AI Copilots in Real Estate: Evidence from China

Lin William Cong　　Xiang Hui　　Yushan Zhou*

First draft: July 2025; this draft: March 2026

Abstract

We study the effect of an AI copilot for financial intermediation, using data from China's largest resale housing platform. Combining a neighborhood-level field experiment in one city and an exposure difference-in-differences design across 18 cities, we find that AI copilot access expedites real estate transactions: listings sell faster, seller time-on-market and buyer search duration fall, prices rise, and post-transaction customer ratings improve. Decomposing exposure by buyer- and seller-side app use shows gains on both sides with cross-side effects, while price effects load mainly on buyer-side exposure. Overall, the evidence is more consistent with reduced search and coordination frictions than with pure redistribution.

Keywords: Artificial Intelligence, Housing Market, Financial Intermediation, Two-Sided Markets, Real Estate Agents

*Cong: Nanyang Technological University, ABFER, CEPR, and NBER (will.cong@cornell.edu); Hui: Washington University in St. Louis (hui@wustl.edu); Zhou: CUHK Business School and DEFT Lab (yushanzhou@cuhk.edu.hk).

We thank Erik Brynjolfsson, Elizabeth Honka, and Meng Liu, as well as participants at the NBER Chinese Economy Working Group Meeting, University of Florida ISOM seminar, and UTD FORMS conference for their very helpful comments and discussions.

1 Introduction

Market designers are increasingly deploying AI decision-support tools to frontline intermediaries who execute transactions on their behalf: customer service representatives, loan officers, real estate agents, and information intermediaries. These tools are often introduced with an efficiency promise of lower search and communication costs, faster execution, and improved decision-making (Brynjolfsson et al., 2025; Bradshaw et al., 2026). However, in *two-sided, intermediated* markets, intermediary-facing AI tools can affect both the level and incidence of trade: it may reduce frictions and speed clearing but also change bargaining and generate cross-side effects that shift who captures value. Therefore, the consequences of deploying intermediary-facing AI tools in two-sided markets are a priori ambiguous.

This paper studies the introduction of an intermediary-facing AI copilot for human real estate agents (hereafter "agents") on China's largest resale housing platform in early 2019. We ask two questions. First, how does access to an intermediary-facing AI copilot affect market outcomes? Second, how do these effects distribute across the buyer side and seller side, and to what extent do improvements on one side affect outcomes on the other side?

In our setting, the copilot is embedded inside the agent's app workflow and is accessible only through the in-app agent–client chat interface. It provides task-specific support for common activities: for buyers, it recommends suitable properties and suggests reply scripts; for sellers, it suggests price adjustments and generates property titles and descriptions. The copilot's outputs are advisory and visible only to agents, so buyers and sellers are affected only through their interactions with agents. Importantly, the sample period surrounding the copilot's introduction provides a relatively uncontaminated empirical setting. Specifically, as the tool provided a discrete technological advance in decision support but preceded the proliferation of generative AI chatbots, there are reduced concerns that control agents or clients were simultaneously using similar tools.

Our empirical setting provides two complementary sources of identifying variation. First, we leverage a randomized field experiment conducted in Jinan (a second-tier city with a population of about 9 million), China, shortly after the platform-wide rollout of the AI copilot. Copilot access was randomly blocked for agents in half of the neighborhoods for roughly 11 months. The experiment provides a causal estimate of copilot access on overall market outcomes. Second, we turn to an 18-city observational sample and use a key institutional feature: because the copilot operates only inside the app channel, its effective reach should be higher in districts where a larger share of transactions were conducted via the app prior to rollout. We use cross-district variation in pre-period app transaction share as an exposure measure and estimate an exposure difference-in-differences design with neighborhood fixed effects and city-by-month fixed effects. The second design assesses whether the experimental patterns generalize beyond Jinan, and allows us to use side-specific exposure measures to study own- and cross-side effects.

To study the own- and cross-side effects, we further use a unique, institutional feature in our setting: app-based exposure is side-specific. Buyer-facing tools are relevant only when buyer–agent interactions occur in the app, while seller-facing tools are relevant only when seller–agent interactions occur in the app. We use pre-period district shares of buyers and sellers transacting via the app as proxies for buyer-side and seller-side exposure. Our empirical object is a 2×2 own- and cross-side effect matrix: at the margin, how buyer-side exposure shifts buyer outcomes and seller outcomes, and how seller-side exposure shifts seller outcomes and buyer outcomes.

We combine these designs with a proprietary dataset covering the 18 cities over the period from early 2018 to the end of 2019, including detailed micro-data on listings, price adjustments, on-site viewings, completed transactions, and post-transaction ratings, labels, and review text. The data allow us to measure speed outcomes (sale within three months and time-on-market for both buyers and sellers), transaction prices, and process-quality proxies on both sides.

Our main findings are threefold. First, copilot access improves overall speed and process efficiency. Specifically, higher exposure increases the probability that a listing sells within three months, and reduces time-on-market for both buyers and sellers. Although transaction prices rise, post-transaction ratings also increase on both sides. These effects are larger for less-experienced agents and in thinner markets, consistent with the copilot being most valuable where intermediation frictions are more binding. Second, the side-decomposition reveals systematic cross-side effects and different loadings across outcomes. Speed improvements are strongly two-sided (i.e., exposure on one side predicting faster outcomes on the other). In comparison, price effects load primarily on buyer-side exposure, rather than seller-side exposure. Third, mechanism evidence is consistent with reduced frictions in both workflows. Buyer-side exposure is associated with a more efficient search process (fewer on-site viewings and shorter buyer time outcomes), while seller-side exposure is associated with more frequent but smaller, negative price adjustments, consistent with faster convergence to market-clearing price.

Does this price increase reflect friction reduction or pure redistribution? Sellers should benefit from selling properties faster at a higher price. For buyers, a higher price could mean pure value extraction for a fixed matching quality, or increased overall value if copilot increases matching quality by allowing buyers to make better decisions and lower search costs. Note that even on platforms with digitized search, resale housing remains a high-friction market because of soft, or uncodifiable, information about property attributes, the difficulty to map high-dimensional preferences into a shortlist, and coordination between parties. These frictions create scope for intermediary skills to affect both efficiency and incidence, and therefore for intermediary-facing AI to change outcomes through channels other than bargaining.

While we do not estimate a full welfare decomposition, we provide a set of empirical facts that discipline interpretation of the price increase: (i) buyer search and time costs fall with exposure; (ii) buyer experience measures improves on dimensions targeted by the tool; (iii) higher prices coincide with faster market clearing; and (iv) seller-side exposure improves pricing workflow but does not drive the transaction price effect. Taken together, these patterns support an interpretation centered on reduced search and coordination frictions rather than a pure redistribution of value from buyers to sellers.

The remainder of the paper proceeds as follows. Section 2 describes the setting and introduces the own- and cross-side effect matrix. Section 3 describes the data and exposure construction. Section 4 reports the field experiment. Section 5 presents the multi-city exposure DiD results, including the decomposition. Section 6 discusses mechanisms and interpretation. Section 7 concludes.

1.1 Related Literature

Our paper is related to several strands of literature on financial intermediation, algorithmic information in asset markets, and the economics of artificial intelligence. Across these strands, our main contribution is to examine how an *intermediary-facing* AI decision-support tool shapes information production, investor decision-making, and transaction outcomes in a *two-sided housing market*, and to explicitly characterize both own-side and cross-side effects.

