobit	(8) BPR Panel Tobit
WMR (Word-of-mouth recommendation)	-0.026*** (0.004)	-0.026*** (0.004)	-0.021*** (0.003)	-0.026*** (0.003)				
BPR (Botler personalization recommendation)					0.078*** (0.006)	0.078*** (0.006)	0.065*** (0.005)	0.078*** (0.004)
Constant	0.362 (0.607)	-0.848 (0.639)	0.621 (0.657)	-1.028*** (0.351)	1.309 (1.167)	-0.051 (1.047)	0.595 (1.079)	-0.118 (0.739)
Control variables	YES	YES	YES	YES	YES	YES	YES	YES
Time trends	YES	YES	YES	YES	YES	YES	YES	YES
Seller effects	YES	YES	NO	YES	YES	YES	NO	YES
Number of sellers	7,883	7,883	7,883	7,883	7,883	7,883	7,883	7,883
Number of observations	1,478,840	1,478,840	1,478,840	1,478,840	1,450,413	1,450,413	1,450,413	1,450,413
Hausman test	χ^2=12,812.39, p<0.001		-		χ^2=10,495.25, p<0.001		-	
LR test	χ^2=8.0e+05, p<0.001		χ^2=8.0e+05, p<0.001		χ^2=9.4e+05, p<0.001		χ^2=9.4e+05, p<0.001	
(Pseudo) R^2	0.2018	0.3861	0.1785	0.0771	0.1594	0.3921	0.1823	0.0672

Table 2. The Impacts of WMR and BPR on Seller Revenues

Note: Clustered robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

The estimation results are shown in Table 2 Columns (1) and (5). Considering that the error terms of sellers within the same district may be correlated, the standard errors are clustered at the district level. As indicated, after controlling for other potential trends, the effects of WMR and BPR are respectively significantly negative and positive, which are both consistent with our hypotheses H1 and H2. We calculated that on average, the WMR algorithm change decreases seller revenues per day by 2.57% ($e^{-0.026} - 1$) and the BPR algorithm change increases seller revenues per day by 8.11% ($e^{0.078} - 1$). To better account for individual unobserved heterogeneity, random effect models incorporating the time-invariant seller characteristics are also employed. Results in Table 2 Columns (2) and (6) suggest consistent results. To ensure consistency of estimators, we use the fixed effects models in the subsequent analysis, which is also supported by the Hausman tests. Furthermore, because of the left censoring of revenue, we conducted analyses with cross-sectional and panel Tobit models. The estimates also suggest consistent results.

As a manipulation check of WMR implementation, we use a pseudo Difference-in-Differences (DiD) model specification to test whether WMR indeed works in the way that it was designed. In this DiD, the first difference is pre- and post-WMR, and the second difference is between highly-rated and lowly-rated sellers. The pseudo DiD model is specified in Equation (3) below.

$$\ln\left(1 + S\_REV_{i,t}\right) = \delta_{o,i} + \delta_1 WMR_t + \delta_2 High\_Rating_{i,t} + \delta_3 WMR_t \times High\_Rating_{i,t} + \gamma_3 T_t + \phi_3 Z_{i,t} + \mu_{3,i,t} \quad (3)$$

The dummy variable $High\_Rating_{i,t}$ is coded as 1 for sellers with high star rating (defined by 5 stars), and 0 for otherwise. The results suggest that the DiD coefficient is statistically significant and positive. This confirms that the WMR algorithm indeed gives more prominence to the highly-rated sellers, i.e., works as intended, in the sense that if a seller provides top-rating quality food, s/he gets a higher recommendation rank and therefore more clickthroughs and sales on the platform (similar to the way that Google's page rank algorithm gives more prominence to websites that match a set of keywords, and higher-ranked websites get more conversions). In addition, we also ran further sub-sample regressions to clarify this heterogeneity. As we suspected in the hypothesis section, WMR did not have a revenue-boosting effect even for those sellers with a high star rating, although the negative impact was much smaller than those with lower ratings, which implies our "*demand spillover issue*". In order to enhance comparability between highly-rated and lowly-rated sellers and uphold the basic assumptions of the DiD design, we conducted an additional pseudo PSM-DiD analysis, which yielded consistent results. We also examined the parallel trends between matched samples using a relative-time model (RTM) (Autor 2003). It is worth mentioning that seller ratings may also be influenced by algorithm changes, affecting our estimation. However, in our pseudo DiD design, we encode seller ratings as a binary choice variable, where only sellers with a rating of 5 are considered highly-rated. Due to the high threshold, this group of sellers is relatively stable, and therefore, the impact from WMR can be negligible.

