Supplier Responses to Platform Recommendations in Business-to-Business Electronic Commerce

Yi Chen[*] Ziwei Cong[†] Yixing Chen[‡] Liwen Hou[§]

April 18, 2026

Abstract

Business-to-business electronic commerce (B2B e-commerce) platforms increasingly use recommendations to connect suppliers with potential business opportunities in the form of buyers' requests for quotation (RFQs). Yet little is known about whether and how these platform recommendations affect supplier bidding participation and downstream transaction outcomes. Using data from a construction-sector B2B e-commerce platform, we find that receiving a recommendation increases not only the probability of bidding on the recommended RFQs by 1.5 percentage points (direct effect) but also the probability of bidding on non-recommended RFQs by 0.8 percentage points (spillover effect). The positive spillover effect emerges because platform recommendations prompt suppliers to broaden their search and pursue additional opportunities beyond focal RFQs. However, the positive spillover effect does not translate into a higher probability of winning because spillover bids tend to target popular yet less relevant RFQs. Further, losing bids is positively associated with the likelihood of supplier churn over the subsequent six months. Collectively, we show that the engagement-oriented recommender systems in B2B may broaden supplier participation without improving downstream outcomes, highlighting the importance of evaluating multiple business objectives and accounting for trade-offs among them in platform design.

Keywords:
business-to-business e-commerce, platforms, digital nudging, quasi-experimental design

[*]Ross School of Business, University of Michigan, yichenyc@umich.edu
[†]McDonough School of Business, Georgetown University, zc260@georgetown.edu
[‡]Mendoza College of Business, University of Notre Dame, ychen43@nd.edu
[§]Antai College of Economics and Management, Shanghai Jiao Tong University, lwhou@sjtu.edu.cn

INTRODUCTION

The business-to-business electronic commerce (B2B e-commerce) sector has transformed how firms procure products and services (Chakravarty et al. 2014; Grewal et al. 2010). In 2025, global B2B e-commerce was valued at $32 trillion, more than six times the size of the business-to-consumer (B2C) sector (Bledsoe 2023). Central to this sector are B2B e-commerce platforms such as Alibaba, IndiaMART, and SAP Business Network, which facilitate the procurement process by enabling buyers to post requests for quotation (RFQs) that suppliers can evaluate and bid on. As transaction volume grows, these platforms increasingly use supplier-oriented recommendations (hereinafter, "platform recommendations") to reduce search costs, improve opportunity discovery, and encourage bidding participation in B2B procurement (Allu and Gaur 2024; Gaur and Liu 2020).

Yet the impact of platform recommendations on supplier bidding participation and downstream transaction outcomes is ambiguous for several reasons. First, their direct effect on bidding participation is not guaranteed. Although platform recommendations may improve opportunity discovery, bidding participation in B2B procurement is a high-stakes decision (Grewal et al. 2022) affected by factors such as pricing terms (Karlinsky-Shichor and Netzer 2024; Zhang et al. 2014), buyer-supplier relationships (Ganesan 1994; Tuli et al. 2007), and referrals (Hada et al. 2014). Thus, it is unclear if platform recommendations are sufficient to move suppliers from discovery to bidding on recommended RFQs. Second, suppliers often operate under limited time and capacity constraints while considering multiple RFQs simultaneously (Wu et al. 2015). As a result, recommendations may alter how suppliers allocate scarce bidding effort across opportunities: they may expand the set of RFQs a supplier bids on (positive spillovers) or redirect bidding effort to recommended RFQs at the expense of other opportunities (negative spillovers). Third, because B2B transactions are often auction-based, greater bidding participation may not translate into successes. Even if recommendations can encourage suppliers to bid more broadly, those marginal bids generated may be placed on less suitable RFQs that are less likely to convert into successes. Collectively, these features suggest that evaluating platform recommendations in B2B commerce requires going beyond their direct effect on supplier participation in bidding on recommended RFQs to examine spillover effects on non-recommended RFQs and downstream transaction outcomes. Accordingly, we ask three related questions: (1) Do platform recommendations increase suppliers' likelihood of bidding on recommended RFQs? (2) Do these recommendations crowd out or stimulate suppliers' likelihood of bidding on non-recommended RFQs? (3) Do these recommendations affect suppliers' likelihood of winning RFQs?

To answer these questions, we use proprietary data from a construction-sector B2B e-commerce platform. Launched in 2015, the platform has facilitated more than 130,000 RFQs. In April 2020, the platform started sending RFQ recommendations to suppliers via short message service (SMS). Between April 2020 and August 2023, the platform sent 621,827 recommendations to 32,234 suppliers, covering 62.2% of all RFQs posted and reaching nearly 80% of existing suppliers during this period.

Estimating the causal effect of platform recommendations is challenging because recommendation delivery is not random. Instead, recommendations are generated by a curation procedure that conditions on supplier characteristics, past activities on the platform, and RFQ–supplier relevance. A naive comparison of outcomes between recipients and non-recipients may confound the recommendation effects with selection induced by this procedure. The ideal experiment would identify a set of similar suppliers, randomly recommend an RFQ to a subset of them, and compare subsequent bidding participation between treated and control suppliers.[1] Our empirical strategy approximates this ideal experiment by leveraging our institutional knowledge of the curation procedure. For each recommendation in which a RFQ is recommended to a supplier on a given day, we construct a set of control observations that are matched along key inputs used by the curation procedure, as well as recent recom-

[1] According to the platform, they were not able to implement such an experiment because it would be costly and require infrastructure capabilities beyond the platform's current resources.

mendation histories.[2] Our identifying assumption is that, conditional on curation-relevant input and recent treatment history, treated and control observations are sufficiently comparable (Imai et al. 2023). We justify the credibility of this assumption through a series of analyses, including the bias-adjusted coefficient approach developed by Oster (2019), which evaluates the influence of unobserved confounders.

We find that receiving a recommendation increases the probability of bidding on the recommended RFQ by 1.5 percentage points, a 121% lift over the baseline. Moreover, recommendations generate a positive spillover effect: after receiving a recommendation, recipients are 0.8 percentage points more likely to bid on non-recommended RFQs, a 65% lift. These results suggest that platform recommendations lift supplier bidding participation in not only recommended RFQs but also other opportunities on the platform.

We then explore the underlying factors driving the spillover effect. A natural starting point is the matching quality of recommendations, i.e., how well recommended RFQs fit suppliers' profiles. We find that spillover is stronger when matching quality is lower, indicating that suppliers respond to poor matches by searching for alternative RFQs on their own. Nevertheless, matching quality alone does not fully explain the effect because the spillover persists even after accounting for it. This motivates us to explore a complementary mechanism: recommendations draw suppliers' attention to the marketplace and prompt broader scanning of opportunities. In support of this mechanism, the spillover effect is stronger among suppliers with greater incentives and bandwidth to pursue additional opportunities.

However, this positive spillover effect does not translate into greater supplier success, i.e., recommendations do not significantly increase the probability of winning bids on non-recommended RFQs. Specifically, spillover bids are disproportionately directed to more popular yet less well-matched opportunities, suggesting that recommendations induce broader rather than more effective exploration. More concerningly, losing bids substantially increases the likelihood of supplier churn over the subsequent six months. These findings suggest that

[2]The latter is to mitigate carryover effects from past treatments. See more details in Section section 3.

recommendations may carry an unintended consequence: they may weaken supplier retention in the longer term by expanding participation without improving success.

Our research offers practical implications for B2B platforms. First, we document a positive spillover effect of platform recommendations on supplier participation in bidding on non-recommended RFQs. Thus, platform managers should monitor and account for this spillover effect when evaluating the overall impact of recommender systems. Second, in auction-based settings, bidding participation alone provides an incomplete basis for evaluating recommendation effects because greater bidding activity does not guarantee greater supplier success. Platform managers need to augment bidding participation with downstream outcomes—conversion and retention—in their evaluation framework and incorporate them into the objective function of recommender systems. Third, because spillover effects are concentrated among suppliers with broad product portfolios, platforms should move beyond a "one-size-fits-all" recommendation strategy. For this segment, managers may wish to monitor whether platform recommendations induce broad yet low conversion bidding and calibrate recommendation intensity or targeting criteria accordingly. Doing so can help platforms balance the objective of encouraging supplier participation with that of sustaining longer-term engagement and retention.

This paper contributes to two strands of literature. First, it contributes to the literature on user responses to recommender systems on digital platforms (Banerjee et al. 2026). Prior work in B2C settings shows that recommendations increase clicks and purchases on recommended items and also generate spillovers to non-recommended items (Oestreicher-Singer and Sundararajan 2012; Liang et al. 2019; Bairathi et al. 2024). However, it is unclear whether these findings carry over to B2B settings, given the distinctive features of B2B e-commerce noted earlier. While recent work has begun to explore recommender systems in B2B, most studies focus on algorithm design but provide little guidance on quantifying the impact of such systems (Nia et al. 2019; Yoon et al. 2021; Gaur and Liu 2020; Cho et al. 2023; Meng et al. 2024; Zhang and Liu 2025). Our paper is among the first to evaluate supplier responses to platform recommendations in B2B across both direct and spillover effects, offering a more comprehensive account of recommendation effects in B2B markets.

Second, it contributes to the design of auction-based B2B procurement platforms. In such settings, a bid only represents an entry into a competition rather than a guaranteed transaction. Prior work examining how design levers such as bidding policy and ranking rules shape supplier behavior tends to focus on either bidding participation or transaction outcomes, but not on both jointly (Allu and Gaur 2024; Tucker and Zhang 2010; Pilehvar et al. 2017; Bimpikis et al. 2020). This separation is limited because an intervention that drives bidding activity may not improve transaction success, and evaluating either metric in isolation risks an incomplete and potentially misleading picture of their effectiveness. We provide a more comprehensive assessment of how platform recommendations shape the supplier journey—from bidding participation to transaction outcomes—in B2B procurement.

The rest of this paper is organized as follows. We first introduce the research context and data, followed by a description of the empirical approach. We then present supplier responses to the recommended and non-recommended RFQs and explore the mechanisms underlying these behaviors. We conclude with a discussion.

EMPIRICAL SETTING AND DATA

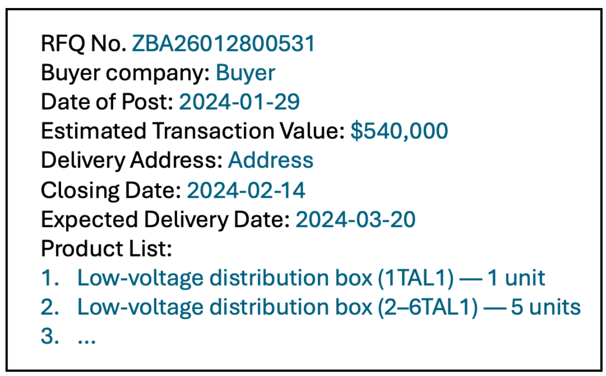
Empirical Setting

Our empirical setting is Platform Z, a construction-sector B2B e-commerce platform in China launched in 2015. By the end of 2025, Platform Z had connected more than 61,000 registered suppliers with more than 130,000 RFQs from over 400 buyers. The total transaction value on the platform exceeds 140 billion CNY (19 billion USD). At registration, suppliers provide profile information including service area coverage, product categories, and ownership type.

Buyers on Platform Z post RFQs that specify the requested products, order quantities, estimated transaction value, delivery date, delivery address, and bid closing date. Some

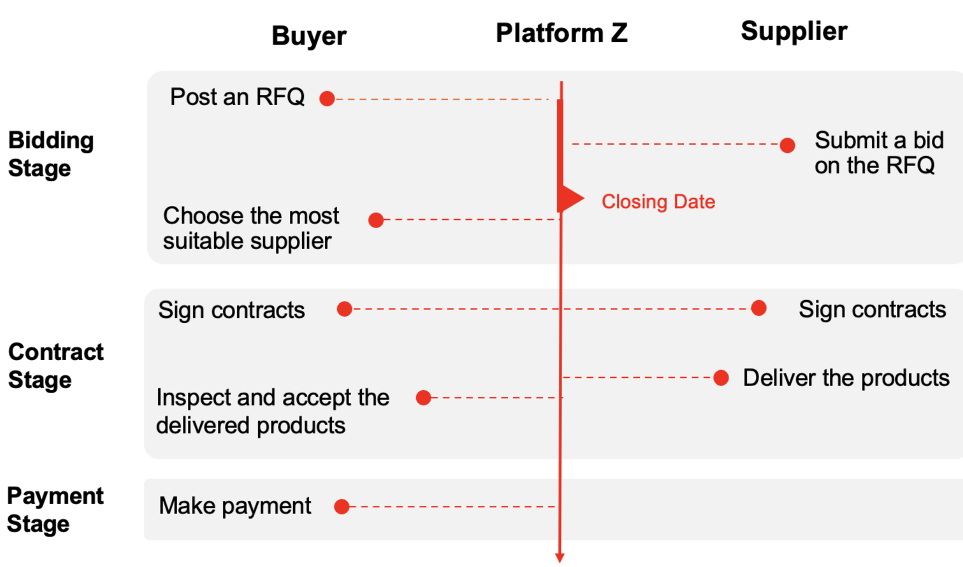
RFQs may request multiple products. For example, an RFQ from a real estate developer may request 10 tons of Round Steel (Q235) and 20 rolls of Wire Rod. Figure 1 shows an example of an RFQ. Once posted, RFQs become publicly visible to suppliers that can browse listings and submit bids before the closing date.[3] After bidding closes, buyers review submitted bids, screen potential suppliers, and select a single winning supplier for each RFQ. A transaction is completed once the supplier fulfills the delivery requirements and the buyer makes the corresponding payment. Figure 2 illustrates this process.