Intermediaries and information frictions in financial markets. A large volume of prior literature studies the role of intermediaries in mitigating information frictions in asset markets (Glode and Opp, 2016; Allen et al., 2025; Bourveau et al., 2022; Agarwal et al., 2019a). In housing markets, real estate brokers play an important role in facilitating transactions, providing market information, and matching buyers and sellers, but may also face incentive conflicts or agency problems (Agarwal et al., 2019b; Levitt and Syverson, 2008). Related work mainly examines how intermediaries influence market outcomes through information provision, negotiation, and matching frictions (Aiello et al., 2026; Braggion et al., 2025; Agarwal et al., 2019b; Gilbukh and Goldsmith-Pinkham, 2024). While this literature has traditionally focused on human intermediaries, our study examines how AI-assisted decision support changes both intermediaries' information production and transaction outcomes.

Algorithmic information and valuation in housing markets. Previous studies on housing markets show that consumer-facing algorithmic valuation tools, such as Zillow's Zestimate, influence listing and selling behavior and can generate feedback loops between algorithmic recommendations and human decisions (Fu et al., 2022; Malik and Manzoor, 2023). Related research also highlights distributional consequences and disparities arising from algorithmic valuation tools (Yu, 2020; Fu et al., 2025). More broadly, algorithmic pricing, recommendation, and ranking systems are studied in platform-mediated markets including ridesharing (Hall et al., 2015; Liu et al., 2023; Castillo, 2023), online rental markets (Zhang et al., 2021b; Pan and Wang, 2021; Huang, 2024), and e-commerce platforms (Zhang et al., 2021a; Fong et al., 2023; Shi et al., 2024). Additionally, existing work studies how algorithmic signals affect decision-making in asset markets (Weller, 2018; Erel et al., 2021). However, much of this literature examines algorithms that directly influence a single market side (for example, valuation signals for sellers or recommendations for consumers). In contrast, our setting studies an AI tool deployed to intermediaries who simultaneously interact with both sides of the market, thereby allowing us to examine how algorithmic information propagates through intermediaries to affect both buyers and sellers.

Algorithmic decision-making and financial technology. Previous studies examine how machine learning and algorithmic tools influence decisions in domains such as credit scoring, underwriting, and lending markets, with implications for efficiency, discrimination, and financial inclusion (Fuster et al., 2022; Cao et al., 2024; Bartlett et al., 2022; Fuster et al., 2019). More broadly, research on financial technology adoption studies how digital technologies reshape financial intermediation and market structure (Hau et al., 2024; Goldstein et al., 2019; Agarwal et al., 2024). Our setting focuses on a new financial technology tool—an AI copilot that assists human intermediaries rather than replacing them—allowing us to study how AI augments intermediary decision-making and how such augmentation affects market outcomes in a two-sided environment.

Economics of AI and information production. Finally, our study is related to the literature documenting how AI technologies affect productivity, decision-making, and information production across tasks and industries. Prior work documents the effects of AI systems on writing (Noy and Zhang, 2023; Wiles et al., 2025; Hui et al., 2024; Demirci et al., 2025), translation (Brynjolfsson et al., 2019; Zhu and Walker, 2024), coding (Cowgill et al., 2020; Peng et al., 2023), knowledge platforms (Burtch et al., 2024; Shan and Qiu, 2025), and professional services (Dell'Acqua et al., 2023; Zhou and Lee, 2024; Peukert et al., 2024; Goldberg and Lam, 2025). Related evidence highlights the growing use of AI in high-stakes decision environments such as risk assessment (Cong et al., 2024; Bertomeu et al., 2025), customer engagement (Luo et al., 2019; Schanke et al., 2021; Brynjolfsson et al., 2025), and healthcare (Agrawal et al., 2024). At the firm level, recent studies also document the effects of AI adoption on firm value and growth (Babina et al., 2024; Eisfeldt et al., 2023; Cao et al., 2023). While this literature primarily studies productivity and task- or firm-level outcomes, it typically does not examine how AI shocks propagate through intermediaries to influence outcomes across multiple sides of a market. Our study examines how AI-assisted intermediaries affect information transmission and transaction outcomes in a two-sided asset market.

2 Setting

2.1 China's Real Estate Market

Scale: The Chinese real estate market is among the largest globally, with the value of total commercial housing sales reaching 17.36 trillion RMB in 2020.[1] Housing is the primary form of household wealth in China, with an urban homeownership rate of 96% and residential housing assets accounting for about 69% of total household assets.[2]

Brokerage organization: Large brokerage firms in China often employ frontline real-estate agents (salary plus commissions) and frequently intermediate both sides of a deal (dual agency). This organizational form matters for our context because it gives the firm and platform leverage to standardize workflows, training, and digital tools across agents.[3]

[1]Source: https://www.stats.gov.cn/sj/zxfb/202302/t20230203_1900967.html

[2]Source: https://m.163.com/dy/article/FB0AF3P20539ADBX.html; https://stock.hexun.com/2023-08-24/209822345.html

[3]A closely related organizational form exists in the U.S. in platform-brokerage models with employed agents (e.g., Redfin's W-2 agent model), though the dominant U.S. model remains independent contractor agents.

The platform: Starting in the late 2010s, leading firms built integrated online–offline platforms that combine consumer search, agent–client communication, and brokerage execution. In this environment, the relevant economic distinction is not online listings vs. offline search but rather information vs. execution: even when buyers can browse listings digitally, completing a transaction still requires screening, coordination, and trust-building between parties.

Value of intermediation: Digitization lowers the cost of accessing listings, but it does not eliminate three major frictions. First, soft information from local knowledge remains important. Information on micro-location disamenities, school-zone boundaries and local rules, and building/management quality is difficult to codify credibly in listings and costly for buyers to learn through independent search. Second, the matching problem is high-dimensional: even with full listings, mapping a buyer's preferences and constraints into a short, relevant consideration set is costly, making "recommendation quality" a natural margin for reducing wasted search efforts. Third, transactions involve coordination and trust: repeated communication, scheduling, and documentation exchanging create scope for breakdowns, and improving the reliability and coherence of agent-client interaction can speed transactions on both sides. These frictions create scope for both human intermediation and an agent-facing copilot to raise efficiency through better screening, smoother communication, and faster convergence in listing strategy.

2.2 Empirical Context

We study one of the largest nationwide platforms. By 2019 it operated thousands of offline stores across more than 100 cities and employed on the order of 100,000 agents, facilitating millions of annual transactions. Sellers and small agencies list on this platform to access its verified, up-to-date inventory and buyer traffic. Buyers use it for trusted, up-to-date listings and professional intermediation. As the user base grows, the platform gains access to better data, leading to better matches, which in turn attract even more users, creating a reinforcing loop that increases the on-platform share of transactions.

Transactions happen through both offline (walk-ins to a store, phone) and online (app/web) entry points. The firm defines a transaction to be offline if the entire process remains offline end-to-end (about 40% transactions in 2018); otherwise, a transaction is defined as online. The app is central to the workflow when clients engage via the app, agents message, recommend homes, manage listings, and schedule viewings there. The AI copilot is available *only within the app's agent-client chat.*

There are clear service boundaries on this platform. Agents serve clients within a local neighborhood, which is a platform-defined commercial or residential area. Districts are strictly nested within cities, but neighborhoods can straddle district borders. Company policy restricts agents from operating outside their local areas. If a client wishes to view a property in another district, the case is handed off to an agent responsible for that district. This limits cross-district spillovers, which we detail in Section 5.2.