Variable	(1) Pseudo DiD for Rating	(2) High Rating (5 stars)	(3) Low Rating (<5 stars)	(4) Pseudo DDD for Capacity	(5) Low Rating Low Capacity (Last 50%)	(6) High Rating High Capacity (Top 50%)	(7) High Rating High Capacity (Top 20%)	(8) High Rating High Capacity (Top 10%)	(9) High Rating High Capacity (Top 5%)
WMR × High_Rating	0.153*** (0.015)			0.105*** (0.016)					
WMR × High_Rating × Capacity				0.0002** (0.0001)					
WMR × Capacity				-0.0004*** (0.00004)					
High_Rating × Capacity				-0.0002 (0.001)					
WMR	-0.050*** (0.005)	-0.015* (0.007)	-0.026*** (0.004)	-0.027*** (0.009)	-0.024*** (0.004)	-0.015 (0.010)	-0.048 (0.027)	0.022 (0.035)	0.053* (0.021)
High_Rating	-0.281*** (0.068)			-0.224** (0.081)					
Capacity				0.0004** (0.0002)					
Constant	0.382 (0.598)	1.625** (0.557)	-0.273 (0.682)	0.314 (0.610)	0.161 (0.604)	1.846** (0.799)	0.985 (0.870)	3.982 (2.339)	1.678 (1.571)
Control variables	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time trends	YES	YES	YES	YES	YES	YES	YES	YES	YES
Seller effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Number of sellers	7,883	1,849	6,930	7,883	7,577	723	172	61	25
Number of observations	1,478,840	227,446	1,251,394	1,478,840	1,402,104	76,736	18,201	7,497	3,886
R^2	0.2026	0.2306	0.2079	0.2051	0.2021	0.2287	0.2754	0.3080	0.4163

Table 3. Heterogeneous Effect of WMR Algorithm Change

Note: Clustered robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Furthermore, we also examine the moderating role of the seller's capacity on the impact of WMR. As explained by the "*Constrained capacity presumption*", if the negative impact of WMR is due to a limitation in seller capacity, then sellers, especially those highly-rated sellers, with a higher dish capacity should receive less negative impact, as they are better equipped to cope with a sudden influx of demand. To capture this heterogeneity and mitigate the potential issue of incomparability between highly-rated and lowly-rated sellers in the pseudo DiD analysis, we introduced a pseudo Difference-in-Difference-in-Differences (DDD) specification. We introduce a moderating variable $Capacity_{i,t}$, which is a continuous variable on seller i's dishes capacity, measured by the sum of his/her stock quantities for each dish at time t. As we expected, the coefficient on the triple interaction term is still significantly positive. High capacity significantly enhances the protective effect of high ratings for sellers. In other words, for those highly-rated sellers, high capacity can further mitigate the negative impact of WMR and even demonstrate a positive effect on revenue. We also validate our findings through a series of sub-sample regressions. For lowly-rated sellers or highly-rated sellers with low capacity (below the median), WMR, as expected, exhibits a significant negative economic effect. However, for highly-rated sellers with high capacity, the negative impact of WMR is less pronounced and no longer significant. We further narrow down the range of high capacity and find that for the top 5% of highly-rated sellers with the highest capacity, WMR shows a significant positive effect on revenue promotion. Unfortunately, the number of sellers who achieve a revenue boost is minimal. The above analysis provides strong evidence for hypothesis H1 and our "*Constrained capacity presumption*".

Random Coefficient and Hierarchical Bayesian Estimations

Because latent individual-specific differences may not be observed in the data (such as sellers' personality values), some unobserved heterogeneity may bias the estimates. Thus, we further develop a random coefficient model (RCM) and a Hierarchical Bayesian (HB) model that both can allow for individual-specific heterogeneity of parameters to avoid incorrect inferences (Sun et al. 2017; Swamy 1970) and provide validation for our findings. Also, HB gauges the effects of WMR and BPR on seller revenue with a distribution of estimation, rather than a point of estimate. At the top level, we model the factors of each seller's revenue, after accounting for individual-specific parameters in Equations (4) and (5):

$$\ln(1 + S\_REV_{i,t}) = \alpha_{o,i,t} + \alpha_{1,i}WMR_t + \gamma_{4,i}T_t + \phi_{4,i}Z_{i,t} + \mu_{4,i,t} \tag{4}$$

$$\ln(1 + S\_REV_{j,t}) = \beta_{o,j,t} + \beta_{1,j}BPR_t + \gamma_{5,j}T_t + \phi_{5,j}Z_{j,t} + \mu_{5,j,t} \tag{5}$$

where the models include the base level parameters $(\alpha_{o,i,t}, \beta_{o,j,t})$, which capture all of the other individual seller-day specific factors. (The RCM and HB model specifications exclude the time-invariant covariates.) Because $\alpha_{o,i,t}$ and $\beta_{o,j,t}$ may exhaust the degree-of-freedom and cannot be identified, we decompose them into two components: $\alpha_{0,i,t} = \alpha_{0,i} + \alpha_{0,t}$ and $\beta_{0,i,t} = \beta_{0,i} + \beta_{0,t}$. The parameters $\alpha_{0,i}$ and $\beta_{0,i}$ measure the baseline of individual seller i and j, and $\alpha_{0,t}$ and $\beta_{0,t}$ capture the baseline day effect. At the lower level, we model the individual-specific effects of algorithm implementation with two parameters: the grand mean effect and error term as shown in Equation (6) below.

$$\begin{bmatrix}\alpha_{0,i}\\ \alpha_{1,i}\\ \gamma_{4,i}\\ \phi_{4,i}\end{bmatrix} = \begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\\ \omega_4\end{bmatrix} + \begin{bmatrix}\tau_{1,i}\\ \tau_{2,i}\\ \tau_{3,i}\\ \tau_{4,i}\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}\beta_{0,j}\\ \beta_{1,j}\\ \gamma_{5,j}\\ \phi_{5,j}\end{bmatrix} = \begin{bmatrix}\partial_1\\ \partial_2\\ \partial_3\\ \partial_4\end{bmatrix} + \begin{bmatrix}\varepsilon_{1,i}\\ \varepsilon_{2,i}\\ \varepsilon_{3,i}\\ \varepsilon_{4,i}\end{bmatrix} \tag{6}$$

Figure 4 shows the trace plots of 10000 iterations and histograms of the estimated Bayesian posterior means of the coefficients for WMR and BPR. The mass of the posterior mean distributions for WMR and BPR are indeed negative and positive in the 95% credible interval ($CI_{0.95}^{WMR} = [-.0210516, -.0210467]$, $CI_{0.95}^{BPR} = [0.0646085, 0.064611]$). The RCM and Bayesian posterior estimates of individual-specific parameters provide consistent support that the WMR targeting review ratings leads to a decline in seller revenue, while the BPR targeting own preference personalization drives seller revenue on the platform. Due to page limitations, additional results of RCM and HB estimations are not shown in the body of the paper.