Figure 1: Example of a Request-for-Quotation (RFQ) Posted on Platform Z


RFQ No. ZBA26012800531
Buyer company: Buyer
Date of Post: 2024-01-29
Estimated Transaction Value: $540,000
Delivery Address: Address
Closing Date: 2024-02-14
Expected Delivery Date: 2024-03-20
Product List:

1. Low-voltage distribution box (1TAL1) — 1 unit
2. Low-voltage distribution box (2–6TAL1) — 5 units
3. ...

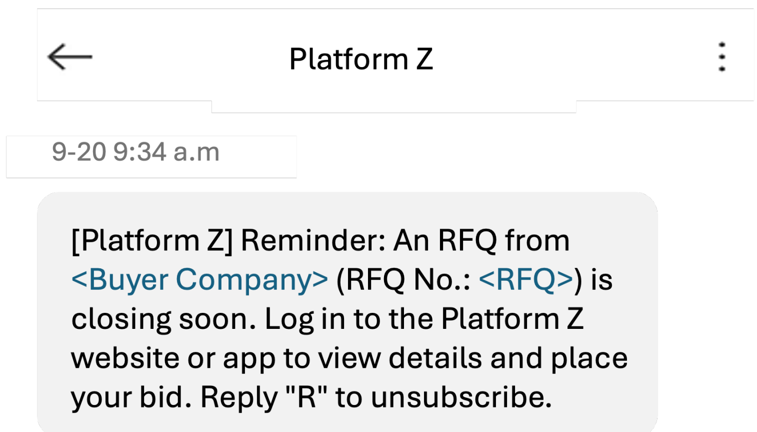
Figure 2: Procurement Process on Platform Z



[3] The platform's webpage lists RFQs in reverse chronological order, with the same ranking shown to all suppliers.

On April 27, 2020, Platform Z introduced RFQ recommendations delivered to suppliers via SMS. Figure 3 provides an exemplary text message, which highlights a recommended RFQ and prompt the focal supplier to bid. All messages follow a standardized template and vary only in the RFQ being recommended. The primary objective of these messages is to facilitate supplier participation by improving opportunity discovery in a crowded marketplace.

Figure 3: An Exemplary Personalized Recommendation



Recommendations are generated using machine learning models trained on supplier profile, recent platform activities (e.g., past bid count), and supplier–RFQ relevance based on product categories and geographic coverage.[5] For example, when an RFQ requests Stainless Steel for delivery in Hangzhou, the algorithm prioritizes suppliers whose product offerings and service regions meet these requirements. Table 1 summarizes the variables used in this recommendation process, those used in our matching procedure in Section 3.

[4]Note that variable BidPriorWeek is not explicitly used in the recommendation process, but we include it to enhance the comparability between treated and control units. This is crucial because recent bidding activity serves as a proxy for a supplier's short-term engagement with the platform, which is not fully captured by other variables and can influence both the likelihood of receiving a recommendation and the probability of responding to it.

[5]According to the platform, the recommendation system relies on Support Vector Machines and logistic regression models trained on historical bidding participation data. Throughout the study period, the systems typically recommend RFQs that have received fewer than three bids as of two days prior to the bidding deadline. This threshold is not particularly restrictive: the mean number of bids per RFQ at closing is three, and 62.2% of RFQs satisfy this criterion. During the study period, a small fraction of recommendations were generated manually by platform staff using a rule-based filtering approach based on the same set of characteristics.

Table 1: Key Variables Used in the Recommendation Process

Panel A: Supplier Characteristics	
Supplier Company Profile	
Ownership	Whether the company is non-private (state-owned) or private (1 = Non-private, 0 = Private)
BusinessFunction	Agent, Contractor, Manufacturer, Trading, and Other
Membership	Whether the company has a premium membership on the platform (1 = Premium, 0 = Basic)
Supplier Activities within the Platform	
ActiveStatus	Whether the supplier placed at least one bid in the past 180 days (1 = Active, 0 = Inactive)
PastResponseRate	Share of recommended RFQs for which the supplier submitted a bid
PastBidValue	Total value of all bids the supplier has placed
PastBidCount	Total number of bids placed by the supplier
WinningBidValue	Total value of bids won by the supplier
WinningBidCount	Total number of winning bids
BidPriorWeek[4]	Whether the supplier placed a bid in the previous week
Panel B: Supplier–RFQ Relevance	
ProductRelevance	Match between RFQ required products and supplier specialization
LocationRelevance	Match between RFQ delivery location and the supplier geographic service coverage

Note: This table summarizes the key variables used in the recommendation process. Membership indicates whether a supplier subscribes to the platform's premium service; premium members receive a waiver on transaction fees and financial service fees.

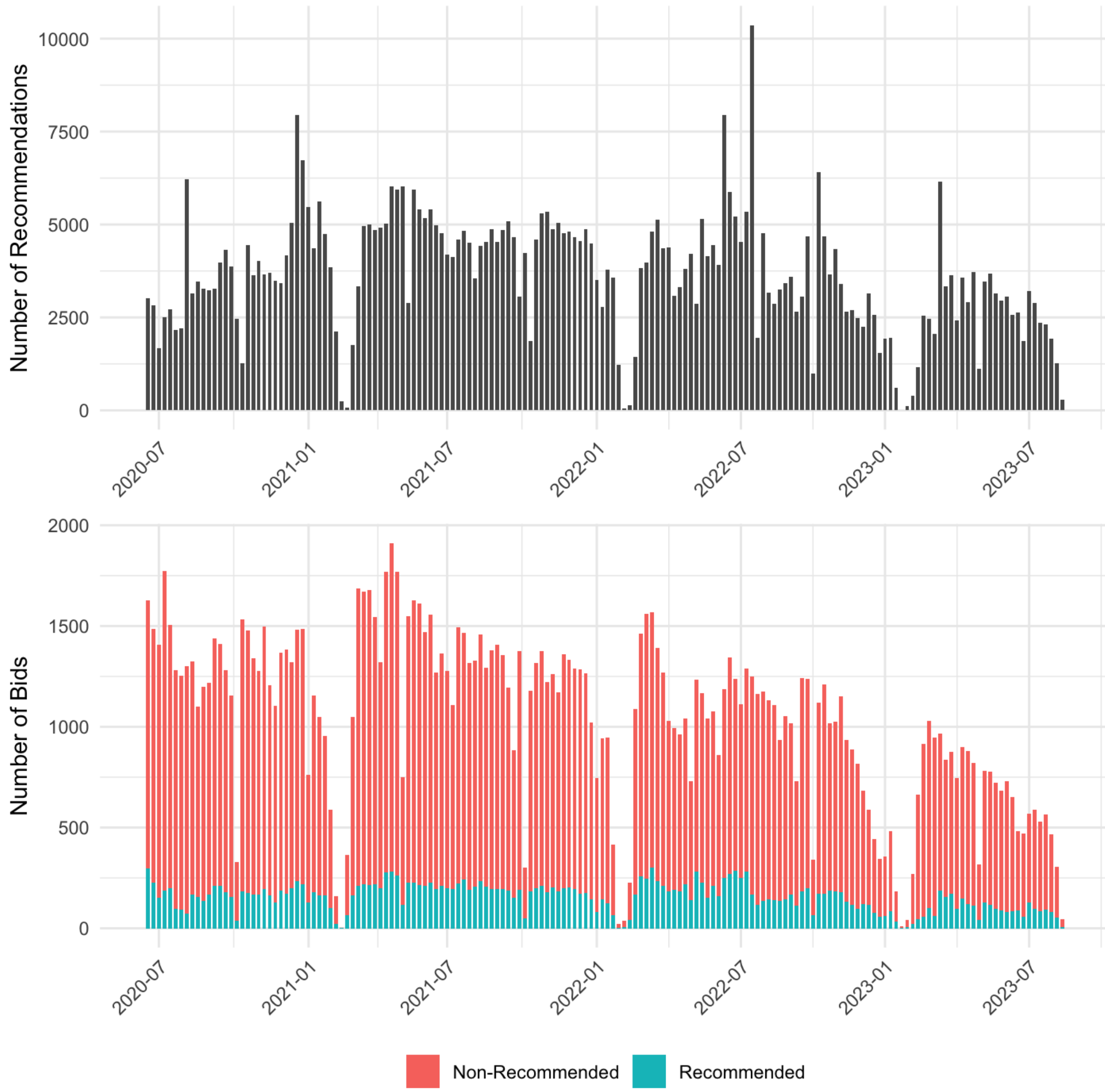
Data

The platform provided us with comprehensive data covering bids, RFQs, recommendations, and user profiles. The bid dataset records the universe of bidding activities on the platform, including the supplier submitting the bid, the associated RFQ, the submission timestamp, and the bid outcome (i.e., whether the bid won the contract). The RFQ dataset provides detailed information on each procurement request, including the buyer identity, posting time, estimated transaction value, requested products, bid closing date, and the winning supplier, and transaction price. The recommendation dataset captures every recommendation delivered since the introduction of personalized recommendations and documents the recommended RFQ, the recipient supplier, and the recommendation timestamp. Finally, the user profile dataset offers rich information on both sides of the marketplace. For suppliers, we observe registered capital, tenure, ownership type, geographic service coverage, product category specialization, and membership tier. For buyers, we observe registered capital, location, tenure, and procurement category.

Our analysis covers the period from April 27, 2020 (the launch of platform recommendations) to August 11, 2023. By the end of this period, the platform hosted 32,372 suppliers and 53 buyers. On the demand side, 32 buyers posted 52,538 RFQs during this period. The average expected transaction value of an RFQ is 2 million CNY (approximately 0.3 million USD). Each RFQ attracted an average of 3 bids (25th percentile = 1; 75th percentile = 4), and 8% of RFQs received no bids. Of all RFQs posted, 38,164 (72.65%) resulted in successful transactions. On the supply side, 25,209 suppliers submitted at least one bid during the study period. An average supplier placed approximately 6 bids—roughly one every two months—and secured about 1.5 winning bids. This relatively low bidding frequency is consistent with the high-stakes and selective nature of participation in B2B procurement. The average supplier's bidding success rate is approximately 33.3%, implying that roughly one in three bids results in a successful transaction.

During the study period, the platform sent 621,827 recommendations to 32,234 suppliers, covering 36,280 RFQs (62.2% of all RFQs posted). On average, each recommended RFQ was sent to 17 suppliers. Each supplier received about 19 recommendations, although the distribution was highly skewed (Median = 4; S.D. = 67.5). The top panel of Figure 4 plots the volume of weekly recommendations on the platform over time.

Of the 621,827 recommendations, 21,011 resulted in bids, corresponding to an average response rate of 3.4%. This response rate is consistent with the relative low baseline level of supplier participation. The bottom panel of Figure 4 plots the weekly bid volume on the platform separately for bids where the supplier received a recommendation to the corresponding RFQ (Recommended) and bids where the supplier did not receive a relevant recommendation (Non-recommended). Although the majority of bids arise from the latter, recommendation-driven bids account for a substantial and stable share of overall bidding activity on the platform.



Note: The top panel plots the weekly number of recommendations sent on the platform. The bottom panel plots the weekly number of bids submitted by suppliers, where red bars denote non-recommended (spontaneous) bids and blue bars denote bids submitted following a recommendation. The temporary declines in activity around February each year coincide with the Chinese New Year holiday, during which RFQ activity on the platform was substantially lower.