The typical transaction process on the platform is summarized in Appendix Figure A1. A buyer initiates communication in-person or via the app. The agent elicits preferences, recommends listings, schedules on-site viewings, and facilitates negotiation. A seller lists a property through an agent in-person or via the app. The agent helps set an initial price, writes titles and descriptions for the ad, monitors interest, and recommends price adjustments. If a deal is reached, the platform records the transaction date and price, and both sides review the agent's service.

2.3 Policy Change: Introduction of an AI Copilot for Agents

On January 16, 2019, the firm introduced an app-exclusive AI copilot to agents in our sample cities, which became popular quickly.[4] The AI copilot gives recommendations targeting four main agent challenges. For buyers, it recommends suitable properties and reply scripts. For sellers, it suggests price adjustments and property titles and descriptions. Importantly, the copilot is accessible only inside the in-app agent–client chat, and tool availability is side-

[4]By the end of 2019, agents followed approximately 60% of the AI copilot's recommendations. While the firm provided these post-rollout adoption rates at the district level, the data cannot be linked to individual transactions.

specific. For example, if the seller uses the app but the buyer does not, the agent can use only seller-side features and not buyer-side features.

To better illustrate the treatment, we map the AI copilot onto the baseline transaction workflow (illustrated in Appendix Figure A1). The rollout did not automate transactions or replace agent discretion end-to-end. Instead, it inserted side-specific suggestions into pre-existing tasks handled manually by agents: on the buyer side, property recommendation and reply-script assistance in the agent–buyer chat; on the seller side, price-adjustment advice and title/description drafting in the agent–seller chat. The technological change is that these workflow steps now include platform-generated suggestions powered by transformer and deep-learning models. While distinct from the open-ended generative large language models prevalent today, these models represented a major leap in targeted decision-support capability at the time.

In all four cases, the output is advisory rather than automatic, so agents can adopt, revise, or ignore the suggestion. Buyers and sellers do not generally observe the raw model output; they are affected through the agent's adopted actions, messages, and listing changes, with the exception that listing text becomes public when adopted. Table 1 summarizes how each module maps into the workflow and clarifies how our buyer- and seller-side exposure measures relate to these functions.

Figure A1 illustrates how the AI copilot integrates into transaction processes. For the buyer-facing workflow, the copilot recommends to agents suitable listings based on conversation text and buyers' browsing histories (Figure 1a). In addition, the AI copilot suggests professional reply scripts to improve communication (Figure 1b). For the seller-facing workflow, the copilot suggests dynamic price adjustments using comparable listings and market information (Figure 1c), and proposes optimized titles and descriptions to highlight key selling points (Figure 1d).

2.4 Empirical Objects: Overall Effects, Own- and Cross-Side Effects

Because the copilot is accessible only inside the in-app agent–client chat and tool availability is side-specific, both effective market-level exposure and buyer- and seller-side exposure to the technology vary across locations with different baseline app usage. We exploit this institutional feature to study two empirical objects: (i) the overall effect of increasing copilot reach on market outcomes and (ii) the own- and cross-side effects for the buyer side and seller side.

We conceptually decompose exposure into seller-side exposure, E^S, capturing the intensity of copilot availability in seller–agent interactions, and buyer-side exposure, E^B, capturing the intensity of copilot availability in buyer–agent interactions. Let Y^S denote seller-side outcomes (e.g., selling speed and price) and Y^B denote buyer-side outcomes (e.g., search and post-transaction experience). Our main own- and cross-side object is the 2×2 response matrix

$$\mathcal{M} = \begin{pmatrix} \partial Y^S / \partial E^S & \partial Y^S / \partial E^B \\ \partial Y^B / \partial E^S & \partial Y^B / \partial E^B \end{pmatrix}.$$

The diagonal elements summarize own-side responses, while the off-diagonal elements capture cross-side effects (i.e., how changes in exposure on one side affect outcomes on the other side). We view $\mathcal{M}$ as a reduced-form, marginal equilibrium response to access to the technology.

Our empirical analysis combines two complementary designs. In Section 4 we use a neighborhood-level field experiment from one city to provide clean evidence on the causal effect of having access to the copilot. In Section 5 we then use cross-district variation in buyer- and seller-side exposure to characterize own- and cross-side effects using data from our sample of 18 cities.

3 Data

We use proprietary data from a large real estate platform. The AI copilot was rolled out on the platform across our sample on January 16, 2019, which falls near the midpoint of our main observation window. The data link listings, within-listing weekly activity, transaction outcomes, and post-transaction evaluations. We analyze two complementary samples. The first is a neighborhood-level field experiment conducted in Jinan, a second-tier city in China, in 2019. The second is a multi-city rollout sample covering 18 cities (not including Jinan) from late February 2018 to the end of December 2019.

Geographic units: We use three geographic levels: city, district, and neighborhood. These three levels are roughly analogous to state, metropolitan statistical area (MSA), and county in the U.S., respectively. More precisely, cities are prefecture-level units or direct-administered municipalities. Districts are county-level units nested within cities; this is the level at which baseline app penetration varies and therefore the level at which exposure measures are defined in the rollout design. The finest location unit is the platform-defined neighborhood, which we use for location fixed effects. Neighborhoods are smaller than districts but can occasionally straddle district borders.

Listings and weekly activity panels (process outcomes): For all active listings in our sample, we observe a weekly panel of listing activity, including on-site buyer viewings and seller price adjustments, along with time-invariant property characteristics (e.g., size, rooms, floor, age, decoration). At the listing level, our main conversion outcome is an indicator for whether the listing sells within three months of publication. In our sample, more than 95% of listings sell within 12 months after publication, and about 54% of them sell within three months.

Transaction outcomes (market outcomes): For listings that sell, we observe transaction outcomes including closing date, transaction price, last listed price, time-on-market for sellers, and time-on-market for buyers (from first contact to closing). We also observe agent attributes (e.g., gender, age, tenure, and performance scores) and property characteristics. The transaction sample contains approximately 230,000 closings (dropping around 20 observations with missing core fields).

Post-transaction feedback: For roughly one-third of transactions, we also observe post-transaction feedback from buyers and sellers, including an overall rating on a 5-point scale, review text, and structured labels evaluating specific dimensions of the agent's service (e.g., "high-quality recommendation," "communication inefficiency," "trustworthy," "reasonable price adjustment recommendations," "accurately capturing selling points"). These post-transaction feedback entries include 75,596 buyer evaluations and 20,336 seller evaluations. Buyer evaluation rates are higher because the platform actively reminded buyers to submit feedback during the sample period.

Pre-period exposure measures: The firm provides three district-level measures of pre-rollout app usage, which we treat as predetermined exposure proxies (used in Section 5.1):

- **Aggregate exposure** E_d^{Total} (empirically, Ratio_tran_d): the pre-period share of transactions in district d conducted through the app channel.
- **Buyer-side exposure** E_d^B (empirically, Ratio_buyer_d): the pre-period share of buyers in district d who transact via the app.
- **Seller-side exposure** E_d^S (empirically, Ratio_seller_d): the pre-period share of sellers in district d who transact via the app.