(a) WMR

(b) BPR

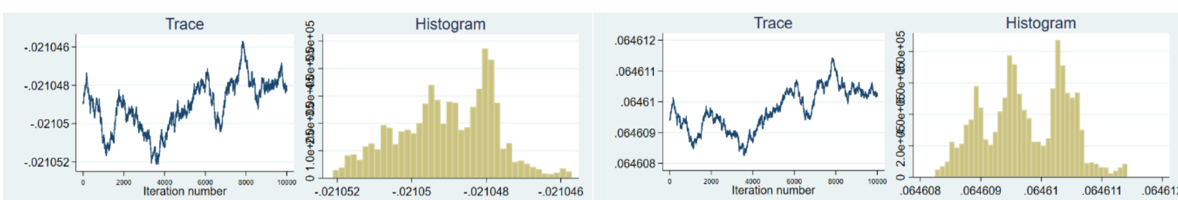


Figure 4. Graphical Summaries for Hierarchical Bayesian Posterior Distributions

Robustness Checks

In order to ensure the credibility of our conclusions, we conducted extensive robustness checks. For example, following the suggestion of Hausman and Rapson (2018), we conducted robustness tests on our RDiT design. We also enhanced comparability between sellers before and after algorithm changes through multiple PSM designs. Additionally, we removed the control variables that could be influenced by the algorithm changes and found the consistent significant revenue effects. This supports the partial mediation mechanisms we proposed. Due to space limitations, we summarized these tests in Table 4 below.

Facets	Motivation / Concerns	Additional Robustness Check	Results
RDiT	To confirm the validity of the assumption that there is no precise platform manipulation of the assignment variable (time)	Covariate continuity test (RDiT plots on control variables)	Continuous
	The model may be incorrectly specified under simple linear regressions	Regressions with polynomial global specifications (1st to 9th order)	Tenable (All)
	There may be a potential serial correlation in the dependent variable (error term)	Regressions with lagged dependent variables (1st to 15th orders)	Tenable (All)
	It may be closer to local randomization at a narrower bandwidth setting	Samples from a shorter time window (±30 days, ±60 days)	Tenable (All)
	We may observe non-zero jumps at other cutoff points if RDiT does not work	Placebo test: a counterfeit policy (-10 days, -30 days)	Neither sig.
	To mitigate concerns about short-run selection/anticipation/avoidance effects	'Donut' regression dropping central samples (±10 days, ±30 days)	Tenable (All)
Heterogeneity	Pseudo DiD specification: Highly-rated sellers and lowly-rated sellers may not be comparable (e.g., with different pre-policy trends)	Pseudo PSM-DiD estimates	Tenable
		Parallel trend testing: relative time modelling (RTM)	Pass
		Pseudo DDD specification for capacity	Tenable
	Some unobservable heterogeneity may bias the estimates	Random coefficient estimates	Tenable
		Hierarchical Bayesian estimates	Tenable
Matching	Sellers before and after WMR and BPR may be incomparable	Full-panel PSM	Tenable
	Basic PSM cannot mitigate concerns about unobservable seller characteristics	PSM with the same seller	Tenable
	Basic PSM cannot mitigate concerns about time trends	Relative-time PSM (Wang and Goldfarb 2017)	Tenable
Model Specification	Some control variables may be significantly influenced by algorithm changes, affecting the credibility of the estimates	We estimated models with and without rating (mean & s.d.) and dishes amount (new & total) as control variables	Partial Mediation
	Seller revenue is truncated by 0, which can lead to inconsistent estimates	Tobit estimates (both cross-sectional and panel Tobit specifications)	Tenable (All)
	Given that there are some 0 values for revenue, we want to check whether the results are robust to the setting of the log transformation function for revenue	Different specifications for seller revenue, e.g., arcsinh(y)	Tenable
Sample Selection	To keep the statistical power high while reducing bias from possible extreme data	75%, 50% and 25% random smaller sample subset	Tenable (All)
	The ceiling effect of 5-star sellers may affect the conclusion since they could not further improve quality	Sub-sample excluding 5-star sellers	Tenable
	Some 'professional' sellers may affect the estimation	Sub-sample of only 'unprofessional' sellers with less than 50 dishes	Tenable
Panel setting	Showing coarse-grained effects	Week-level panels	Tenable
	Showing effects on sellers who have been active on the platform throughout each completed observation period	Balanced panels	Tenable
	Showing effects in the entire observation period	Integrated panels (WMR + BPR)	Tenable
Standard Error	Possible correlation within the same seller's strategy and performance	Robust standard error clustered at seller level	Tenable
	Possible cross-sectional correlation	Driscoll Kraay standard error	Tenable

Table 4. Overview of Potential Concerns and Corresponding Robustness Checks

Note: "Tenable" means that the conclusion is consistent with our baseline results, where WMR shows a significant negative impact and BPR shows a significant positive impact.