Figure 4: Weekly Number of Recommendations and Bids at Platform Level

EMPIRICAL STRATEGY

Our goal is to estimate the causal effects of platform recommendations on supplier bidding participation. In an ideal experiment, the platform would identify a set of candidate suppliers with similar characteristics and randomly recommend an RFQ to a subset of them (treatment), while withholding the recommendation from the remaining suppliers (control). We would then compare bidding participation between recommendation recipients and non-recipients to identify the causal effects of platform recommendations. In practice, however, implementing such an experiment is costly and operationally challenging in a high-stakes B2B procurement environment.

We approximate this ideal experiment using a matching-based design. For each recommendation in which RFQ j is recommended to supplier i on day t, we aim to identify a set of control suppliers that are comparable along dimensions relevant to treatment assignment and outcome, as informed by the institutional setting. First, we consider the characteristics used by the platform's curation procedure, including supplier profile characteristics, prior participation levels, and RFQ–supplier relevance. Second, we account for suppliers' prior recommendation histories. This is important because suppliers may receive multiple recommendations at different points in time (see Figure 5 for an illustration). Prior recommendations may affect both subsequent treatment assignment and subsequent participation: recommendation assignment depends partly on past participation, which may be shaped by earlier recommendations, and prior recommendations may also exert carryover effects on subsequent bidding participation.[6] We adopt the PanelMatch framework of Imai et al. (2023), as detailed later, to jointly account for dynamic treatment histories and time-varying covariates when constructing control observations.

Our main identifying assumption is that, conditional on the matching dimensions described above, the assignment of RFQ j to treated and control suppliers is plausibly ex-

[6] For example, a supplier who receives a recommendation at day $t-3$ may increase its participation in the subsequent days.

ogenous (Robins et al. 2000). This assumption is credible given our matching-based design, which is informed by detailed institutional knowledge of the platform's recommendation procedure. We validate the assumption through a series of checks, including a balance check on observed covariates and the bias-adjusted coefficient approach (Oster 2019) that evaluates the potential influence of unobservables.

Another identifying assumption is the Stable Unit Treatment Value Assumption (SUTVA) (Rubin 1980). SUTVA requires (1) no interference (i.e., the potential outcome of one unit is unaffected by the treatment status of other units) and (2) no hidden variations in treatment (i.e., each unit receives a well-defined version of the treatment). Both conditions are plausible in our setting. Regarding the first, the platform's bidding process is sealed: suppliers cannot observe who else has bid on a given RFQ or the prices submitted by others. This institutional feature limits the extent to which one supplier's bidding decision can be influenced by whether other suppliers receive recommendations for the same RFQ.[7] Regarding the second, the platform's recommendation algorithm and delivery format remain stable throughout our study period, ensuring that "receiving a recommendation" carries a consistent meaning across suppliers and over time.

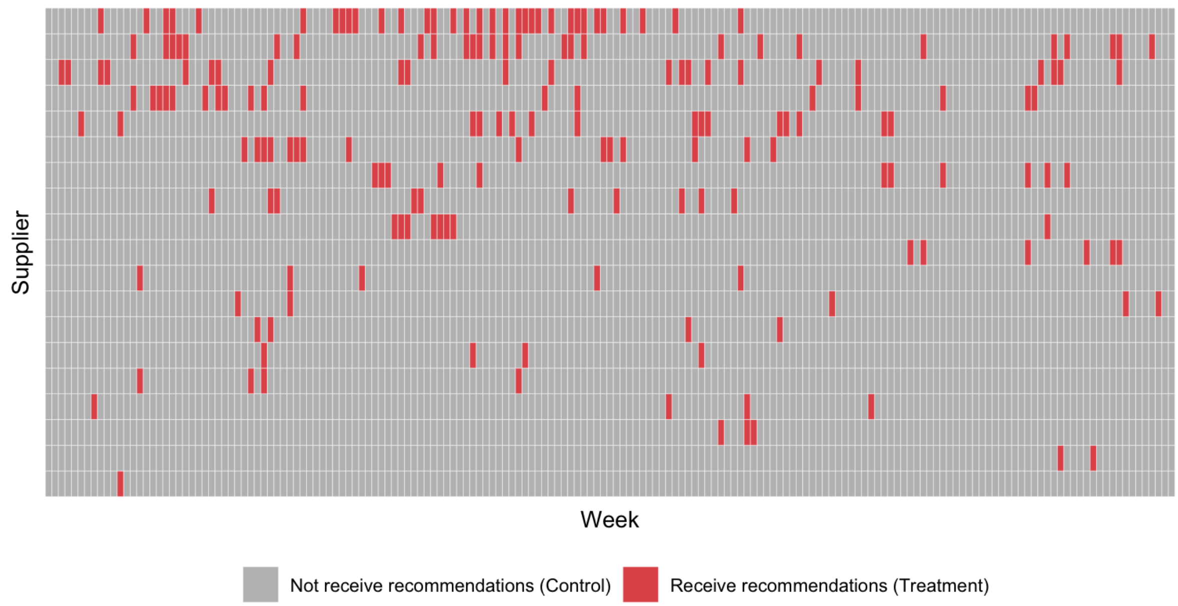
Matching Procedure

Our matching procedure proceeds in two steps.

Step 1: Matching on recommendation history. For supplier i that receives the recommendation of RFQ j at day t, we select a set of control suppliers that do not receive the recommendation and have an identical treatment history for a pre-specified time span, denoted as M_{ij_t}. Figure WA2 in the Web Appendix illustrates an example of this procedure. We aggregate the recommendation history at the weekly level and focus on a window of four weeks prior to the treated observation.[8] This step yields 280,759 treated observations and

[7] This reasoning applies to bidding participation but not necessarily to the bidding outcome, which we discuss in detail in subsequent sections.

[8] Since the platform rarely assigns multiple recommendations to a single supplier within a week, this approach preserves sufficient variation in historical treatment while remaining computationally feasible. For suppliers with less



Note: This figure displays the distribution of treatment for 20 randomly selected suppliers between April 27, 2020 and August 11, 2023, in which the blue (red) rectangle represents a control (treatment) supplier-day observation. We aggregate the data at weekly level for ease of illustration.

Figure 5: Treatment Distribution across Suppliers and over Time

4,144,282 control observations, with each treatment unit matched to an average of 15 control units in M_{ij_t}. We exclude 111 treated observations for which no same-history match exists.

Step 2: Refinement based on supplier characteristics. We refine the matched sets using supplier characteristics that enter the platform's recommendation procedure, including supplier profile, historical activity, and supplier-RFQ relevance, as summarized in Table 1.

We begin with supplier–RFQ relevance. As the relevance is not a supplier-specific variable, it cannot be incorporated directly into the matching framework. We instead require that treated and control observations are comparable in product offerings and service areas. To this end, we construct two sets of supplier embeddings for product categories and geographic service coverage, respectively, and calculate the cosine similarity between the treated observation and each of its candidate controls in M_{ij_t} (for details, see Web Appendix D). We retain only controls whose cosine similarity with the treated observation exceeds 0.9 on both

than four weeks of tenure on the platform, we restrict potential matches to those with the same length of tenure.

dimensions, and denote the refined matched set as M'_{ij_t}. This threshold is selected to balance relevance and sample size; 9.75% of candidate controls are retained at this threshold.[9]

Next, we balance the remaining supplier characteristics using coarsened exact matching (CEM). The platform's recommendation algorithm is based on flexible predictive models that may incorporate nonlinearities and interactions among supplier characteristics. CEM does not impose functional-form assumptions or require pre-specifying interaction terms, allowing our empirical design to better approximate the underlying assignment mechanism and mitigate concerns about misspecification arising from unobserved interaction structures. Because supplier characteristics are predominantly categorical, CEM achieves exact balance on these variables. For each treated supplier i on day t, we pool it with its candidate controls in M_{ij_t} and apply CEM to stratify these observations based on supplier characteristics. We retain controls that fall into the same stratum as the treated observation.[10]

While CEM helps balance categorical variables, it may leave imbalances in continuous variables within strata (Black et al. 2022). To enhance the balance in continuous variables, we apply propensity score matching (PSM) on the CEM-refined sample.[11] In our context, the propensity score is the probability that an individual supplier receives a recommendation, conditional on the observable supplier characteristics used in the recommendation procedure. We estimate propensity scores via a logistic regression on the pooled sample of all treated observations and their candidate controls. We then apply nearest-neighbor matching with replacement to select up to five controls with the closest propensity scores for each treated observation (i.e., $|M'_{ij_t}| \leq 5$). The resulting matched sample contains 25,795 treated observations, with an average of 3.2 control observations for each. Each treated observation and

[9] We also examine additional thresholds of the cosine similarity: 0.8 and 0.95. 10.11% of candidate controls are retained at the threshold of 0.8 and 8.56% of candidate controls are retained at the threshold of 0.95. Sensitivity analyses using alternative thresholds yield similar results, and additional balance checks show no meaningful improvement in covariate balance under stricter thresholds.

[10] The resulting sample contains 26,203 treated units and 286,592 control units, with each treated unit matched to an average of 10.9 controls.

[11] This hybrid matching procedure follows the extension proposed by Iacus et al. (2012) in their discussion of "Combining CEM with Other Methods", which combines the strengths of nonparametric stratification and fine-grained balance refinement. Similar approaches have been adopted in recent Economics and Finance research (e.g., Makioka 2021; Bertoni et al. 2023).

its matched set M'_{ijt} form a matched pair surrounding the focal RFQ j.

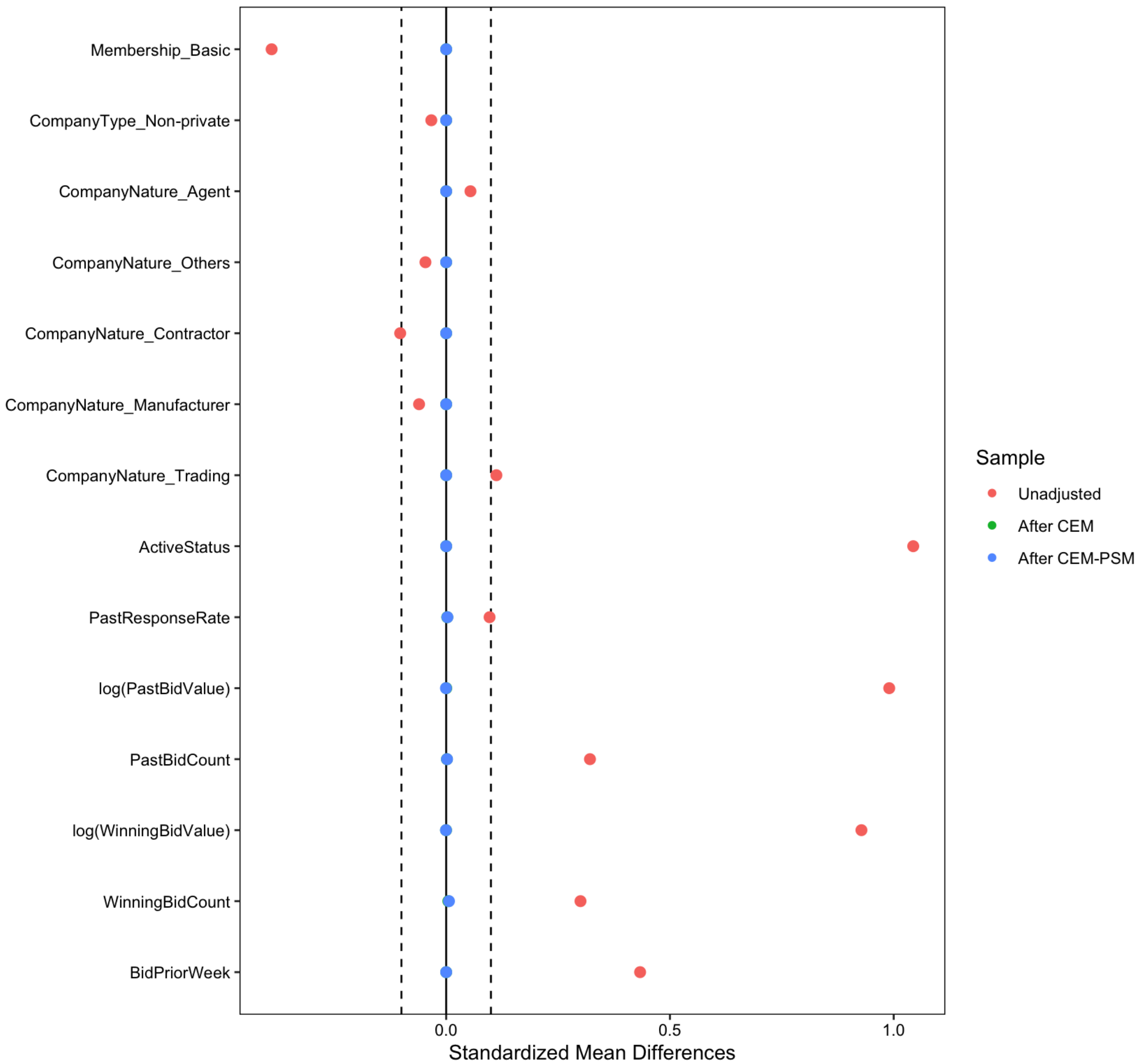
Figure 6 compares the standardized means in covariates between the treatment and control groups before and after matching. The covariate means are substantially different across groups before matching, while they are almost identical after matching. Figure WA3 further shows that the treated and control groups exhibit similar distributions across each covariate. These suggest that the matching procedure successfully removes systematic differences in observed characteristics that may confound the treatment assignment.