We use Ratio_tran_d to estimate overall effects of copilot access (Section 5.3), and Ratio_buyer_d, Ratio_seller_d to estimate own- and cross-side effects via a 2×2 exposure-response decomposition (Section 5.4). Both Ratio_buyer_d and Ratio_seller_d can show independent variation across districts because of idiosyncratic local management practices that historically prioritized either buyer app onboarding for digital search or seller app onboarding for listing management. We do not observe function-by-function copilot usage; these pre-period shares proxy differential exposure because the copilot's buyer- (resp., seller-) facing features are embedded in the agent–buyer (resp., agent–seller) in-app chat flows.

District covariates: We complement platform data with district-level time-varying covariates from the firm's internal sources (i.e., population, income, GDP, unemployment), which we use as controls.

Summary statistics: Table 2 reports summary statistics for the main variables. Unless otherwise noted, we use $\log(y + 1)$ transformation for buyer and seller time outcomes, and logs for transaction prices. The mean of the transformed seller time-on-market is 3.99, or 53 days. For buyers, the transformed mean is 2.66, or 13 days. The mean log transaction price is 10.27, which translates to an average transaction price of 28,800 RMB (about 4,122 USD) per square meter. At the transaction level, the mean overall exposure is 0.446. Notably, this overall measure, along with the side-specific buyer and seller exposure measures, shows substantial variation, providing the necessary identifying variation for our estimates.

4 The Causal Effect: Evidence from a Field Experiment

We begin with a randomized field experiment that estimates the causal effect of copilot access on market outcomes in a single city.

The partner firm ran the experiment in Jinan for 11 months starting in February 2019, roughly one month after the copilot was rolled out platform-wide.[5] The unit of randomization is the neighborhood. The firm randomized 68 neighborhoods, assigning 34 to a group in which access was maintained and 34 to a group in which access was blocked. We refer to the former as the *treatment* group (copilot enabled) and to the latter as the *control* group (copilot disabled), even though the platform-wide status quo at the time was access. In control neighborhoods, the feature was disabled in the agent app, and the agents were informed that the functionality was temporarily under maintenance.

Appendix Table A2 reports randomization checks and shows no meaningful pre-treatment differences across conditions. We estimate intent-to-treat (ITT) effects at the neighborhood level, with standard errors clustered by neighborhood. We focus on four outcome families: the probability that a listing sells within three months, seller and buyer time-on-market outcomes, transaction prices, and post-transaction ratings from buyers and sellers.

Table 3 reports the ITT estimates. In control neighborhoods, about 52.7% of listings sold

[5] The firm conducted this small-scale field experiment to better evaluate the effectiveness of this tool and to inform future improvements and deployment decisions.

within three months. Copilot access increased this probability by about 11 percentage points (roughly a 21% rise relative to the control mean). For sold listings, copilot access reduced seller and buyer time outcomes by 0.37 and 0.45 log points, respectively, implying declines of about 31% and 36%, respectively. Transaction prices rose by roughly 4.8% relative to the control group. Ratings from both buyers and sellers also increased.

One concern is that disabling the tool for the control group might lead to agent frustration. If so, the randomized control trial estimates could capture both the technological benefit of the technology and a psychological cost. To ensure our findings reflect the effects of the technology's introduction rather than those of its removal, we complement this experiment with a multi-city exposure difference-in-differences design that evaluates the copilot rollout (Section 5).

5 Decomposing the Effects: Evidence from Multi-City Rollout

We next use an exposure difference-in-differences (DiD) specification in 18 other cities to assess external validity across these cities and decompose the overall effect into own- and cross-side effects by separately leveraging buyer-side and seller-side copilot access.

5.1 Empirical Strategy

As the copilot is embedded in the in-app agent–client chat interface, its reach across geography depends on whether a transaction is conducted through the app rather than through offline channels. We use cross-district variation in *pre-rollout* app usage as a proxy for treatment exposure. Specifically, districts with higher baseline app shares have more agent–client interactions occurring in the channel where the copilot operates, and therefore experience a larger change in exposure after the rollout.

Similar to Yang (2025), we begin with an exposure difference-in-differences specification that uses a district-level pre-period app-transaction share, E_d^{Total} (empirically, Ratio_tran_d), as a proxy for the aggregate reach of the copilot:

$$y_{indct} = \beta \text{Post}_t \times \text{Ratio\_tran}_d + X_{idt}\delta + \xi_n + \tau_{ct} + \varepsilon_{indct}, \tag{1}$$

where y_{indct} represents whether the listed property is sold within three months, the time on market for sellers, time on market for buyers, and housing transaction prices for transaction i in neighborhood n district d city c in year-month t; Post_t indicates whether the transaction occurred after the deployment of the AI copilot; and Ratio_tran_d indicates the proportion of housing transactions at the district level that were conducted through the app during the pre-treatment period. We include neighborhood fixed effects ξ_n and city-by-month fixed effects τ_{ct}, and control for property characteristics and time-varying district covariates in X_{idt}: housing characteristics (such as building size or number of rooms) and four key district-level variables: income, unemployment, GDP, and population. We also include the corresponding main effects in the implementation; however, with neighborhood fixed effects and city-by-month fixed effects, they are absorbed (or identified only off limited within-month timing) and are not interpreted. Standard errors are clustered at the district level.

To study own- and cross-side effects, we replace $E_{\text{Total},d}$ with two pre-period exposure proxies: buyer-side exposure E_d^B (empirically, Ratio_buyer_d) and seller-side exposure E_d^S (empirically, Ratio_seller_d):

$$y_{indct} = \beta_1 \text{Post}_t \times \text{Ratio\_buyer}_d + \beta_2 \text{Post}_t \times \text{Ratio\_seller}_d + X_{idt}\delta + \xi_n + \tau_{ct} + \epsilon_{indct}. \tag{2}$$

The pre-period shares Ratio_buyer_d and Ratio_seller_d represent the proportion of buyers and sellers using the app in the pre-period, respectively. They proxy different exposure to the copilot's buyer-facing and seller-facing workflow, because the copilot's buyer- (resp., seller-) facing features live in the agent–buyer (resp., agent–seller) in-app chat flows. We also include the corresponding main effects in the implementation; however, they are mostly absorbed by the fixed effects and are not interpreted. Standard errors are clustered at the district level.

5.2 Identification Assumptions and Checks

Our exposure DiD design uses pre-period app shares as a proxy for differential exposure to the copilot after rollout. The key identifying assumption is an exposure-weighted parallel trends condition: absent the copilot, outcomes in districts with higher and lower pre-period app usage would have evolved similarly after conditioning on neighborhood fixed effects, city-by-month fixed effects, and the covariates. We also need that other contemporaneous platform changes are absorbed by city-by-month fixed effects and are not differentially correlated with pre-period app shares, and that cross-district contamination is limited.

Exposure-weighted parallel trends: We assess the parallel-trends assumption using an event-study version of Equation 1:

$$\mathrm{y}_{indct} = \sum_{k \neq -1} \beta_k \text{ Period }_k \times \text{Ratio\_tran}_d + X_{idt}\delta \ + \xi_n + \tau_{ct} + \epsilon_{indct}, \tag{3}$$

where Period_k indexes event time relative to rollout (with $k = -1$ omitted). Under parallel trends, the coefficients β_k should be statistically indistinguishable from zero in pre-period months ($k < 0$). The event-study plots in Figure 2 show flat pre-trends and an immediate post-rollout shift in outcomes.

We also estimate the analogous event-study version of Equation 2. The plots in Figure 3 show relatively flat pre-trends and an immediate post-rollout shift in outcomes for both buyer- and seller-side exposure interactions.