Mechanisms and Additional Analyses

Price, Quantity or Average Order Value as the Driver of Revenue Changes

We now test whether the revenue effects of WMR and BPR are driven by changes in prices, number of orders, or average order values. Our analysis is carried out in two stages. First, we decompose the seller's daily revenue as the product of daily order amount and daily average order value (AOV), which provides an estimate of the potential value by the buyers (Behera et al. 2020). In this stage, we examine the impacts of the algorithm changes on their transaction from the *buyer's* perspective. Prior research notes that recommendation algorithms may change buyers' value (Adomavicius et al. 2018) or influence the buyer orders (Häubl and Trifts 2000). The results reported in Columns (1) and (3) of Table 5 show that the effect of WMR on daily orders is not significant, while the effect of WMR on AOV is significantly negative, suggesting that the revenue decrease by WMR is mainly driven by fewer AOV rather than changes in orders. We did not find evidence of WMR having a promotional effect on order amount, which can be attributed to our "*Constrained capacity presumption*". In addition, Columns (2) and (4) of Table 5 show that the effects of BPR on daily orders and AOV are both significantly positive, suggesting that the revenue increase by BPR is driven by both more orders and higher AOV.

Second, we turn our perspective to the *sellers*, to examine whether the changes in AOV can be attributed to a strategic price change by the seller (Zhou and Zou 2022). The results in Columns (5) and (6) of Table 5 show no evidence that the algorithm changes significantly altered sellers' prices, suggesting that the decline in AOV stems from a change in buyers' willingness rather than sellers' strategic pricing. These findings further support our "*Matching friction presumption*" for the WMR effect: buyers are not reducing their purchases, but they seem to be willing to pay much lower for the kitchens they choose on the platform.

Variable	(1) WMR: Daily orders	(2) BPR: Daily orders	(3) WMR: Daily AOV	(4) BPR: Daily AOV	(5) WMR: Prices	(6) BPR: Prices
WMR (Word-of-mouth recommendation)	-0.001 (0.002)		-0.028*** (0.003)		-0.014 (0.013)	
BPR (Botler personalization recommendation)		0.025*** (0.002)		0.057*** (0.004)		0.002 (0.012)
Constant	0.012 (0.198)	0.090 (0.277)	0.504 (0.451)	1.435 (1.061)	22.714*** (1.516)	21.915*** (1.050)
Control variables	YES	YES	YES	YES	YES	YES
Time trends	YES	YES	YES	YES	YES	YES
Seller effects	YES	YES	YES	YES	YES	YES
Number of sellers	7,883	7,883	7,883	7,883	7,883	7,883
Number of observations	1,478,840	1,450,413	1,478,840	1,450,413	1,478,840	1,450,413
R^2	0.3261	0.2881	0.1182	0.0801	0.0014	0.0002

Table 5. The Impacts of WMR and BPR on Seller Orders, AOV, and Average Prices

Note: Clustered robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

The Pathways for the WMR and BPR Effects

Next, we dive deeper into the mechanisms by explaining the changed AOV and thus sales revenues with two different types of seller's behavioral changes: sellers may respond to the platform algorithm changes by specializing in a higher quality of current products or by innovating with more new products.

We estimate Model (1) by replacing revenue with review ratings and rating standard deviations as the dependent variable, which can be used as signals of seller quality. The model estimation result is shown in Table 6. The positive coefficient suggests that after controlling for other potential trends, the WMR indeed incentivizes sellers to improve the quality of current dishes. On average, sellers' ratings increase by 0.025 star per day (about 0.75 star improvement in one month) after WMR. Moreover, the coefficient of WMR is significantly negative for rating standard deviation. The effect size is nontrivial: sellers' rating dispersion decreases by 0.011 star per day (about 0.3 star less dispersion in one month) after WMR. Together, these findings suggest that the WMR platform algorithm change indeed incentivizes sellers to be more specialized, providing strong support for hypothesis H3a. However, when we turn our attention to BPR algorithm change, things seem to be out of control. Columns (2) and (4) show that the coefficient of BPR is significantly negative for review rating and significantly positive for rating standard deviation, suggesting that after controlling for other potential trends, sellers' review rating and the quality consistency level both decrease after BPR. These results support that although BPR induces the benefits of more new product introductions, it has an unintended negative effect and impairs quality specialization for sellers on the platform, providing strong evidence for our hypothesis H4b. As a robustness check, we further use other methods (delivery ratings) to measure the service quality of sellers and still obtained consistent conclusions.

Similarly, to examine the impacts on seller innovation, we estimate Model (2) by replacing the revenue with sellers' new product introductions as the dependent variable. The estimation results in Table 6 suggest that the coefficient of BPR is statistically significant and positive across the three measures: new dishes, new types of meat (e.g., beef, chicken, pork, or fish), and new cooking methods (e.g., fry, stew, or roast) of the dish offerings. These findings suggest that the BPR platform algorithm change indeed incentivizes sellers to be more innovative, supporting our proposed hypothesis H4a. Again, as we hypothesized for WMR, Columns (5), (7), and (9) of Table 6 show that the coefficient of WMR is significantly negative for each measurement. These results support that although WMR induces benefits with higher quality of existing offerings, it also has an unintended negative effect and impedes new product introductions by sellers on the platform, which provides support for hypothesis H3b. This inhibition of innovation may have also led, in part, to lower seller revenues. Given that the dependent variable is non-negative integers, we also re-estimated our model using Poisson regression, zero-inflated Poisson regression, negative binomial regression, and zero-inflated negative binomial models to ensure the robustness of our findings, and we obtained consistent conclusions.