Model Specification

We estimate the effect of platform recommendations on supplier bidding participation using the matched sample. For each treated supplier who receives a recommendation of RFQ j on day t and its matched control in M'_{ijt}, we measure three types of bidding participation defined over a three-day window following day t.[12] The first metric captures bidding participation in the recommended RFQ ($Direct_{ijt}$), defined as whether the supplier submits a bid on RFQ j. The second captures bidding participation in other RFQs ($Spillover_{ijt}$), defined as whether the supplier submits a bid on other RFQs. The third measures overall bidding participation ($Overall_{ijt}$), i.e., whether the supplier bids on any RFQs.

Our estimation pools observations across all matched pairs surrounding each recommendation and estimate a regression model with one coefficient capturing the average treatment effect across all recommendations (Sahni et al. 2017). Pooling maximizes statistical power by fully exploiting the matched sample. An alternative approach is to estimate a separate treatment effect for each "experiment" and compute a sample-weighted mean effect. This is not feasible in our context, as individual experiments contains too few observations to

[12] We use a three-day window for several reasons. First, RFQs typically have short bidding windows. The average bidding window across RFQs is 4.08 days, with 72.4% of bids occurring within the first three days. Among recommendations that result in bidding, 90.9% occur within three days while only 39% of bids occur on the same day, indicating that bid preparation requires time and that one- or two-day windows are likely too short to capture supplier responses. Second, recommendations to the same supplier generally recur at intervals longer than three days: 90% of suppliers receive their next recommendation more than three days after the previous one. Together, a three-day window captures the majority of supplier responses while limiting the overlap with subsequent recommendations.



Note: This figure compares standardized covariate means between treated and control observations before and after the matching procedure.

Figure 6: Comparison of Covariate Balance Before and After the Matching Procedure

yield precise estimates. We restrict our sample to treated suppliers who receive exactly one recommendation on a given day.[13]

Accordingly, we estimate the average treatment effect as follows:

$$Participation_{ijt} = \beta Recommendation_{ijt} + \alpha_i + \delta_t + \xi_j + \varepsilon_{ijt} \tag{1}$$

where $Participation_{ijt} \in \{Direct_{ijt}, Spillover_{ijt}, Overall_{ijt}\}$ and $Recommendation_{ijt}$ indicates whether supplier i received the recommendation of RFQ j on day t. We include supplier fixed effects, α_i, to account for time-invariant factors that are specific to the supplier, day fixed effects, δ_t, to capture seasonal patterns in suppliers' bidding activities, and RFQ fixed effect, ξ_j, to control for time-invariant characteristics of each RFQ. We cluster standard errors at the supplier level to account for potential correlations among multiple observations of a supplier. We use a linear probability model to facilitate the interpretability of coefficients (Caudill 1988) and handle the number of fixed effects (Goldfarb and Tucker 2011; Guo et al. 2025; Lee and Hosanagar 2021).

EFFECTS OF PLATFORM RECOMMENDATIONS ON SUPPLIER BIDDING PARTICIPATION

Main Results

Table 2 presents the estimation results. We find that recommendations increase the probability of bidding on recommended RFQs by 1.4 percentage-point ($p < 0.01$) and that of bidding non-recommended RFQs by 0.7 percentage-point ($p < 0.01$). As a result, the overall bidding participation of recommendation recipients increases by 1.9 percentage points ($p < 0.01$). These findings indicate that recommendations generate not only a direct effect on targeted RFQs but also a positive spillover effect on the platform.

[13]They account for 92.1% of the full sample. This restriction is necessary for clear identification because it is not possible to attribute the spillover effect to any single recommendation for suppliers receiving multiple recommendations on the same day. Results are consistent when we use the full sample (see Web Appendix E).

Although these effects may appear modest in absolute terms, they imply substantial relative lift estimates given the platform's low baseline bidding rate. Specifically, the average probability of spontaneously bidding on an RFQ in the absence of a recommendation on a given day is only 1.24 percentage points.[14] Relative to this baseline, the estimated effects on *Direct*, *Spillover*, *Overall*, correspond to a lift of 121%, 65%, and 161%, respectively. Given the substantial economic value of B2B transactions, the magnitude of these effects is sizable.

The low baseline response rate reflects key features of the B2B setting. First, B2B transactions typically involve large monetary stakes and firm-wide decision-making, which naturally constrain suppliers' bidding frequency. Second, a large share of registered suppliers remains inactive on the platform—one of the primary motivations for the platform's adoption of personalized recommendation algorithms.

Table 2: Effects of Recommendations on Supplier Bidding

	Direct	Spillover	Overall
Recommendation	0.014***	0.007***	0.019***
	(0.001)	(0.001)	(0.002)
Fixed effects	Yes	Yes	Yes
Num. obs.	114,633	114,633	114,633
R-sq.	0.455	0.442	0.455

Note: This table reports regression estimates of the effects of recommendations on supplier bidding outcomes. The "Overall" column reflects the combined effects from both direct and spillover channels. Standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

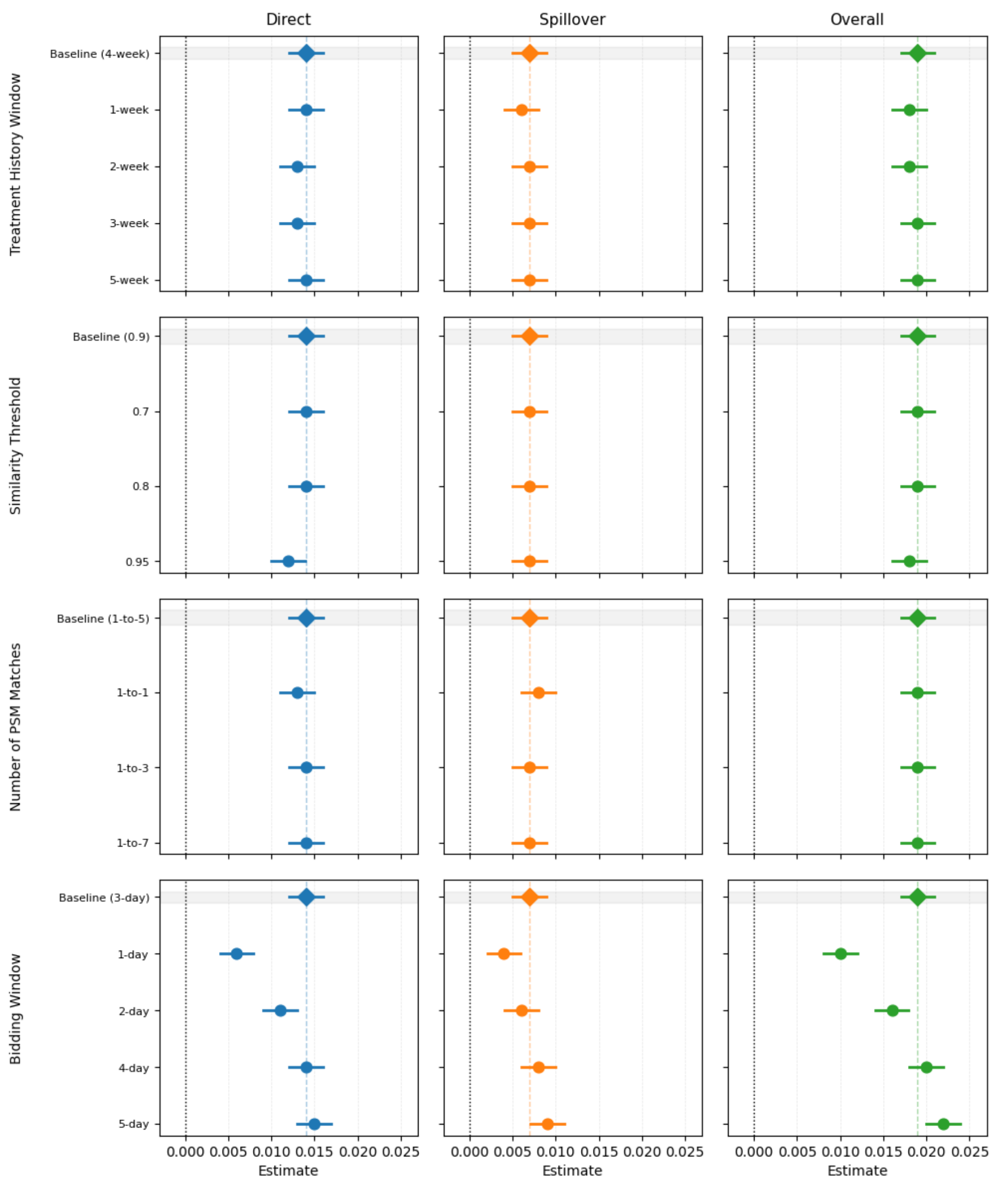
Robustness Checks

Our main identifying assumption is that matching based on supplier characteristics and past recommendation history sufficiently removes systematic differences that may confound

[14]This number is computed using control observations and equals the fraction of supplier–day observations in which the supplier submits bids to any RFQ among all control supplier–day observations (i.e., days on which the supplier does not receive a recommendation).

the treatment assignment. This assumption may be violated if omitted confounders are present. We assess the sensitivity of our findings to omitted variable bias.

Figure 7: Robustness Check: Alternative Matching Specifications



Note: Columns correspond to three outcomes: *Direct*, *Spillover*, and *Overall*. Each row varies one component of the empirical design while holding all other elements fixed. Diamond markers (with shaded background) denote estimates from the baseline specification in each panel, and circular markers represent alternative specifications. Horizontal bars indicate 95% confidence intervals. Vertical dashed lines mark the baseline point estimates, and vertical dotted lines indicate zero.

4.2.1 Alternative Matching Specifications

Omitted variable bias may arise from the matching procedure. We therefore test sensitivity to four matching specifications, with all results shown in Figure 7. First, our main matching procedure requires that treated and control suppliers have identical treatment histories over a prior window of four weeks. To assess sensitivity, we consider alternative windows ranging from one to five weeks, as well as a specification that imposes no treatment history controls. Second, we vary the similarity threshold for product offerings and service area, using alternative cutoffs of 0.7, 0.8, and 0.95 in place of the baseline value of 0.9. Third, we vary the maximum number of retained candidates (nearest neighbors) in matched sets M to 1, 3, and 7, compared to 5 in the baseline. Fourth, we vary the engagement measurement window from one to five days, compared to the baseline three-day window. The estimated effects are robust across all alternative specifications, with one exception: the one-day and two-day engagement measurement windows yield smaller effects for *Direct*, *Spillover*, *Overall*. This is consistent with the notion that suppliers often respond with a short delay, as bid preparation typically requires time for internal coordination and evaluation. Nevertheless, the overall results are not materially sensitive to any of these design choices, reinforcing the robustness of our findings.

4.2.2 Alternative Model Specification

In the recommendation generating process, supplier-RFQ relevance is one of the key inputs, including product relevance and location relevance. Although our matching procedure ensures comparability between treated and control suppliers in terms of product categories and geographic coverage, it does not directly account for supplier-RFQ relevance. To address this concern, we augment the baseline specification in Equation (1) with a vector of supplier-RFQ relevance controls X_{ij}, which includes $LocationRelevance_{ij}$ and $ProductRelevance_{ij}$. $LocationRelevance_{ij_t}$, is a binary indicator equal to 1 if the requested delivery address of RFQ j lies within supplier i's service area, and 0 otherwise. The $ProductRelevance_{ij}$

draws on platform Z's product taxonomy, which organizes products using a hierarchical domain–category–product catalog.[15] While RFQ requests are recorded at the product level, supplier specialization is documented at the category level. To reconcile this mismatch, we map products to their parent categories and define $ProductRelevance$ as the fraction of requested categories in an RFQ that the supplier is able to serve. For example, if an RFQ requests products from three categories and the supplier can provide two, then $ProductRelevance = 0.67$. The average $ProductRelevance$ value across suppliers who receive recommendations is 75% in our sample. Table 3 shows the results. The inclusion of relevance controls leaves our estimates largely unchanged.