Exclusion and concurrent changes: A concern is that pre-period app penetration may proxy for latent district traits that correlate with differential responsiveness to the technology. An interview with an executive of the firm confirmed that app usage varied significantly (from 25% to 70%), driven primarily by younger demographics, lower-priced properties where clients are relatively less cautious, and areas with sparse offline stores. To ensure these latent selection traits do not drive our results through different treatment elasticity, we have that our baseline specification absorbs time-invariant neighborhood factors with ξ_n, city-specific time shocks with τ_{ct}, and observed property and district covariates in X_{idt}.

Next, we show our main treatment effect estimates remain stable when explicitly allowing for differential post-rollout effects by interacting $Post_t$ with key property and district controls (e.g., building age, number of rooms, income, GDP etc.; Appendix Tables A3). This finding suggests that groups with different baseline traits are not simply reacting differently to the post-period.

Additionally, we replace the pre-exposure proxies with district-level copilot usage intensity measured in the month immediately following rollout. Because post-rollout usage is potentially endogenous, we interpret this exercise as supportive rather than causal. Nevertheless, the estimates in Appendix Table A4 of Post$\times$Usage are similar (if anything larger), and the buyer/seller decomposition continues to show strong cross-side effects, suggesting that our baseline exposure measures capture differential copilot reach rather than unrelated latent district trends.

Furthermore, the lack of a sharp change in the exposure measure after the rollout (Appendix Figure A2) suggests that the effect is not entirely driven by increased app usage.

Finally, we show robustness to a range of placebo tests and test additional outcomes (Appendix Table A6). The placebo checks in columns 1–5 provide evidence against coincident platform changes that are correlated with baseline app shares, or confounding from compositional shifts. Column 7 shows no statistically significant exposure-weighted change in log transaction volume, suggesting that the main results are not driven by an exposure-correlated expansion in market activity (e.g., marketing or digitization shocks) but instead primarily reflect faster clearing and process changes conditional on participating.

Spillovers and interference: Because the copilot operates inside in-app chat flows, a threat is interference: treated agents could indirectly benefit offline clients or affect adjacent districts. Note that if such a spillover exists, to the extent that a low-exposure district benefits from a neighboring high-exposure district, the spillover would attenuate our treatment effect estimates, making them conservative. Institutionally, the tool is not standalone and is tied to specific in-app conversations, which limits direct use for offline clients; moreover, platform policy restricts agents from instructing clients to switch channels during our sample period. Empirically, we probe geographic interference by excluding neighborhoods that straddle district borders and find the results are stable (Appendix Table A7). Finally, we check that the estimated *cross-side* effects are not mechanically driven by correlation between buyer- and seller-side exposure: when we exclude districts in the top quartile of buyer exposure, seller-side exposure continues to predict faster buyer outcomes. Similarly, when we exclude districts in the top quartile of seller exposure, buyer-side exposure continues to predict faster seller outcomes (Table A8).

Treatment heterogeneity and functional form: Our main specifications treat exposure as continuous. However, Callaway et al. (2024) show that standard linear estimators can mask heterogeneous treatment effects across different exposure intensities by forcing a constant, linear dose-response relationship, and can produce even the wrong sign. To relax these functional-form assumptions and to mitigate potential measurement error in the exposure proxy, we also report qualitatively similar results when we use binned (indicator) versions of the exposure measure (Appendix Table A9).

5.3 Overall Effect

We first estimate overall effects of copilot access on market outcomes using Equation 1 in the 18-city rollout sample. This multi-city evidence serves two purposes: it assesses whether the effects observed in the single-city experiment generalize across locations, and it provides the variation needed to quantify own- and cross-side effects, as discussed in the next subsection.

We begin with checking the first stage. We regress district-level post-rollout copilot usage on our exposure measure, $Ratio\_tran_d$, and a constant. The estimated slope is 0.84 and highly significant, consistent with actual tool usage scaling well with pre-period in-app chat usage.

Panel A of Table 4 reports estimates of Equation 1. The interaction coefficient β_3 measures how the post-rollout change in outcomes varies with baseline app penetration. To illustrate magnitudes, we scale effects to a 10 percentage-point (pp) increase in the pre-period app transaction share (mean share is 0.446). For sale within three months (column 1), $\beta_3 = 0.4215$ implies that a 10 pp higher baseline app share is associated with a $0.4215 \times 0.10 \approx 0.042$ (4.2 pp) larger post-rollout increase in the probability of sale, which is about 6.3% relative to the pre-period mean of 0.668. Columns 2–3 show that for seller and buyer time on market, the corresponding effects are $-0.4434 \times 0.10 \approx -0.044$ and $-0.2484 \times 0.10 \approx -0.025$ in log points (or -4.3% and -2.5%), respectively.

Transaction prices increase by $0.10 \times 0.10 = 0.010$ log points, i.e., about a 1% increase (column 4). At the sample mean unit price of 28,800 RMB per square meter and mean size of 84.28 m^2, this increase corresponds to $(e^{0.010} - 1) \times 28,800 \approx 289$ RMB/m^2, or about $(e^{0.010} - 1) \times 28,800 \times 84.28 \approx 24,357$ RMB (about 3,400 USD) per transaction.

Although prices rise, both buyer and seller post-transaction ratings increase (columns 5–6). A 10 pp increase in baseline app share raises buyer ratings by $0.2206 \times 0.10 \approx 0.022$ and seller ratings by $0.3474 \times 0.10 \approx 0.035$ rating points.

Overall, the direction of these effects (i.e., faster transactions, higher sale probability, and higher prices) is qualitatively consistent with the randomized evidence in Section 4. This alignment is consistent with the validity of the exposure DiD approach, and suggests that the rollout estimates capture economically meaningful consequences of copilot access that generalize beyond a single market.

We also ask whether the overall effects of copilot access are larger where matching and execution frictions are plausibly more binding. Consistent with a friction-reduction interpretation, the estimated effects are more pronounced for transactions with less-experienced agents, lower-performing agents, and in thinner markets (Tables 5 and Appendix Tables A11 and A12). This finding suggests the copilot serves a "leveling" function, disproportionately benefiting agents who face the highest baseline intermediation frictions. We view these heterogeneity patterns as suggestive evidence for reduced friction.

We conduct a few robustness checks. Event-study estimates show flat pre-trends and immediate post-rollout shifts in outcomes (Figure 2), suggesting a causal effect from copilot access. We find little evidence of compositional changes in property characteristics (Appendix Table A5) and get null effects in placebo outcomes the copilot should not affect (Appendix Table A6). The main estimates are also robust to interacting district covariates with *Post*, binned versions of the exposure proxy, and alternative clustering schemes (Appendix Tables A3, A10, and A9).

5.4 Own- and Cross-Side Effects

We next connect the rollout estimates to the own- and cross-side effect matrix in Section 2.4. Recall that the conceptual object is the 2×2 response matrix linking buyer- and seller-side outcomes to buyer-side and seller-side exposure. To operationalize buyer-side and seller-side exposure, we replace the single aggregate proxy Ratio_tran_d with two pre-period exposure proxies: the district share of buyers transacting via the app ($E_d^B \equiv \text{Ratio\_buyer}_d$) and the district share of sellers transacting via the app ($E_d^S \equiv \text{Ratio\_seller}_d$). Estimating Equation 2, we find the coefficients on $\text{Post} \times \text{Ratio\_buyer}_d$ and $\text{Post} \times \text{Ratio\_seller}_d$ correspond to reduced-form estimates of the post-rollout responses $\partial Y/\partial E^B$ and $\partial Y/\partial E^S$. Comparing these coefficients across buyer- and seller-side outcomes therefore maps directly to the own- and cross-side cells of the matrix.