Variable	**Mechanism: Product Quality**				**Mechanism: Product Innovation**					
	(1) WMR Stars	(2) BPR Stars	(3) WMR; S. D. of Stars	(4) BPR S. D. of Stars	(5) WMR New Dishes	(6) BPR New Dishes	(7) WMR New Types of Meat	(8) BPR New Types of Meat	(9) WMR New Cooking Methods	(10) BPR New Cooking Methods
WMR	0.025*** (0.003)		-0.011*** (0.001)		-0.009*** (0.002)		-0.010*** (0.003)		-0.009** (0.003)	
BPR		-0.009*** (0.001)		0.003*** (0.001)		0.003*** (0.001)		0.006*** (0.002)		0.007*** (0.001)
Constant	0.683*** (0.154)	2.712*** (0.178)	0.288*** (0.089)	-1.109*** (0.062)	0.523*** (0.093)	0.369* (0.206)	0.939*** (0.130)	0.941*** (0.282)	0.710*** (0.075)	0.825*** (0.099)
Control var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time trends	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Seller effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
# of sellers	7,883	7,883	7,883	7,883	7,883	7,883	7,883	7,883	7,883	7,883
# of obs.	1,478,840	1,450,413	1,478,840	1,450,413	1,478,840	1,450,413	1,478,840	1,450,413	1,478,840	1,450,413
R^2	0.6008	0.4570	0.5669	0.4509	0.0315	0.0026	0.0324	0.0049	0.0341	0.0059

Table 6. Mechanisms: Product Quality and Product Innovation

Note: Clustered robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Thus, reasonably but in an interesting manner, WMR or BPR each has some unintended negative effects, despite the benefits. WMR boosts quality specialization at the expense of innovation, while BPR promotes innovation at the expense of quality among the sellers. Clearly, a platform manager would like these algorithm changes to motivate most sellers to improve both specialization and innovation, rather than one at the expense of the other. These double-edged effects inspired us to leverage a machine learning technique for optimal targeting.

Heterogeneous Effects of WMR and BPR on Seller Revenues

As the effects of WMR and BPR are likely to be heterogeneous across sellers, here we provide more nuanced evidence for the pathways by revealing their different effect magnitudes. Extant studies show that firms with a more concentrated product assortment will achieve superior performance in a competitive marketplace. If WMR affects seller revenue via the specialization pathway of improving the current quality ratings, then the revenue reduction of WMR should be less salient for sellers who have a more concentrated product assortment prior to the algorithm change, because they are relatively easier to improve current product quality. We use Herfindahl-Hirschman index (HHI) to measure the concentration of different dish types for each seller. The index is calculated at a seller-dish level on the share of each dish and sums the square of share within each seller. The coefficients of interaction between WMR and concentration are significantly positive, suggesting that sellers with a more concentrated dish assortment indeed receive less negative impact from WMR, in line with the specialization pathway.

Further, the motivation and ability to innovate also vary with each seller, which implies that if BPR drives seller revenue via the innovation pathway of introducing more new products, then the revenue effect of BPR should be amplified for entrepreneurs who are more agile and nimble in their product innovation. We have considered several methods of measuring innovative agility for sellers. First, it is clear that sellers with higher available capacity tend to be more agile as they are able to innovate flexibly while maintaining the supply of existing products. The results in Column (4) of Table 7 suggest that the interaction term between BPR and sellers' capacity is significantly positive, affirming that it is relatively those large-capacity sellers on the platform that indeed benefit more from BPR, consistent with the innovation pathway. Second, extant studies find that entrepreneur age plays an essential role in innovation, and younger firms are more likely to innovate (Ransbotham and Mitra 2010). The results in Column (5) of Table 7 suggest that the interaction term between BPR and sellers' age is significantly negative, affirming that it is relatively younger sellers on the platform that indeed benefit more from BPR, also consistent with the innovation pathway.

Third, entrepreneur experience is also an important factor in the strategic agility of entrepreneurs (Westphal and Zajac 2001). Experienced sellers tend to have better insight into market dynamics and make more agile and flexible innovations in response to market demand (Nerkar and Roberts 2004). Indeed, we observe from Column (6) of Table 7 that more experienced sellers benefit more from BPR. Further, a few studies also point out that the impact of experience may be non-linear: sellers who entered the market very early may already have an advantage in a certain niche and the wealth of experience may reinforce their perception of risk and curb potential incentives to innovate (Cliff et al. 2006). We introduce a quadratic term for experience and its interaction term with BPR to explain this possible effect, and the results in Column (7) show that experience indeed has a significant inverted U-shaped moderating effect on the revenue-boosting effect of BPR, with moderately experienced sellers tending to reap the most from BPR.