Table 3: Robustness Check: Inclusion of Relevance Controls.

	Direct	Spillover	Overall
Recommendation	0.014***	0.008***	0.019***
	(0.001)	(0.001)	(0.002)
Relevance controls	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Num. obs.	114,633	114,633	114,633
R-sq.	0.455	0.442	0.455

Note: This table reports regression estimates of the effects of recommendations on supplier bidding outcomes with relevance controls. The "Overall" column reflects the combined effects from both direct and spillover channels. Standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

4.2.3 Sensitivity Analysis on Unobservables

We next conduct a sensitivity analysis following the approach developed by Oster (2019). This approach examines the robustness of treatment effect estimation by observing coefficient

[15] Domains represent broad material groupings (e.g., Ferrous Metals, Cement and Concrete). Categories divide each domain into functional groups (e.g., Rod Products and Steel Components under Ferrous Metals). Products capture specific material types within each category (e.g., Wire Rod and Round Steel (Q235) under Rod Products).

and R-sq movements after the inclusion of controls. Consider the following regressions:

$$Engagement = \beta_0 Recommendation + \varepsilon \tag{2}$$

$$Engagement = \beta_1 Recommendation + X_o + \varepsilon \tag{3}$$

$$Engagement = \beta^* Recommendation + X_o + X_u + \varepsilon \tag{4}$$

where X_o represents the impact of observables and fixed effects, and X_u represents the impact of unobservables. Thus, Equation (2) is the baseline regression without controls. Equation (3) controls for observables. Equation (4) is a hypothetical regression that accounts for both observables and unobservables, which yields the bias-adjusted treatment effect β^*. Oster (2019) proposes an approximation of the bias-adjusted treatment effect:

$$\beta^* = \beta_1 - \delta\,[\beta_0 - \beta_1]\,\frac{R^2_{max} - R^2_1}{R^2_1 - R^2_0}, \tag{5}$$

where R^2_0, R^2_1, and R^2_{max} are the R-sq values from Equation (2) to (4); the variable δ is the relative degree of selection on unobservables to observables.[16] A value of $\delta = 1$ suggests that the observables are at least as important as the unobservables.

Table 4 reports the results and the parameters used for the sensitivity analysis. For each dependent variable, β_0 and R^2_0 are estimated by removing all covariates and fixed effects from Equation (1). The estimates for β_1 and R^2_1 are taken from Table 2. Following Oster (2019), we set $\delta = 1$ and $R^2_{max} = 1.3 \times R^2_1$, and compute the bias-adjusted treatment effect β^* under these assumptions.

We find that adding observable controls and fixed effects leads to a substantial increase in explanatory power. Across specifications, the R^2 increases sharply from the baseline regressions to the controlled regressions. In contrast, the estimated treatment effects exhibit only modest changes from β_0 to β_1. Notably, the value of β^* remains similar to our main

[16]Specifically, define the proportional selection relationship as $\delta\frac{\sigma_{o,price}}{\sigma_o^2} = \frac{\sigma_{u,price}}{\sigma_u^2}$, where $\sigma_{s,price} = cov(X_s, price)$, $\sigma_s^2 = var(X_s)$ for $s \in \{o, u\}$. Thus, δ can be interpreted as to what extent the bias from observed characteristics is informative about the bias from unobserved characteristics.

Table 4: Robustness Check: Bias-adjusted Coefficients and R^2_{max} Values

	β_0	β_1	R^2_0	R^2_1	$R^2_{\mathbf{max}}$	β^*	δ
Direct	0.018	0.014	0.013	0.455	0.592	0.013	4.450
Spillover	0.012	0.008	0.002	0.442	0.575	0.007	5.476
Overall	0.022	0.019	0.007	0.455	0.592	0.018	3.958

Note: This table reports the Oster (2019) sensitivity analysis for the effect of recommendations on engagement. β_0 denotes the coefficient from the uncontrolled regression (Equation (2)), and β_1 denotes the coefficient from the specification including observables and fixed effects (Equation (3)). R^2_0 and R^2_1 are the corresponding R-squared values. We set $R^2_{\max} = 1.3 \times R^2_1$ following Oster (2019). β^* reports the bias-adjusted treatment effect under $\delta = 1$. The final column reports the value of δ required to nullify the estimated effect. Values of $\delta > 1$ indicate that selection on unobservables would need to be stronger than selection on observables to fully explain away the effect.

results β_1. The interval formed by β^* and β_1 also excludes zero, further confirming the robustness of our results (Oster 2019).[17]

Oster (2019) also proposes another sensitivity analysis by computing the value of δ required to nullify the results, given an assumed R^2_{max} value. A required value of $\delta > 1$ indicates that unobservables must be more important than observables to produce a treatment effect of zero, thereby implying robustness to omitted variable bias. In our setting, the required δ is 4.450, 5.476, and 3.958 for *Direct*, *Spillover*, and *Overall*, assuming $R^2_{max} = 1.3 \times R^2_1$. These magnitudes imply that unobserved factors would need to be nearly four times as important as observables to explain away the effects. Given that our main specification already accounts for the key inputs of the platform's recommendation algorithm, along with rich fixed effects, it appears implausible that unobserved factors would exert such a disproportionately large influence in this setting.

[17] It is noteworthy that the bias-adjusted treatment effect is slightly smaller than our estimates. This indicates that, though negligible, the impact of unobservables likely includes desirable supplier features (e.g., supplier capability or supplier relevance with the focal RFQ) that are positively correlated with both treatment assignment and supplier engagement.

UNDERSTANDING THE SPILLOVER EFFECTS OF PLATFORM RECOMMENDATIONS

Our findings suggest positive direct effects on recommended RFQs, as well as positive spillover effects on non-recommended ones. The former is more intuitive and similar with prior evidence on the effectiveness of personalized recommendations on targeted users behavior (Kumar and Hosanagar 2019). The latter, however, is less straightforward and warrants further investigation.

We consider two complementary explanations. First, the recommendations themselves are poorly matched to suppliers, which prompts suppliers to search more broadly across the platform. Second, recommendations may work as an activation trigger, prompting suppliers to increase platform exploration as part of their strategic search for market opportunities.

Recommendation Relevance

A natural explanation is the relevance between the recommendation and the supplier profile. When a supplier receives recommendations that appear irrelevant, it may prompt suppliers to search for alternative RFQs on their own (Knijnenburg et al. 2012; Pu et al. 2011). If this mechanism is at play, we would expect spillover effects to be stronger when recommendations are less relevant to suppliers.

To test this, we extend our main specification by interacting the treatment indicator with $LocationRelevance_{ij_t}$, as defined in Section 4.2:

$$\begin{aligned} Spillover_{ij_t} = {} & \beta_0 Recommendation_{ij_t} + \beta_1 LocationRelevance_{ij_t} \\ & + \beta_2 Recommendation_{ij_t} \times LocationRelevance_{ij_t} + \alpha_i + \delta_t + \xi_j + \varepsilon_{ij_t}, \end{aligned} \tag{6}$$

where all notation follows the main specification in Equation (1).

Results are reported in Columns (1) of Table 5. The coefficient on *Recommendation* is 0.009 ($p < 0.01$), and that on the interaction term is -0.004 ($p < 0.05$). These estimates indicate that spillover effects are stronger when the recommended RFQ falls outside the supplier's service province, consistent with the proposed mechanism. Nevertheless, the absolute spillover effect remains positive in both the location-relevant and location-irrelevant groups, suggesting that while the difference is meaningful, spillovers persist regardless of location relevance. This implies that recommendation relevance alone does not fully account for the observed positive spillover.

We also estimate an analogous specification using $ProductRelevance_{ij_t}$, with results reported in Column (2) of Table 5. The interaction between recommendation and product relevance is not statistically significant. This null result likely reflects the limited variation in $ProductRelevance_{ij_t}$, which has a mean of 0.75 and a median of 1.00, as well as the granularity mismatch between supplier capabilities (recorded at the product level) and RFQ requirements (recorded at the category level).

Table 5: Heterogeneous Spillover Effects Across Suppliers.

Dep. Var.: *Spillover*	(1)	(2)	(3)
Recommendation	0.009***	0.005***	0.003*
	(0.001)	(0.002)	(0.002)
Recommendation × LocationRelevance	-0.004**		
	(0.002)		
Recommendation × ProductRelevance		0.003	
		(0.002)	
Recommendation × Breadth			0.006***
			(0.002)
Relevance control	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Num. obs.	114,633	114,633	114,633
R-sq.	0.452	0.452	0.452

Note: This table compares spillover effects of recommendations across suppliers with different characteristics. Column (1) examines heterogeneity by location relevance, Column (2) examines heterogeneity by product relevance, and Column (3) examines heterogeneity by the breadth of suppliers' product offerings. Standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Recommendation as An Activation Trigger

In B2B markets, suppliers manage multiple product categories and seek to maximize expected revenue under uncertainty in bidding outcomes. This creates strong incentives to diversify search across opportunities rather than focusing narrowly on any single RFQ. In this view, recommendations may act as an activation trigger, prompting suppliers to engage with the platform and proactively expand their exploration to other RFQs beyond the one recommended.

If this mechanism is at play, spillover effects should be stronger among suppliers with broader product portfolios, who face richer opportunity sets and stronger incentives to seek out outlets for their products. To test this, we measure product breadth as the number of categories a supplier serves and classify suppliers based on whether their breadth exceeds the sample median (three categories). We define an indicator $Breadth_i$ that equals one for above-median suppliers and zero otherwise, and interact it with the recommendation indicator. Column (3) of Table 5 reports the results. The interaction term is positive and statistically significant ($p < 0.01$), indicating that spillover effects are stronger for suppliers with broader product portfolios, a pattern consistent with the activation-based mechanism.

To further corroborate this interpretation, we examine whether spillover effects vary with suppliers' capacity utilization, proxied by the number of bids submitted in the prior week, denoted as $\#PastWeekBids_{it}$. If recommendations trigger proactive exploration rather than mechanical reactions, the spillover effects should depend on whether suppliers have sufficient capacity to act on them. Because bidding requires substantial attention and effort (e.g., quotation preparation), suppliers that have recently submitted many bids are likely operating near their capacity limits, leaving less bandwidth to pursue additional opportunities. Consistent with this prediction, we find that suppliers with higher bidding activity in the prior week exhibit attenuated spillover effects in the current period.

Table 6: Heterogeneous Spillover Effects by Recent Workload

	Spillover
Recommendation	0.007***
	(0.001)
#PastWeekBids	0.090***
	(0.007)
Recommendation × #PastWeekBids	-0.027***
	(0.008)
Relevance control	Yes
Fixed effects	Yes
Num. obs.	122,626
R-sq.	0.233

Note: This table examines how spillover effects vary with suppliers' recent bidding intensity. $\#PastWeekBids$ measures the number of bids submitted by a supplier in the seven days (one week) prior to the recommendation. All specifications include supplier and time fixed effects and control for product and location relevance. Standard errors are clustered at the supplier level and reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Ruling Out Intertemporal Substitution

An alternative interpretation of the spillover effect is intertemporal substitution: recommendations may simply shift the timing of bidding activity rather than generate additional engagement. Under this account, a supplier who would have bid on a non-recommended RFQ later in the week submits that bid earlier upon receiving a recommendation, producing the appearance of a spillover effect within our observation window without reflecting a genuine increase in overall platform engagement.

To rule out this explanation, we exploit the temporal structure of our data. Our baseline specification measures spillover as bidding on non-recommended RFQs within the three-day window following a recommendation (t to $t+2$). If intertemporal substitution is driving the results, we would expect to observe a compensating reduction in bidding activity in the period immediately after this window, as suppliers have already exhausted bids they would

Table 7: Ruling Out Intertemporal Substitution

	Spillover in Delayed Window
Recommendation	0.001
	(0.001)
Fixed effects	Yes
Num. obs.	107,504
R-sq.	0.474

Note: This table tests whether the spillover effect reflects intertemporal substitution rather than genuine additional engagement. The dependent variable measures bidding on non-recommended RFQs during the three-day window following the baseline observation period ($t+3$ to $t+5$). Suppliers who receive any new recommendations during this subsequent window are excluded. All specifications include supplier and time fixed effects and control for product and location relevance. Standard errors are clustered at the supplier level and reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

otherwise have placed later. We therefore construct an alternative outcome variable that captures bidding on non-recommended RFQs in the subsequent three-day window ($t+3$ to $t+5$)[18], and re-estimate the baseline specification on this delayed window. To ensure a clean test, we exclude suppliers who receive any new recommendations during this subsequent period, isolating the lingering effect of the original recommendation.