Panel B of Table 4 reports the decomposed estimates. There are three key findings.

First, efficiency gains are two-sided, with cross-side effects. Both $\text{Post} \times \text{Ratio\_buyer}_d$ and $\text{Post} \times \text{Ratio\_seller}_d$ increase the probability that a listing sells within three months and reduce time outcomes for both buyers and sellers (columns 1–3). Scaling to a 10 pp increase in exposure, we find that buyer-side exposure raises the likelihood of a sale within three months by about 1.7 percentage points and reduces seller and buyer time by about 2.3% and 1.4%, respectively. Seller-side exposure raises a sale within three months by about 4.1 percentage points and reduces seller and buyer time by about 4.8% and 2.5%, respectively. In the matrix, both off-diagonal cells are economically meaningful: buyer-side exposure improves seller-side speed outcomes, and seller-side exposure improves buyer-side time outcomes, consistent with cross-side effects in a high-friction market.

Second, price effects load primarily on buyer-side exposure. In column 4, the price increase is concentrated in Post $\times$ Ratio_buyer_d: a 10 pp increase in buyer-side exposure implies about a 1.2% increase in prices. In contrast, the coefficient on Post $\times$ Ratio_seller_d is small and statistically indistinguishable from zero. In matrix terms, the response of prices loads on $\partial P/\partial E^B$ rather than $\partial P/\partial E^S$. One interpretation is that buyer-side workflows mainly affect match quality and willingness-to-pay.

Third, satisfaction rises on both sides, including cross-side. Columns 5–6 show that buyer-side exposure increases seller ratings, and seller-side exposure increases buyer ratings, alongside own-side gains. For example, a 10 pp increase in buyer exposure raises buyer (resp., seller) ratings by about 0.013 (resp., 0.029) points, while a 10 pp increase in seller exposure raises buyer (resp., seller) ratings by about 0.015 (resp., 0.014) points. This finding suggests that improvements in the transaction process for one side translate into higher post-transaction experience for the other side.

As a test for the parallel trends assumption, event-study specifications for both Ratio_buyer_d and Ratio_seller_d show flat pre-trends and immediate, persistent post-rollout changes (Figure 3), supporting a causal interpretation of these two-sided and cross-side patterns.

6 Mechanisms: Reduced Frictions vs. Redistribution

We interpret the own- and cross-side patterns documented in Section 5.4. Recall that we do not observe function-by-function usage, and we proxy exposure to buyer- versus seller-facing workflows using pre-period district shares of buyers and sellers transacting via the app (Section 3). We proceed in three steps. First, we map each cell of the 2×2 own- and cross-side effect matrix to behavioral margins that the copilot could plausibly affect. Second, we document how these margins change. Third, we present a bundle of evidence to help interpret the price increase, distinguishing between friction reduction and pure redistribution.

6.1 Mapping Matrix Cells to Behavioral Margins

Recall the conceptual own- and cross-side effect matrix from Section 2.4. Let E_d^B and E_d^S denote buyer-side and seller-side exposure proxies (empirically, Ratio_buyer_d and Ratio_seller_d). For each outcome Y, the coefficients on Post $\times$ E_d^B and Post $\times$ E_d^S in Equation 2 summarize reduced-form responses $\partial Y/\partial E^B$ and $\partial Y/\partial E^S$ (up to scaling). As summarized in Table 6, we use this matrix to organize which behavioral margins are most informative for interpreting each cell.

6.2 Evidence on Behavioral Margins

This subsection examines the behavioral margins highlighted in Table 6. We first study buyer-side workflow margins and then seller-side workflow margins.

6.2.1 Buyer-Side Workflow Margins

Search and process costs: Consistent with reduced search and coordination frictions in the agent–buyer workflow, we find that buyer-facing exposure is associated with a more efficient search process. In particular, on-site viewings per transaction decline when buyer-side exposure is higher (column 5 of Table 7). We also find improvements in buyer-side time outcomes in the main rollout estimates (Table 4), indicating that buyers complete transactions faster in higher-exposure districts.

Buyer feedback: We next study buyer feedback on agents' service quality. Because these labels and review texts are recorded post-transaction and may be selected, we interpret them as supportive evidence rather than as primary mechanism measures. Buyer-facing exposure is associated with higher perceived recommendation quality and improved communication. As reported in Table 7, buyers are more likely to select the "high-quality recommendation" label (column 1), less likely to report "communication inefficiency" (column 2), and more likely to provide an explicit positive evaluation of communication skills in review texts (column 3). Buyer trust also rises (column 4).

Panel B of Table 7 shows that these improvements are concentrated in exposure to buyer-facing workflows, consistent with buyer workflows affecting buyer process outcomes. For example, a 10 pp increase in buyer-side exposure raises the "high-quality recommendation" label by about 2.3 pp (from a baseline rate of 36.8%), reduces "communication inefficiency" by about 0.6 pp (from 1.0%), increases "communication comfort" by about 0.7 pp (from 1.2%), and raises trust by about 0.6 pp (from 4.2%).

Placebo checks support a workflow-specific interpretation. Attributes not targeted by the buyer-facing tools (e.g., on-site presentation skills, appearance/temperament) show null effects, and buyer-facing communication tools do not shift seller-side communication labels (Appendix Table A6).

6.2.2 Seller-Side Workflow Margins

Pricing: On the seller side, pricing adjustments become more frequent and smaller in magnitude (Table 8, columns 3–4), consistent with faster convergence toward market-clearing prices and reduced trial-and-error in listing strategy. Panel B implies that a 10 pp increase in seller-side exposure increases weekly adjustment frequency by about 0.008 (relative to a baseline mean of 0.187 adjustments per week) and reduces total adjustment magnitude by about 0.079 (baseline mean 4.151). Additional results show that the increase in adjustment frequency is driven disproportionately by downward adjustments (Appendix Table A13), consistent with improved pricing discipline rather than mechanical relabeling.

Seller feedback: We then study seller feedback on the agent's service, again treating these as supportive because they are post-transaction and potentially selected. Seller-facing exposure increases the probability that sellers choose labels indicating "reasonable price adjustment recommendations" and "accurately capturing the property's key selling points" (Table 8, columns 1–2). Panel B confirms that these improvements load on seller-facing workflows: a 10 pp increase in seller-side exposure raises these labels by about 3.3 pp (baseline 25.5%) and 4.1 pp (baseline 24.1%), respectively.

As on the buyer side, placebo and composition checks suggest these patterns are not driven by broad shifts in listing mix or unrelated platform changes (Appendix Tables A5 and A6).

6.3 Mechanism Diagnostics

The main interpretation question is whether the observed increase in transaction prices reflects value-creating channels, such as lower search and coordination costs, improved matching, or faster price discovery, or instead a redistribution of value from buyers to sellers. This concern is especially relevant in our setting, where the copilot assists agents in a market with already-digitized search and communication. A natural alternative interpretation is therefore that buyer-facing assistance primarily strengthens persuasion or bargaining rather than improving search and match formation.