Variable	(1) Concentration of meat types	(2) Concentration of vegetable types	(3) Concentration of cooking methods	(4) Measurement 1 for seller innovation agility	(5) Measurement 2 for seller innovation agility	(6) Measurement 3 for seller innovation agility	(7) Measurement 3 for seller innovation agility
WMR × Seller assortment concentration	0.000*** (0.000)	0.000*** (0.000)	0.000** (0.000)				
WMR	-0.055*** (0.007)	-0.038*** (0.005)	-0.032*** (0.005)				
Seller assortment concentration	-0.000 (0.000)	-0.000 (0.000)	-0.000*** (0.000)				
BPR × Seller innovation agility				0.000** (0.000)	-0.001*** (0.000)	0.000*** (0.000)	0.001** (0.000)
BPR × (Seller innovation agility)^2							-0.000*** (0.000)
BPR				0.072*** (0.007)	0.116*** (0.007)	0.068*** (0.007)	0.080*** (0.010)
Seller innovation agility				0.013 (0.016)		-0.022*** (0.003)	-0.025*** (0.003)
(Seller innovation agility)^2							0.000*** (0.000)
Constant	0.381 (0.609)	0.366 (0.607)	0.506 (0.597)	1.123 (1.052)	1.316 (1.166)	8.916*** (0.831)	8.812*** (0.869)
Control variables	YES	YES	YES	YES	YES	YES	YES
Time trends	YES	YES	YES	YES	YES	YES	YES
Seller effects	YES	YES	YES	YES	YES	YES	YES
Number of sellers	7,883	7,883	7,883	7,883	7,883	7,883	7,883
Number of observations	1,478,840	1,478,840	1,478,840	1,450,413	1,450,413	1,450,413	1,450,413
R^2	0.2021	0.2020	0.2020	0.1594	0.1594	0.1598	0.1603

Table 7. Heterogeneous Revenue Effects of WMR and BPR Algorithm Changes

Note: Clustered robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. Measurement 1 for seller innovation agility = seller's available stock level; measurement 2 for seller innovation agility = seller age, which is time-invariant; measurement 3 for seller innovation agility = seller experience on the food-sharing platform.

Optimal Targeting for WMR and BPR

As we mentioned earlier, the dual nature of both WMR and BPR motivates us to capture the heterogeneous impact of the two algorithm changes on seller specialization and innovation through machine learning techniques. Specifically, we use *causal random forest* with honest tree (Wager and Athey 2018) to learn heterogeneous responses to WMR and BPR among different sellers and allow the platform manager to craft an optimal targeting rule. The optimal targeting rule here is to learn and leverage the best combination of sellers' individual heterogeneity based on their feature variables (such as prior review ratings and the number of dish offerings). The causal forest approach to the optimal targeting rules is appealing here because it does not assume any linear or nonlinear interactive relationships for the combination of heterogeneity variables that regulate the effects of WMR and BPR. Rather, it is a nonparametric tool and can decompose the average effects of WMR and BPR into infinite combinations of heterogeneous effects by relentlessly (brutal force) learning and splitting the data on the individual heterogeneity variables.

We denote (X_i, Z_i) as independent samples that include individual sellers' feature variables of X_i and the sales revenue variable of Z_i and W_i as the dichotomous variable of WMR or BPR, when building a causal regression tree. The random forest algorithm can recursively split the feature space of samples until we have a set of leaves $L(x) \coloneqq L$, each of which only contains a few training samples. Then, given a test point x, we can evaluate the prediction $\hat{\delta}(x)$ by identifying the leaf $L(x)$ and setting $\hat{\delta}(x) = \frac{1}{|\{i:X_i \in L\}|} \sum_{\{i:X_i \in L\}} Z_i$. The causal trees assure that the leaves are small enough so that the (Z_i, W_i) pairs correspond to the indices i for $i \in L(x)$ as if in a randomized experiment with a balanced sampling and splitting. Then, it estimates the causal effect for $\forall\, X_i \in L(x)$ as: $\hat{z}(x) = \frac{1}{|\{i:W_i=1, X_i \in L\}|} \sum_{\{i:W_i=1, X_i \in L\}} Z_i - \frac{1}{|\{i:W_i=0, X_i \in L\}|} \sum_{\{i:W_i=0, X_i \in L\}} Z_i$. After causal random forest generates an ensemble of B trees, each of which casts a vote with an estimate $\hat{z}_b(x)$, the forest then aggregates their predictions by averaging these votes: $\hat{z}(x) = B^{-1} \sum_{b=1}^{B} \hat{z}_b(x)$.

The results indicated that, for WMR to attain both highest specialization and highest innovation, the platform should recommend the optimal targets—the sellers whose prior review rating stars are from 4.51 to 5 and whose number of dishes is from 26 to 83 simultaneously. In addition, for BPR, the platform should recommend the optimal targets—the sellers whose prior review rating stars are from 3.42 to 4.31 and whose number of dishes is from 16 to 55, in order to attain both highest specialization with rating reputation and highest innovation with new products after the algorithm change. Thus, these findings show that while both WMR and BPR have advantages and limitations, the inherent trade-off can be mitigated if using the causal forest technique to craft optimal targeting rules for identifying seller segments on the platform.

Discussion

This paper examines the economic effects of two recommendation algorithms changes, i.e., WMR and BPR, on a large-scale food-sharing platform. We reveal several insightful findings: compared to naïve geo-fencing recommendation, WMR has a negative economic effect on sellers' revenues; while compared to WMR, BPR has a positive economic effect. WMR incentivizes sellers to become specialists with high quality of current products, however, this comes at the expense of reduced new product offerings. In contrast, BPR motivates sellers to innovate and introduce more new products to suit diverse customer tastes, but at the same time makes them do less well in product and service quality. Consistent with the specialization pathway, WMR is more beneficial for sellers who have a concentrated product assortment. In contrast and in line with the innovation pathway, it is relatively more agile and flexible entrepreneurs that reap more benefits from BPR.