The results in Table 7 show no significant effect of the original recommendation on bidding participation on non-recommended RFQs in the $t+3$ to $t+5$ window. This absence of a delayed effect is inconsistent with intertemporal substitution, which would predict a negative effect (if bids were pulled forward). Instead, the pattern is consistent with a contemporaneous activation effect: recommendations stimulate exploration within a narrow window, after which bidding returns to its baseline level. We therefore conclude that the spillover effects documented in our main analysis reflect genuine additional engagement rather than

[18] This window is chosen to immediately follow our baseline observation period while remaining within the natural bidding horizon: in our data, the average bidding window across RFQs is 4.08 days, with 72.4% of bids occurring within the first three days, suggesting that any substituted bids would most plausibly materialise in this adjacent window.

a reallocation of pre-existing bidding intentions across time.

EFFECTS OF PLATFORM RECOMMENDATIONS ON TRANSACTION OUTCOMES

Moving beyond the effects of recommendations on supplier engagement, we now examine whether recommendation-induced engagement effectively translates into successful transactions. Answers to this are essential for understanding the economic value of recommendations in B2B markets.

Do Recommendations Increase Transactions?

In our setting, a supplier wins an RFQ if the buyer selects it as the final transaction partner. Each RFQ has at most one winning supplier. We define three success-related variables, each analogous to the engagement measures. The first is $DirectSuccess_{ijt}$, which equals one if the supplier wins the recommended RFQ within three days after receiving a recommendation. The second, $SpilloverSuccess_{ijt}$, equals one if the supplier wins any non-recommended RFQs within the three-day window following a recommendation. The third is $OverallSuccess_{ijt}$, equals one if the supplier wins any RFQs within the same three-day window. For each outcome, we re-estimate Equation (1) by replacing the dependent variable with the corresponding success-related variables.

Table 8 reports the estimation results. On average, suppliers who receive recommendations are 0.4 percentage points more likely to win RFQs ($p < 0.001$; Column (1)). Given that the average monetary value of an RFQ on the platform is approximately 2.0 million CNY (about 0.3 million USD), this effect corresponds to an expected increase of roughly 800 CNY (115 USD) in transaction value per recommendation exposure. This increase in *OverallSuccess* is largely driven by success in the recommended RFQs, which rises by 0.3 percentage points (Column (1)). In contrast, we do not observe a statistically significant increase in *SpilloverSuccess*, despite a significant increase in spillover bidding activity documented in Table 2. Importantly, this pattern suggests that supplier engagement with non-recommended RFQs is less likely to translate into successful transactions. We investigate the underlying reasons in the following section.

One concern is the violation of the SUTVA assumption. Recommendations stimulate increased bidding participation among treated suppliers. Because treated and control suppliers compete for the same RFQs, increased bidding intensity from the treated group can mechanically reduce the opportunity for control suppliers to win. As a result, the observed difference between treatment and control groups may be overstated (Kohavi et al. 2009, 2020; Blake and Coey 2014).

Given that our focus is on the spillover effect of recommendations on bidding success, this implies that the estimated effect on spillover success is likely an upper bound. In other words, the true effect of recommendation-induced spillover bidding on success is likely even smaller than our estimates, reinforcing the conclusion that such spillover bidding is inefficient in translating into successful outcomes.

Another concern is the possibility that this null result is driven by insufficient statistical power. To rule out this possibility, we conduct a post-hoc power analysis. With a sample size of 114,633 observations, our analysis achieves a statistical power of 0.863, well above the conventional threshold of 0.80, which would require only 96,340 observations. This confirms that the absence of a significant effect on $SpilloverSuccess$ reflects a genuine null result rather than an underpowered test.

Why Are Spillover Bids Inefficient?

To understand why spillover bidding is less effective at converting into successful transactions, we compare RFQs targeted by spillover bids with those targeted by other bid types.

We classify each supplier's bids into three mutually exclusive types. Recommended bids are bids placed on RFQs directly recommended to the supplier. Spillover bids are bids placed

Table 8: Effects of Recommendations on Successful Transactions.

	Overall Success	Direct Success	Spillover Success
Recommendation	0.004***	0.003***	0.001
	(0.001)	(0.000)	(0.001)
Relevance control	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Num. obs.	114,633	114,633	114,633
R-sq.	0.405	0.425	0.393

Note: This table reports regression estimates of the effects of recommendations on successful transaction outcomes. The "Overall Success," "Direct Success," and "Spillover Success" columns correspond to success in any RFQs, recommended RFQs, and non-recommended RFQs, respectively. Standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

on non-recommended RFQs within three days of receiving a recommendation.[19] Regular bids are bids that do not fall into the above categories and therefore reflect suppliers' baseline, spontaneous bidding activity.

During our study period, regular bids account for 54.4% of all bids, spillover bids for 30.6%, and recommended bids for 15.0%. For each bid type, we examine the average characteristics of the targeted RFQs along three dimensions: RFQ-supplier fit in product category and location ($ProduceRelevance$, $LocationRelevance$), and RFQ popularity, measured by the total number of bidders for the RFQ ($NumBidder$).

Table 9 shows that the RFQs targeted by recommended bids exhibit substantially higher product and location relevance than both regular bids and spillover bids ($p < 0.01$), indicating that the recommendations effectively directs suppliers toward well-matched opportunities. In contrast, RFQs targeted by spillover bids show relatively low RFQ-supplier relevance, with location relevance being the lowest among all bid types ($p < 0.01$). Moreover, spillover bids tend to target more popular RFQs that attract a larger number of bidders than those targeted by other bid types ($p < 0.01$). Taken together, these patterns suggest that spillover

[19]Because spillover bids are defined based on temporal proximity to recommendation exposure, they may include some regular bids that cannot be perfectly separated.

bids are disproportionately directed toward popular but poorly matched RFQs. This pattern suggests that spillover bidding is driven more by salience and opportunistic attention than by deliberate evaluation, which may help explain its lower conversion rate.

Table 9: Comparison of Bid Types

	# Obs	Product Relevance	Location Relevance	# Bidders
Recommended bids	27,701 (15.0%)	78.1%	47.3%	5.09
Spillover bids	56,561 (30.6%)	51.8%	31.1%	5.47
Regular bids	100,474 (54.5%)	48.4%	44.3%	5.08

Note: This table reports summary statistics by bid type. The unit of analysis is a bid. Product relevance measures whether the supplier's product offering matches the RFQ requested product in category level. Location relevance measures whether the supplier's service area covers the RFQ's requested delivery address. # Bidders denotes the total number of bidders participating in the corresponding RFQ.

Downstream Consequences of Bidding Failure

The above findings suggest that spillover bidding is more likely to result in unsuccessful outcomes. We next examine the downstream consequences of bidding failure by investigating its impact on supplier churn, a behavior that is central to the sustainability of the platform ecosystem.

We conduct a bid-level analysis. For each bid submitted by supplier i on day t for RFQ j, we observe whether the bid wins the targeted RFQ, denoted as $IfSuccess_{ijt}$. We define churning as supplier i not submitting any bids for at least 180 consecutive days following day t, denoted as $Churn_{ijt}$. We model supplier churn as a function of bidding outcomes:

$$Churn_{ijt} = \lambda IfSuccess_{ijt} + \epsilon_{ijt} \tag{7}$$

where the parameter λ captures the effect of bidding success on subsequent churn.

A simple OLS estimate of λ may be biased because $IfSuccess_{ijt}$ is endogenous: more capable or motivated suppliers are both more likely to win bids and less likely to churn. To address this concern, we exploit a regression-discontinuity-style instrumental variable (RD-IV) based on suppliers' price ranking within each RFQ. Empirically, 80% of winning bids are those offering the lowest price. Thus, suppliers whose bids narrowly lose (i.e., those ranking second-lowest on bidding price) provide a natural comparison group: they are observationally similar to winners but differ only by a narrow price margin.

Guided by the RD-IV design, we restrict the sample to bids with the lowest and second-lowest prices within each RFQ and construct a binary indicator, $IfLowestPrice_{ijt}$, which equals one if the bid offers the lowest price. We use $IfLowestPrice_{ijt}$ as an instrument for bidding success, $IfSuccess_{ijt}$. Under this design, our estimates identify a local average treatment effect (LATE), capturing the causal impact of narrowly winning rather than narrowly losing a bid.

We estimate Equation (7) using two-stage least squares (2SLS). Table 10 reports the first- and second-stage regressions. The results show that winning a bid significantly reduces the probability of subsequent churn by 8.3 percentage points ($p < 0.01$), which equivalently implies that narrowly losing a bid significantly increases the likelihood of churn relative to narrowly winning. This suggests that unsuccessful bidding experiences may discourage continued platform participation. Taken together, these findings highlight that while recommendations can activate additional bidding activity, the relatively low success rate of spillover bids may undermine suppliers' long-term engagement. This underscores the importance for platforms to monitor and facilitate suppliers' post-activation behavior following recommendation exposure.

DISCUSSION

Our study provides new evidence on how personalized recommendations operate in B2B procurement environments, a setting characterized by high stakes and multidimensional supplier capabilities. The results reveal that recommendations on these platforms function very differently than in consumer-facing markets. Rather than serving solely as cues directing users toward specific opportunities, recommendations act as reactivation triggers that shift suppliers into a broader search mode. Supplier responses are not limited to the focal opportu-

Table 10: Effect of Winning versus Losing on Supplier Churn

	First-Stage Regression	Second-Stage Regression
	$IfSuccess$	$Churn$
$IfLowestPrice$	0.351^{***}	
	(0.016)	
$IfSuccess$		-0.083^{**}
		(0.031)
Supplier FE	Yes	Yes
Num. obs.	3,797	3,797
R-sq.		

Note: The unit of analysis is a bid. The first stage regresses $IfSuccess_{it}$ on an indicator for being the lowest-price bidder. The second stage instruments $IfSuccess_{it}$ with the lowest-price indicator to estimate the causal impact of winning on churn behavior. Standard errors are reported in parentheses. The second-stage estimate represents a Local Average Treatment Effect (LATE) for suppliers whose bidding outcomes are determined by small price differences near the cutoff. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

nity but spill over into the wider landscape of RFQs, reflecting the strategic, revenue-seeking orientation of B2B participants. The activation effect we document is substantial. After receiving a recommendation, suppliers significantly increase bidding not only on the recommended RFQ but also on non-recommended ones.

However, the increased spillover engagement stimulated by recommendations does not significantly improve economic outcomes. Although suppliers' responses spill over to the wider landscape of RFQs, these efforts are less effective in converting into success than their responses to the recommended RFQs. Further analysis shows that such failures carry important implications for the platform. Unsuccessful bids increase the likelihood of supplier churn over the subsequent six months. Thus, by inducing low-efficiency spillover engagement, recommendations may unintentionally elevate long-term disengagement risk.

Managerial Implications

Our findings offer several insights for managers of B2B procurement platforms seeking to improve supplier participation, match quality, and marketplace efficiency. First, while recommendations are effective at stimulating supplier engagement, engagement alone is not a reliable proxy for performance in auction-based settings. Traditional recommendation systems developed for consumer platforms focus on maximizing views or clicks. In contrast, in B2B procurement, what ultimately matters is whether induced bids are competitive enough to succeed. Platforms should therefore complement engagement-based metrics with outcome-based metrics, such as the win probability of recommended bids or the competitiveness of bids submitted after recommendations.

Second, because recommendations trigger broader reactivation and spillover bidding, platform managers should consider designing recommendation systems that guide suppliers toward opportunities where they are likely to be competitive, rather than encouraging broad exploratory activity. Incorporating supplier-specific signals—such as historical win rates in particular product categories, prior interactions with buyers, or past bidding competitiveness—may help filter recommended RFQs so that suppliers focus on opportunities aligned with their capabilities. Such filters may be especially valuable for newer suppliers or larger suppliers that tend to over-engage following recommendations.

Finally, because winning RFQs increases future activity while losing reduces it, the quality of bids induced by recommendations affects long-term supplier retention. Platforms should therefore balance the goal of stimulating engagement with the risk of inadvertently increasing supplier churn by encouraging low-probability bids. Recommendation systems that emphasize match quality rather than bid quantity may enhance both short-run efficiency and long-run marketplace stability.

Limitations and Future Research

Several limitations of this study offer opportunities for future research. First, although we document that recommended and spillover bids have lower win rates, we cannot fully disentangle the underlying causes of this performance gap. Buyer requirements, unobserved technical specifications, and competitive dynamics within each RFQ may all contribute, but these factors are only partially observable in our data. Future work could incorporate richer data on buyer preferences or RFQ-level evaluation criteria to better understand the drivers of bid performance.