Our data do not allow for a full welfare decomposition, nor do they separately identify bargaining, persuasion, and matching channels. We therefore use the behavioral margin evidence as a set of reduced-form diagnostics, moving from outcomes closest to the buyer-side workflow to broader market-level and seller-side patterns.

(1) Buyer-side process outcomes improve. Buyer-side exposure is associated with a more efficient search process. Higher exposure reduces on-site viewings per transaction (Table 7, column 5) and shortens buyer-side time outcomes in the main rollout estimates (Table 4). Since we do not detect post-rollout changes in housing composition, these patterns are consistent with lower search and coordination costs in the agent–buyer workflow rather than with a mechanical change in the mix of transactions.

(2) Buyer feedback improves on dimensions targeted by the tool. We next examine post-transaction feedback, which we treat as supportive rather than primary evidence because it is well-known that ratings and review are selectively provided. Exposure to buyer-facing workflows increases the probability that buyers select the "high-quality recommendation" label, decreases reports of "communication inefficiency," increases explicit positive evaluations of communication skills in review texts, and raises trust (Table 7, columns 1–4).

Overall post-transaction ratings also rise in the rollout estimates (Table 4). These improvements are concentrated in buyer-facing exposure and do not extend to placebo attributes outside the scope of the tool (Appendix Table A6), which is consistent with improvements in the targeted buyer workflow rather than with a broad shift in review propensity.

(3) Higher prices coincide with faster market clearing. At the market level, higher exposure is associated with both higher transaction prices and faster clearing: homes are more likely to sell within the horizon, and transaction times fall (Table 4). While these are not direct estimates of demand, the joint pattern of higher prices and faster clearing is more consistent with stronger effective demand or improved matching than with an inward shift of supply. Combined with the absence of composition changes, it suggests that the price increase does not simply reflect worse matches at higher prices.

(4) Seller-side exposure improves pricing workflow but does not drive the transaction price effect. The side-specific decomposition shows that the transaction-price effect loads primarily on Post $\times$ Ratio_buyer_d, while Post $\times$ Ratio_seller_d has little effect on prices (Table 4, Panel B, column 4). Instead, seller-side exposure mainly shows up in seller-workflow margins: price adjustments become more frequent and smaller in magnitude, seller-facing feedback improves on price recommendations and listing presentation, and the increase in adjustment frequency is driven disproportionately by downward adjustments (Table 8; Appendix Table A13). This pattern is less consistent with a story in which seller-facing tools are the primary source of rent extraction, and more consistent with the copilot improving pricing discipline and listing execution on the seller side while the main transaction-price effect operates through buyer-side search and coordination, and possibly match formation.

Taken together, these diagnostics are more consistent with value-creating channels, especially lower search and coordination costs, and possibly improved matching or faster price discovery, than with a pure redistribution of value from buyers to sellers. At the same time, the evidence remains reduced-form: we do not observe willingness-to-pay or feature-level usage, and we do not estimate a full welfare decomposition. The mechanism interpretation should therefore be read as disciplined but not definitive.

7 Conclusion

This paper studies an intermediary-facing AI copilot introduced on a large resale housing platform in China. Combining a neighborhood-level field experiment with a multi-city exposure difference-in-differences design, we find that copilot access improves speed and process efficiency: listings are more likely to sell quickly, and time-on-market falls for both buyers and sellers. Exploiting side-specific exposure, we show that efficiency gains are two-sided, with economically meaningful cross-side effects, while price effects load primarily on buyer-side exposure rather than seller-side exposure.

A key lesson is conceptual: in two-sided, intermediated markets, evaluating intermediary-facing tools requires separating own-side effects from cross-side effects and characterizing how effects load differently on speed vs. prices. At the same time, what we do not learn is equally important. We do not observe micro-level tool usage, so mechanism attribution is indirect; the intervention is bundled, so we cannot isolate feature-level effects. More broadly, the mapping from efficiency gains to prices and value redistribution is not mechanical: it depends on market design features—including disclosure and transparency, the strength of competition among intermediaries, and bargaining institutions—that shape how cost reductions and information improvements pass through to each side. Finally, our estimates are best interpreted as adoption-phase effects; equilibrium outcomes under universal access may differ as competitive advantages attenuate and participation incentives adjust.

These limitations suggest disciplined implications. The speed, process efficiency, and the "leveling" patterns are plausibly portable because they reflect reductions in information and coordination frictions and augmentation of intermediary skill, which are margins that arise even when listings are digitized. In comparison, the price response and value split are likely context-specific, shaped by the relative access to and effectiveness of buyer- vs. seller-facing tools and by institutions governing incentives and bargaining. In our setting, price effects load primarily on buyer-side exposure, while speed gains load on both sides, suggesting that different tool investments can tilt outcomes toward liquidity vs. prices. More broadly, value distribution matters for platform design: persistent skew in value capture can discourage participation on the disadvantaged side and reduce market thickness over time, so platforms have incentives to consider both own-side and cross-side effects when choosing the tool mix and rollout strategy. Overall, the evidence indicates that intermediary-facing AI can reduce search and coordination frictions, and possibly improve matching efficiency in complex marketplaces, but that own- and cross-side effects are first-order objects for understanding both aggregate performance and who captures value.

References

Agarwal, S., Choi, H.-S., He, J., and Sing, T. F. (2019a). Matching in housing markets: The role of ethnic social networks. *The Review of Financial Studies*, 32(10):3958–4004.

Agarwal, S., Ghosh, P., Li, J., and Ruan, T. (2024). Digital payments and consumption: Evidence from the 2016 demonetization in india. *The Review of Financial Studies*, 37(8):2550–2585.

Agarwal, S., He, J., Sing, T. F., and Song, C. (2019b). Do real estate agents have information advantages in housing markets? *Journal of Financial Economics*, 134(3):715–735.

Agrawal, A., Gans, J., Goldfarb, A., and Tucker, C. E. (2024). *The economics of artificial intelligence: Health Care Challenges*. University of Chicago Press.

Aiello, D., Garmaise, M., and Nadauld, T. (2026). What problem do intermediaries solve? evidence from real estate markets. *The Review of Financial Studies*, 39(2):562–604.

Allen, J., Clark, R., Houde, J.-F., Li, S., and Trubnikova, A. (2025). The role of intermediaries in selection markets: Evidence from mortgage lending. *The Review of Financial Studies*, 38(11):3284–3328.

Babina, T., Fedyk, A., He, A., and Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151:103745.

Bartlett, R., Morse, A., Stanton, R., and Wallace, N. (2022). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics*, 143(1):30–56.

Bertomeu, J., Lin, Y., Liu, Y., and Ni, Z. (2025). The impact of generative ai on information processing: Evidence from the ban of chatgpt in italy. *Journal of Accounting and Economics*, page 101782.

Bourveau, T., De George, E. T., Ellahie, A., and Macciocchi, D. (2022). The role of disclosure and information intermediaries in an unregulated capital market: Evidence from initial coin offerings. *Journal of Accounting Research*, 60(1):129–167.

Bradshaw, M. T., Ma, C., Yost, B., and Zou, Y. (2026). Generative ai use by capital market information intermediaries: Evidence from seeking alpha. *Available at SSRN 5226562*.

Braggion, F., Manconi, A., Pavanini, N., and Zhu, H. (2025). The value of financial intermediation: Evidence from online debt crowdfunding. *Journal of Financial Economics*, 172:104113.

Brynjolfsson, E., Hui, X., and Liu, M. (2019). Does machine translation affect international trade? evidence from a large digital platform. *Management Science*, 65(12):5449–5460.