Our study contributes to prior literature. First, extant studies focus on the benefits of a recommendation system for consumer welfare and sellers' profit. Our research shows for the first time that, under certain market conditions, platform recommendation algorithms can also have a negative economic effect. Second, current research on sharing economy identified the key features of shared goods and the substitution or complementarity with the incumbents at aggregated levels. Using a unique and rich dataset, we focus on the design of recommendation algorithms embedded in sharing platforms at a more fine-grained level. Third, we also identify several underlying mechanisms and heterogeneity for the impact of recommendation algorithms on sales effectiveness, which has also received little attention in prior studies.

Our study also has several practical implications. For complementors, our research highlights the potential systemic impacts of platform algorithms on their revenues and operational strategies, which are necessary for their decision-making. Also, when they have options for multi-homing, alignment of their own goals with those of the platform should be an important consideration. For sharing platforms, managers need to be aware that the economic effects of algorithms may be quite different from those of traditional online platforms. Although managers may adopt different algorithms for different objectives (Resnick and Varian 1997; Rietveld and Eggers 2018), they may overlook their potential negative effects, ultimately leading to undesirable outcomes. Our work suggests that WMR may be an effective incentive when platforms urgently need to improve the product and service quality, but this may come at the cost of lower diversity and lower revenue, while BPR, although seemingly more beneficial, may bring concerns about product quality.

Our work has some limitations that could motivate further research. First, our study focuses on the economic effects of two specific algorithms—i.e., WMR and BPR. Additional studies examining other algorithms (e.g., expert recommendations) could be a useful contribution. A second limitation is related to data coverage. Our data could not cover long-run effects on sellers' behaviors. Examining the performance of sellers over a longer time period and considering the gradual evolution of the algorithms is an important research direction. Finally, limited by a single scenario, further studies on the impacts of recommendation algorithms across sharing platforms (Perren and Kozinets 2018) are needed to extend our findings.

References

Abrate, G., and Viglia, G. 2019. "Personal or Product Reputation? Optimizing Revenues in the Sharing Economy," *Journal of Travel Research* (58:1), pp. 136-148.

Adomavicius, G., Bockstedt, J. C., Curley, S. P., and Zhang, J. 2018. "Effects of Online Recommendations on Consumers' Willingness to Pay," *Information Systems Research* (29:1), pp. 84-102.

Andrews, M., Luo, X., Fang, Z., and Ghose, A. 2016. "Mobile Ad Effectiveness: Hyper-Contextual Targeting with Crowdedness," *Marketing Science* (35:2), pp. 218-233.

Autor, D. H. 2003. "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing," *Journal of Labor Economics* (21:1), pp. 1-42.

Babar, Y., and Burtch, G. 2020. "Examining the Heterogeneous Impact of Ride-Hailing Services on Public Transit Use," *Information Systems Research* (31:3), pp. 820-834.

Behera, R. K., Gunasekaran, A., Gupta, S., Kamboj, S., and Bala, P. K. 2020. "Personalized Digital Marketing Recommender Engine," *Journal of Retailing and Consumer Services* (53), p. 101799.

Berger, J. 2014. "Word of Mouth and Interpersonal Communication: A Review and Directions for Future Research," *Journal of Consumer Psychology* (24:4), pp. 586-607.

Bernal, J. L., Cummins, S., and Gasparrini, A. 2017. "Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial," *International Journal of Epidemiology* (46:1), pp. 348-355.

Bodapati, A. V. 2008. "Recommendation Systems with Purchase Data," *Journal of Marketing Research* (45:1), pp. 77-93.

Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43:3), pp. 345-354.

Cliff, J. E., Jennings, P. D., and Greenwood, R. 2006. "New to the Game and Questioning the Rules: The Experiences and Beliefs of Founders Who Start Imitative Versus Innovative Firms," *Journal of Business Venturing* (21:5), pp. 633-663.

Davis, L. W. 2008. "The Effect of Driving Restrictions on Air Quality in Mexico City," *Journal of Political Economy* (116:1), pp. 38-81.

Ernst, H., Hoyer, W. D., and Rübsaamen, C. 2010. "Sales, Marketing, and Research-and-Development Cooperation across New Product Development Stages: Implications for Success," *Journal of Marketing* (74:5), pp. 80-92.

Filippas, A., Horton, J. J., and Zeckhauser, R. J. 2020. "Owning, Using, and Renting: Some Simple Economics of the "Sharing Economy"," *Management Science* (66:9), pp. 4152-4172.

Fleder, D., and Hosanagar, K. 2009. "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity," *Management Science* (55:5), pp. 697-712.

Häubl, G., and Trifts, V. 2000. "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids," *Marketing Science* (19:1), pp. 4-21.

Hausman, C., and Rapson, D. S. 2018. "Regression Discontinuity in Time: Considerations for Empirical Applications," *Annual Review of Resource Economics* (10), pp. 533-552.

Henderson, R. M., and Clark, K. B. 1990. "Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms," *Administrative Science Quarterly* (35), pp. 9-30.