Second, our analysis relies on observational data, which limits our ability to fully identify causal effects. Although we apply matching methods to mitigate selection bias and conduct a range of robustness checks, unobserved confounders may still influence both the probability of receiving a recommendation and subsequent bidding behavior. Future research could build on our findings by leveraging randomized controlled trials or natural experiments that offer stronger causal identification.

Third, while we identify a reactivation mechanism that explains spillover bidding, we do not have direct measures of the cognitive or strategic processes underlying this behavior. Experimental or survey-based methods could complement our findings by probing how suppliers interpret recommendations, how they evaluate RFQs, and how they decide where to allocate attention after receiving recommendations.

Fourth, although we observe that winning reduces churn and losing increases it, our data do not allow us to model long-term strategic adjustments suppliers may make in response to persistent failures. Future research could examine whether suppliers learn to self-correct over time by becoming more selective in the recommendations they follow, or whether repeated exposure to low-success recommendations causes disengagement.

Finally, our analysis centers on a specific recommendation algorithm deployed by a single B2B platform in the construction industry. While this context is representative of B2B procurement environments, generalizability to other sectors or algorithmic designs remains an open question. As recommendation technologies continue to evolve, future work could explore how different design objectives, such as maximizing profitability, promoting fairness, or increasing diversity, affect supplier behavior and platform efficiency.

By acknowledging these limitations, we hope to encourage further research into how personalized interventions shape supplier behaviors and the whole platform in complex B2B environments, an area still relatively underexplored in the marketing literature.

REFERENCES

Allu, Rakesh and Vishal Gaur (2024), 'User engagement on an online b2b platform: How does quality of rankings matter?'. Available at SSRN: https://ssrn.com/abstract=4684782.

Bairathi, Mimansa, Xu Zhang and Anja Lambrecht (2024), 'The value of platform endorsement', *Marketing Science* . Forthcoming.

Banerjee, Shrabastee, Ishita Chakraborty, Hana Choi, Hannes Datta, Remi Daviet, Chiara Farronato, Minkyung Kim, Anja Lambrecht, Puneet Manchanda, Aniko Öry et al. (2026), 'Digital platforms 2.0: Emerging topics, opportunities, and challenges', *International Journal of Research in Marketing* .

Bertoni, Fabio, Massimo G. Colombo and Anita Quas (2023), 'The long-term effects of loan guarantees on sme performance', *Journal of Banking & Finance* **142**, 106572.

Bimpikis, Kostas, Wedad J Elmaghraby, Ken Moon and Wenchang Zhang (2020), 'Managing market thickness in online business-to-business markets', *Management Science* **66**(12), 5783–5822.

Black, Bernard, Parth Lalkiya and Joshua Y. Lerner (2022), The trouble with coarsened exact matching. Working paper.
URL: *https://ssrn.com/abstract=3694749*

Blake, Thomas and Dominic Coey (2014), Why marketplace experimentation is harder than it seems: The role of test-control interference, *in* 'Proceedings of the fifteenth ACM conference on Economics and computation', pp. 567–582.

Bledsoe, James (2023), 'ecommerce sales & size forecast', https://www.trade.gov/ecommerce-sales-size-forecast. Accessed: 2025-12-06.

Caudill, Steven B (1988), 'An advantage of the linear probability model over probit or logit', *Oxf. Bull. Econ. Stat.* **50**(4), 425–427.

Chakravarty, Anindita, Alok Kumar and Rajdeep Grewal (2014), 'Customer orientation structure for internet-based business-to-business platform firms', *Journal of Marketing* **78**(5), 1–23.

Cho, G., P.-s. Shim and J. Kim (2023), 'Explainable b2b recommender system for potential customer prediction using kgat', *Electronics* **12**(17), 3536.

Ganesan, Shankar (1994), 'Determinants of long-term orientation in buyer-seller relationships', *J. Mark.* **58**(2), 1.

Gaur, Vishal and Xiaoyan Liu (2020), 'Personalized recommendation system design for an online b2b platform'. Available at SSRN: https://ssrn.com/abstract=3902710.

Goldfarb, Avi and Catherine Tucker (2011), 'Online display advertising: Targeting and obtrusiveness', *Marketing Science* **30**(3), 389–404.

Grewal, Rajdeep, Anindita Chakravarty and Amit Saini (2010), 'Governance mechanisms in business-to-business electronic markets', *Journal of Marketing* **74**(4), 45–62.

Grewal, Rajdeep, Gary L Lilien, J Andrew Petersen and Stefan Wuyts (2022), Business-to-business marketing: Looking back, looking forward, *in* 'Handbook of business-to-business marketing', Edward Elgar Publishing, pp. 2–11.

Guo, Xue, Zhi (Aaron) Cheng and Paul A. Pavlou (2025), 'Skill-biased technical change, again? online gig platforms and local employment', *Information Systems Research* **36**(3), 1354–1374.

Hada, Mahima, Rajdeep Grewal and Gary L Lilien (2014), 'Supplier-selected referrals', *Journal of Marketing* **78**(2), 34–51.

Iacus, Stefano M., Gary King and Giuseppe Porro (2012), 'Causal inference without balance checking: Coarsened exact matching', *Political Analysis* **20**(1), 1–24.

Imai, Kosuke, In Song Kim and Erik H. Wang (2023), 'Matching methods for causal inference with time-series cross-sectional data', *American Journal of Political Science* **67**(3), 587–605.

Karlinsky-Shichor, Yael and Oded Netzer (2024), 'Automating the b2b salesperson pricing decisions: A human-machine hybrid approach', *Marketing Science* **43**(1), 138–157.

Kipf, Thomas N and Max Welling (2016), 'Semi-supervised classification with graph convolutional networks', *arXiv preprint arXiv:1609.02907* .

Knijnenburg, Bart P, Martijn C Willemsen, Zeno Gantner, Hakan Soncu and Cameron Newell (2012), 'Explaining the user experience of recommender systems', *User Modeling and User-Adapted Interaction* **22**(4-5), 441–504.

Kohavi, Ron, Diane Tang and Ya Xu (2020), *Trustworthy online controlled experiments: A practical guide to a/b testing*, Cambridge University Press.

Kohavi, Ron, Roger Longbotham, Dan Sommerfield and Randal M Henne (2009), 'Controlled experiments on the web: survey and practical guide', *Data mining and knowledge discovery* **18**(1), 140–181.

Kumar, Anuj and Kartik Hosanagar (2019), 'Measuring the value of recommendation links on product demand', *Information Systems Research* **30**(1), 129–145.

Lee, Dokyun and Kartik Hosanagar (2021), 'How do product attributes and reviews moderate the impact of recommender systems through purchase stages?', *Management Science* **67**(1), 524–546.

Liang, Chen, Zhan (Michael) Shi and T.S. Raghu (2019), 'The spillover of spotlight: Platform recommendation in the mobile app market', *Information Systems Research* **30**(4), 1296–1318.

Makioka, Ryo (2021), 'The impact of export promotion with matchmaking on exports and service outsourcing', *Review of International Economics* **29**(5), 1418–1450.

Meng, Wenjun, Lili Chen and Zhaomin Dong (2024), 'The development and application of a novel e-commerce recommendation system used in electric power b2b sector', *Frontiers in Big Data* **7**, 1374980.

Nia, A. G., Jie Lu, Qian Zhang and M. Ribeiro (2019), A framework for a large-scale b2b recommender system, *in* '2019 IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)', pp. 337–343.

Oestreicher-Singer, Gal and Arun Sundararajan (2012), 'The visible hand? demand effects of recommendation networks in electronic markets', *Management Science* **58**(11), 1963–1981.

Oster, Emily (2019), 'Unobservable selection and coefficient stability: Theory and evidence', *Journal of Business & Economic Statistics* **37**(2), 187–204.

Pilehvar, Ali, Wedad J Elmaghraby and Anandasivam Gopal (2017), 'Market information and bidder heterogeneity in secondary market online b2b auctions', *Management Science* **63**(5), 1493–1518.

Pu, Pearl, Li Chen and Rong Hu (2011), A user-centric evaluation framework for recommender systems, *in* 'Proceedings of the fifth ACM conference on Recommender systems', ACM, pp. 157–164.

Robins, James M, Miguel Angel Hernan and Babette Brumback (2000), 'Marginal structural models and causal inference in epidemiology'.

Rubin, Donald B. (1980), ‘Discussion of ”randomization analysis of experimental data: The fisher randomization test” by d. basu’, *Journal of the American Statistical Association* **75**(371), 591–593.

Sahni, Navdeep S, Dan Zou and Pradeep K Chintagunta (2017), ‘Do targeted discount offers serve as advertising? evidence from 70 field experiments’, *Management Science* **63**(8), 2688–2705.

Tucker, Catherine and Juanjuan Zhang (2010), ‘Growing two-sided networks by advertising the user base: A field experiment’, *Marketing Science* **29**(5), 805–814.

Tuli, Kapil R, Ajay K Kohli and Sundar G Bharadwaj (2007), ‘Rethinking customer solutions: From product bundles to relational processes’, *J. Mark.* **71**(3), 1–17.

Veličković, Petar, William Fedus, William L Hamilton, Pietro Liò, Yoshua Bengio and R Devon Hjelm (2018), ‘Deep graph infomax’, *arXiv preprint arXiv:1809.10341* .

Wu, Dianshuang, Guangquan Zhang and Jie Lu (2015), ‘A fuzzy preference tree-based recommender system for personalized business-to-business e-services’, *IEEE Transactions on Fuzzy Systems* **23**(1), 29–43.

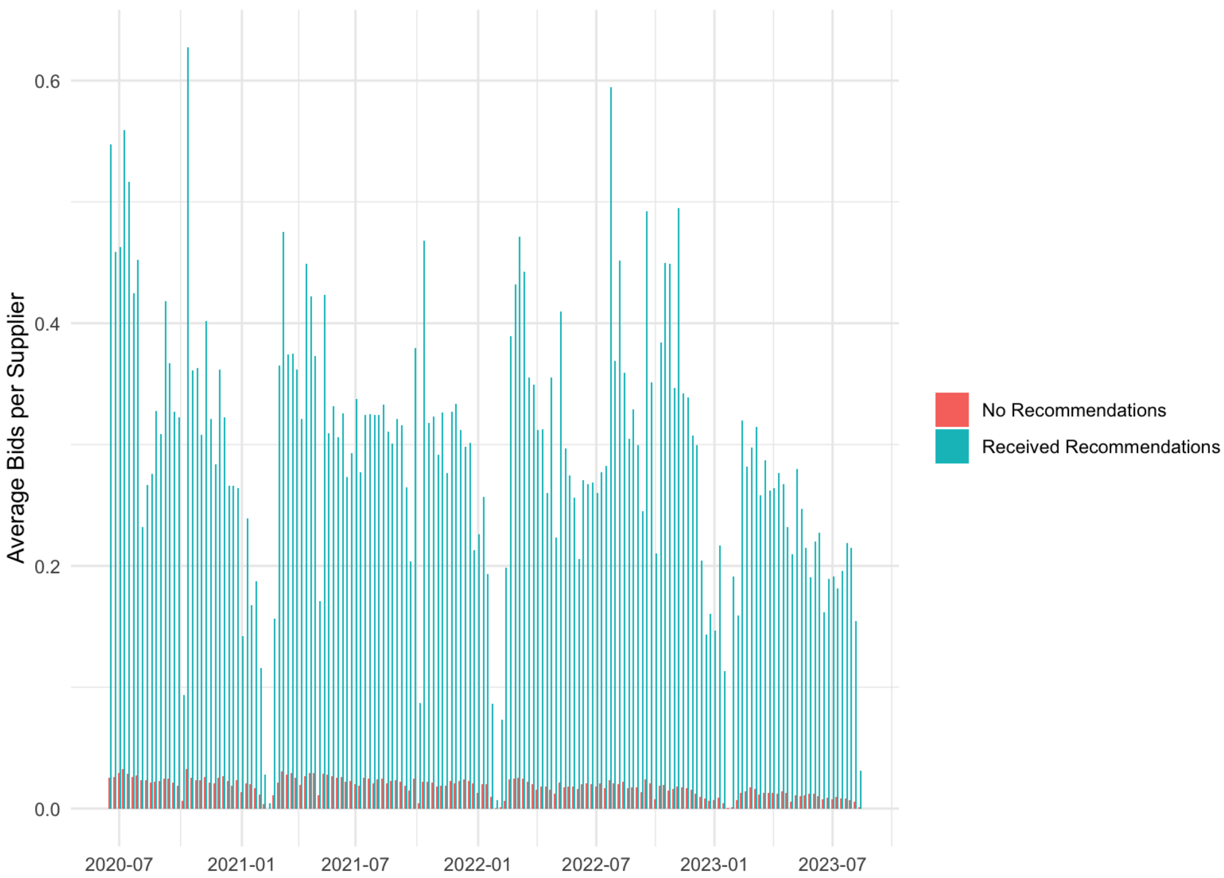
Yoon, Yeo Lim, Yeohong Yoon, Hyoryung Nam and Jeonghye Choi (2021), ‘Buyer–supplier matching in online b2b marketplace: An empirical study of small- and medium-sized enterprises (smes)’, *Industrial Marketing Management* **93**, 90–100.

Zhang, Jonathan Z., Oded Netzer and Asim Ansari (2014), ‘Dynamic targeted pricing in b2b relationships’, *Marketing Science* **33**(3), 317–337.

Zhang, Y. and J. Liu (2025), ‘B2b marketing recommendation system based on improved genetic algorithm model’, *Journal of Business Analytics* **8**(3), 135–146.
URL: *https://doi.org/10.1080/2573234X.2025.2458331*

Appendix A. Summary Statistics

Figure WA1 compares the weekly average number of bids per supplier between suppliers who received at least one recommendation in a given week and those who did not. Suppliers receiving recommendations tend to have higher bidding intensity relative to their counterparts. While this data pattern is consistent with a positive correlation between recommendation exposure and supplier bidding participation, it may also reflect nonrandom recommendation allocation. We address this identification challenge in the following sections using our matching procedure and regression analysis.



Note: This figure plots the weekly average number of bids per supplier. The blue line denotes suppliers who received at least one recommendation in the given week, and the red line denotes those who did not.

Figure WA1: Weekly Number of Bids at Supplier Level

Appendix B. Matching Procedure

Figure WA2: Illustration of Matching on Treatment History (Step I)

	i=1	i=2	i=3	i=4	i=5
t+1	0	1	1	0	0
t	1	0	1	0	0
t-1	0	0	0	0	0
t-2	1	1	0	1	0
t-3	0	0	1	0	1
t-4	1	1	0	1	0
t-5	0	1	0	0	1

Notes: Suppose supplier 1 receives a recommendation at time t and we regard it as a treated supplier. We first consider suppliers who do not receive recommendations at t (suppliers 2, 4, and 5). We then compare their treatment histories during the specified time window ($t-4$ to $t-1$ here) with that of the treated supplier. Only suppliers who share the same past treatment (suppliers 2 and 4) are retained to form the matched set M_{1t}.

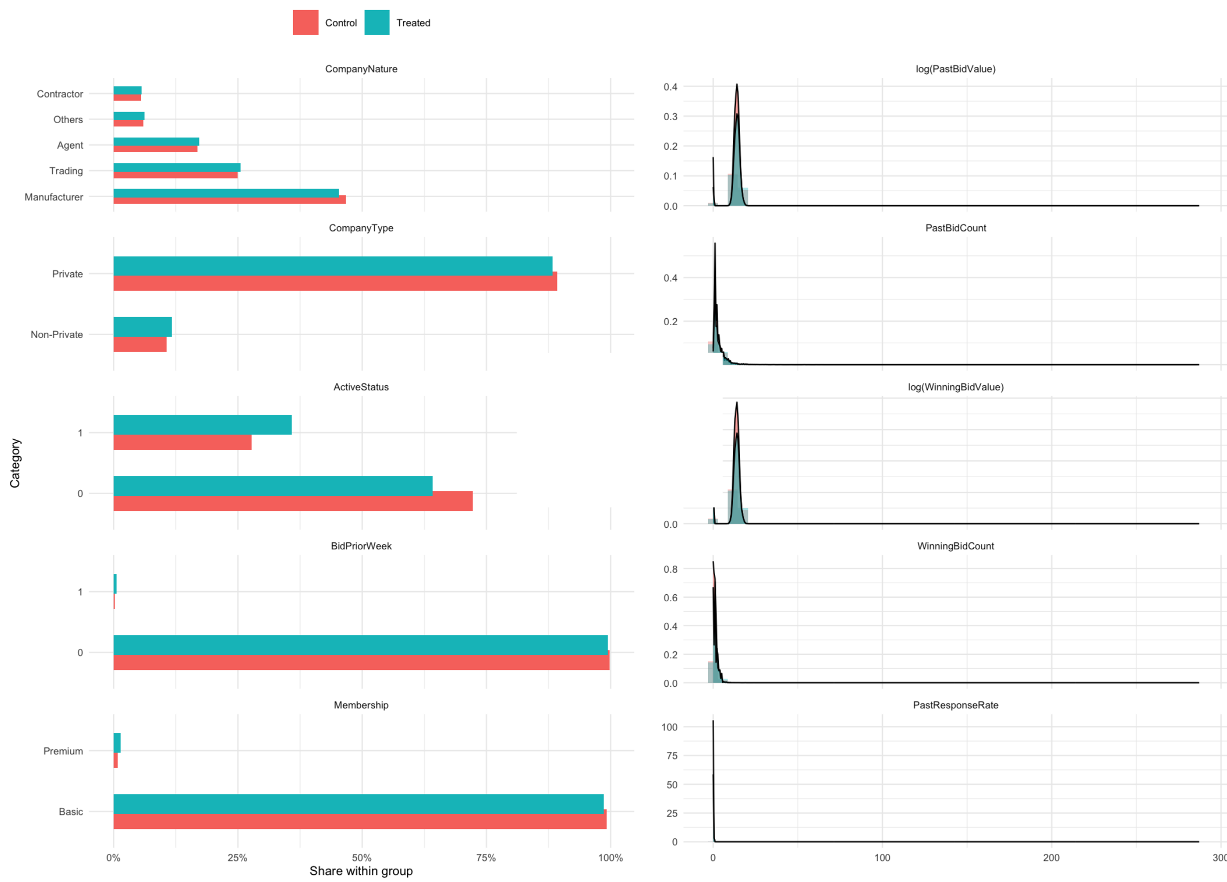
Appendix C. Summary Statistics of Covariates Before vs. After CEM

Table WA1: Summary Statistics of Covariates Before vs. After CEM

	log(Total Bid Value)		log(Total Winning Bid Value)		Total Bid Count		Total Winning Bid Count		Response Rate	
Statistic	Before	After	Before	After	Before	After	Before	After	Before	After
Min	0	0	0	0	0	0	0	0	0	0
1st quantile	0	0	0	0	0	0	0	0	0	0
Median	13.30	12.90	12.90	12.71	1	1	0	0	0	0
3rd quantile	15.07	14.29	14.74	14.13	4	2	1	1	0	0
Max	23.16	20.79	26.34	20.67	4003	287	1168	24	1	1

	Active Status		Bid at Prior Week		Membership			Company Type		
Level	Before	After	Before	After	Level	Before	After	Level	Before	After
0	.761	.876	.978	.999	Basic	.938	.989	Private	.831	.907
1	.239	.124	.022	.001	Premium	.062	.011	Non-private	.169	.093

	Company Nature	
Level	Before	After
Agent	.169	.155
Contractor	.082	.057
Manufacturer	.404	.466
Trading	.237	.245
Others	.108	.076



Note: This figure plots the distributions of key covariates for treated and control observations before and after matching. The left panel is the distributions for categorical covariates, and the right panel is those for continuous covariates.

Figure WA3: Covariate Distributions of Treated Group vs. Control Group

Appendix D. Deriving Supplier Embeddings

Categorical Embeddings

To quantify the similarity between supplier product portfolio, we develop a machine learning framework that captures latent characteristics based on the categories each supplier provides. Let $C(i)$ denote the set of categories offered by supplier i. We define the embedding of supplier i as the average of the latent embeddings of the categories in $C(i)$:

$$I_i = \frac{1}{|C(i)|} \sum_{c \in C(i)} h_c, \tag{8}$$

where h_c denotes the latent representation of category c. Thus, the core task is to obtain meaningful latent representations for each category.

To derive h, we leverage ChatGPT 4.0 to generate a comprehensive text description for every category. This step offers us valuable initial information because many category terms are niche and lack comprehensive documentation in either platform Z or other public sources. We employ natural language processing (NLP) methods to generate initial category embeddings based on the text descriptions. While these descriptions provide a starting point, they are often too generic. To tailor the embeddings to the platform's taxonomy, we leverage the platform's category catalog. Platform Z uses a hierarchical product–category–domain catalog. Individual products are grouped into categories based on shared functions or processing characteristics (e.g., Round Steel (Q235) and Wire Rod both belong to the Rod Products category). These categories are further aggregated into broader domains defined by material type (e.g., Rod Products fall under the Ferrous Metals domain).

We combine the text-based embeddings with the catalog structure by employing a Graph Convolutional Network (GCN), which allows us to refine category embeddings by aggregating information from neighboring nodes in the taxonomy graph (Kipf and Welling 2016). Specifically, we represent the category catalog as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with N nodes $v_i \in \mathcal{V}$, edges $(v_i, v_j) \in \mathcal{E}$, and adjacency matrix $A \in \mathbb{R}^{N \times N}$. Specifically, we define edges based on two types of relationships. First, we consider edges between categories with a direct child–parent relationship (e.g., there is an edge between "Round Steel (Q235)" and its parent category "Rod Products"). Second, we assign edges between sibling categories that share the same parent (e.g., there is an edge between "Round Steel (Q235)" and "Wire Rod," both of which fall under "Rod Products"). We implement a one-layer GCN with the propagation rule:

$$H = \delta(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}X\Theta), \tag{9}$$

where Θ is a learnable weight matrix, $\hat{A} = A + I_N$ adds self-loops to the adjacency matrix, and $\hat{D}$ is the corresponding degree matrix with $\hat{D}ii = \sum_j \hat{A}ij$. The matrix $X = \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N$ contains the initial text-based embeddings. We apply a ReLU activation function, δ, to introduce nonlinearity.

We train two versions of the GCN model. First, we conduct supervised learning by predicting the existence of an edge between two arbitrary categories in the category catalog. We use the trained GCN model to generate 512-dimensional embeddings for the 279 categories. The resulting embeddings are referred to as "surprised embeddings." Alternatively, we also train the GCN in an unsupervised way using the Deep Graph Infomax (DGI) algorithm (Veličković et al. 2018), which introduces a contrastive learning objective, where the model aims to distinguish between a node's neighborhood (i.e., positive examples) and other random nodes in the graph (i.e., negative examples). The derived embeddings are represented as "unsupervised embeddings." In our main analysis, we use the supervised embeddings, and the unsupervised embeddings serve as robustness checks.

Geographical Embeddings

To quantify the similarity between supplier service areas, we construct geographical embeddings based on their declared coverage regions. On Platform Z, suppliers report the specific regions they are able to serve, which may include multiple cities or provinces (e.g., Nantong City in Jiangsu Province, Shanghai, and Anhui Province).[20] To standardize these regions and reduce sparsity, we aggregate all location information to the province level.

We apply a one-hot encoding approach, where each of China's provinces is represented as a separate dimension. Each supplier's service coverage is encoded as a binary vector, with a value of one indicating availability in that province and zero otherwise. For suppliers that declare nationwide coverage, all entries in the vector are set to one.

The resulting binary vectors serve as the suppliers' geographical embeddings. To measure similarity between suppliers in terms of their geographic reach, we compute cosine similarity between their respective one-hot vectors.

Appendix E. Results Using Full Sample

Table WA2: Effects of Recommendations on Supplier Bidding.

	Direct	Spillover	Overall
Recommendation	0.015***	0.007***	0.020***
	(0.001)	(0.001)	(0.001)
Relevance control	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Num. obs.	122,626	122,626	122,626
R-sq.	0.260	0.255	0.258

Note: This table reports regression estimates of the effects of recommendations on supplier bidding outcomes. The "Overall" column reflects the combined effects from both direct and spillover channels. Standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

[20] In China, Shanghai is a municipality directly under the central government and is administratively equivalent to a province.

Table WA3: Effects of Recommendations on Successful Transactions.

	Direct Success	Spillover Success	Overall Success
Recommendation	0.004^{***}	0.001	0.005^{***}
	(0.000)	(0.001)	(0.001)
Relevance control	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Num. obs.	122,626	122,626	122,626
R-sq.	0.170	0.186	0.158

Note: This table reports regression estimates of the effects of recommendations on successful transaction outcomes. The "Overall Success," "Direct Success," and "Spillover Success" columns correspond to success in any RFQs, recommended RFQs, and non-recommended RFQs, respectively. Standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.