Insider. 2016. "Netflix Lifted the Lid on How the Algorithm That Recommends You Titles to Watch Actually Works." from http://www.businessinsider.com/how-the-netflix-recommendation-algorithm-works-2016-2

Kerin, R. A., Varadarajan, P. R., and Peterson, R. A. 1992. "First-Mover Advantage: A Synthesis, Conceptual Framework, and Research Propositions," *Journal of Marketing* (56:4), pp. 33-52.
Kumar, A., and Hosanagar, K. 2019. "Measuring the Value of Recommendation Links on Product Demand," *Information Systems Research* (30:3), pp. 819-838.
Lee, D., and Hosanagar, K. 2019. "How Do Recommender Systems Affect Sales Diversity? A Cross-Category Investigation Via Randomized Field Experiment," *Information Systems Research* (30:1), pp. 239-259.
Levinthal, D. A., and Wu, B. 2010. "Opportunity Costs and Non - Scale Free Capabilities: Profit Maximization, Corporate Scope, and Profit Margins," *Strategic Management Journal* (31:7), pp. 780-801.
Li, X., Grahl, J., and Hinz, O. 2022. "How Do Recommender Systems Lead to Consumer Purchases? A Causal Mediation Analysis of a Field Experiment," *Information Systems Research* (33:2), pp. 620-637.
Luo, X., Tong, S., Lin, Z., and Zhang, C. 2021. "The Impact of Platform Protection Insurance on Buyers and Sellers in the Sharing Economy: A Natural Experiment," *Journal of Marketing* (85:2), pp. 50-69.
Nerkar, A., and Roberts, P. W. 2004. "Technological and Product - Market Experience and the Success of New Product Introductions in the Pharmaceutical Industry," *Strategic Management Journal* (25:8 - 9), pp. 779-799.
Pathak, B., Garfinkel, R., Gopal, R. D., Venkatesan, R., and Yin, F. 2010. "Empirical Analysis of the Impact of Recommender Systems on Sales," *Journal of Management Information Systems* (27:2), pp. 159-188.
Perren, R., and Kozinets, R. V. 2018. "Lateral Exchange Markets: How Social Platforms Operate in a Networked Economy," *Journal of Marketing* (82:1), pp. 20-36.
Ransbotham, S., and Mitra, S. 2010. "Target Age and the Acquisition of Innovation in High-Technology Industries," *Management Science* (56:11), pp. 2076-2093.
Resnick, P., and Varian, H. R. 1997. "Recommender Systems," *Communications of the ACM* (40:3), pp. 56-58.
Rietveld, J., and Eggers, J. 2018. "Demand Heterogeneity in Platform Markets: Implications for Complementors," *Organization Science* (29:2), pp. 304-322.
Rossi, M. 2023. "Competition and Reputation in an Online Marketplace: Evidence from Airbnb," *Management Science* (Forthcoming).
Shi, Z., and Raghu, T. 2020. "An Economic Analysis of Product Recommendation in the Presence of Quality and Taste-Match Heterogeneity," *Information Systems Research* (31:2), pp. 399-411.
Skiti, T., Luo, X., and Lin, Z. 2022. "When More Is Less: Quality and Variety Trade - Off in Sharing Economy Platforms," *Journal of Management Studies* (59:7), pp. 1817-1838.
So, K. K. F., Oh, H., and Min, S. 2018. "Motivations and Constraints of Airbnb Consumers: Findings from a Mixed-Methods Approach," *Tourism Management* (67), pp. 224-236.
Sun, Y., Dong, X., and McIntyre, S. 2017. "Motivation of User-Generated Content: Social Connectedness Moderates the Effects of Monetary Rewards," *Marketing Science* (36:3), pp. 329-337.
Swamy, P. A. 1970. "Efficient Inference in a Random Coefficient Regression Model," *Econometrica* (38:2), pp. 311-323.
Tae, C. J., Luo, X., and Lin, Z. 2020. "Capacity - Constrained Entrepreneurs and Their Product Portfolio Size: The Response to a Platform Design Change on a Chinese Sharing Economy Platform," *Strategic Entrepreneurship Journal* (14:3), pp. 302-328.
Verma, S., and Yadav, N. 2021. "Past, Present, and Future of Electronic Word of Mouth (eWOM)," *Journal of Interactive Marketing* (53), pp. 111-128.
Wager, S., and Athey, S. 2018. "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests," *Journal of the American Statistical Association* (113:523), pp. 1228-1242.
Wang, K., and Goldfarb, A. 2017. "Can Offline Stores Drive Online Sales?," *Journal of Marketing Research* (54:5), pp. 706-719.
Westphal, J. D., and Zajac, E. J. 2001. "Decoupling Policy from Practice: The Case of Stock Repurchase Programs," *Administrative Science Quarterly* (46:2), pp. 202-228.
Wired. 2017. "This Is How Netflix's Top-Secret Recommendation System Works." from http://www.wired.co.uk/article/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like
Xiao, B., and Benbasat, I. 2007. "E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact," *MIS Quarterly* (31:1), pp. 137-209.
Xu, X. 2020. "How Do Consumers in the Sharing Economy Value Sharing? Evidence from Online Reviews," *Decision Support Systems* (128), p. 113162.
Zervas, G., Proserpio, D., and Byers, J. W. 2017. "The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry," *Journal of Marketing Research* (54:5), pp. 687-705.
Zhou, B., and Zou, T. 2022. "Competing for Recommendations: The Strategic Impact of Personalized Product Recommendations in Online Marketplaces," *Marketing Science* (42:2), pp. 360-376.