Brynjolfsson, E., Li, D., and Raymond, L. (2025). Generative ai at work. *The Quarterly Journal of Economics*, page qjae044.

Burtch, G., Lee, D., and Chen, Z. (2024). The consequences of generative ai for online knowledge communities. *Scientific Reports*, 14(1):10413.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. (2024). Difference-in-differences with a continuous treatment. Technical report, National Bureau of Economic Research.

Cao, S., Jiang, W., Wang, J., and Yang, B. (2024). From man vs. machine to man+ machine: The art and ai of stock analyses. *Journal of Financial Economics*, 160:103910.

Cao, S., Jiang, W., Yang, B., and Zhang, A. L. (2023). How to talk when a machine is listening: Corporate disclosure in the age of ai. *The Review of Financial Studies*, 36(9):3603–3642.

Castillo, J. C. (2023). Who benefits from surge pricing? *Available at SSRN 3245533*.

Cong, L. W., Guo, Y., Zhao, X., and Zhou, W. (2024). Writing quality and soft information in the genai age: Evidence from online credit markets. *Available at SSRN 4959535*.

Cowgill, B., Dell'Acqua, F., Deng, S., Hsu, D., Verma, N., and Chaintreau, A. (2020). Biased programmers? or biased data? a field experiment in operationalizing ai ethics. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 679–681.

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., and Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of ai on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (24-013).

Demirci, O., Hannane, J., and Zhu, X. (2025). Who is ai replacing? the impact of generative ai on online freelancing platforms. *Management Science*.

Eisfeldt, A. L., Schubert, G., and Zhang, M. B. (2023). Generative ai and firm values. Technical report, National Bureau of Economic Research.

Erel, I., Stern, L. H., Tan, C., and Weisbach, M. S. (2021). Selecting directors using machine learning. *The Review of Financial Studies*, 34(7):3226–3264.

Fong, J., Manchanda, P., and Song, Y. (2023). How effective is suggested pricing?: Experimental evidence from an e-commerce platform. *Experimental Evidence from an E-Commerce Platform (November 10, 2023)*.

Fu, R., Huang, Y., Mehta, N., Singh, P. V., and Srinivasan, K. (2025). Unequal impact of zestimate on the housing market. *Marketing Science*.

Fu, R., Jin, G. Z., and Liu, M. (2022). Does human-algorithm feedback loop lead to error propagation? evidence from zillow's zestimate. Technical report, National Bureau of Economic Research.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., and Walther, A. (2022). Predictably unequal? the effects of machine learning on credit markets. *The Journal of Finance*, 77(1):5–47.

Fuster, A., Plosser, M., Schnabl, P., and Vickery, J. (2019). The role of technology in mortgage lending. *The review of financial studies*, 32(5):1854–1899.

Gilbukh, S. and Goldsmith-Pinkham, P. (2024). Heterogeneous real estate agents and the housing cycle. *The Review of Financial Studies*, 37(11):3431–3489.

Glode, V. and Opp, C. (2016). Asymmetric information and intermediation chains. *American Economic Review*, 106(9):2699–2721.

Goldberg, S. and Lam, H. T. (2025). Generative ai in equilibrium: Evidence from a creative goods marketplace. *Available at SSRN 5152649*.

Goldstein, I., Jiang, W., and Karolyi, G. A. (2019). To fintech and beyond. *The Review of Financial Studies*, 32(5):1647–1661.

Hall, J., Kendrick, C., and Nosko, C. (2015). The effects of uber's surge pricing: A case study. *The University of Chicago Booth School of Business*.

Hau, H., Huang, Y., Lin, C., Shan, H., Sheng, Z., and Wei, L. (2024). Fintech credit and entrepreneurial growth. *the Journal of Finance*, 79(5):3309–3359.

Huang, Y. (2024). Pricing frictions and platform remedies: the case of airbnb. *Available at SSRN 3767103*.

Hui, X., Reshef, O., and Zhou, L. (2024). The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organization Science*, 35(6):1977–1989.

Levitt, S. D. and Syverson, C. (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *The Review of Economics and Statistics*, 90(4):599–611.

Liu, M., Tang, X., Xia, S., Zhang, S., Zhu, Y., and Meng, Q. (2023). Algorithm aversion: Evidence from ridesharing drivers. *Management Science*.

Luo, X., Tong, S., Fang, Z., and Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, 38(6):937–947.

Malik, N. and Manzoor, E. (2023). Does machine learning amplify pricing errors in the housing market?–the economics of machine learning feedback loops. *arXiv preprint arXiv:2302.09438*.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Pan, Q. and Wang, W. (2021). Costly price adjustment and automated pricing: The case of airbnb. *Available at SSRN 4077985*.

Peng, S., Kalliamvakou, E., Cihon, P., and Demirer, M. (2023). The impact of ai on developer productivity: Evidence from github copilot. *arXiv preprint arXiv:2302.06590*.

Peukert, C., Abeillon, F., Haese, J., Kaiser, F., and Staub, A. (2024). Strategic behavior and ai training data. *arXiv preprint arXiv:2404.18445*.

Schanke, S., Burtch, G., and Ray, G. (2021). Estimating the impact of "humanizing" customer service chatbots. *Information Systems Research*, 32(3):736–751.

Shan, G. and Qiu, L. (2025). Examining the impact of generative ai on users' voluntary knowledge contribution: Evidence from a natural experiment on stack overflow. *Information Systems Research*.

Shi, A., Chan, T., and Zhang, D. (2024). Using algorithmic scores to measure the impacts of targeting promotional messages. *Available at SSRN 4262156*.

Weller, B. M. (2018). Does algorithmic trading reduce information acquisition? *The Review of Financial Studies*, 31(6):2184–2226.

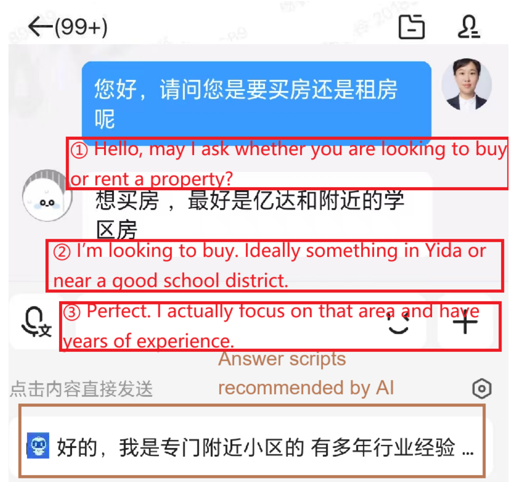
Wiles, E., Munyikwa, Z., and Horton, J. (2025). Algorithmic writing assistance on jobseekers'

resumes increases hires. *Management Science*.
Yang, K. (2025). Trust as an entry barrier: Evidence from fintech adoption. *Journal of Financial Economics*, 169:104062.
Yu, S. (2020). Algorithmic outputs as information source: The effects of zestimates on home prices and racial bias in the housing market. *Available at SSRN 3584896*.
Zhang, M., Sun, T., Luo, L., and Golden, J. (2021a). Consumer and ai co-creation: When and why human participation improves ai creation. *USC Marshall School of Business Research Paper Sponsored by iORB, No. Forthcoming*.
Zhang, S., Mehta, N., Singh, P. V., and Srinivasan, K. (2021b). Frontiers: Can an artificial intelligence algorithm mitigate racial economic inequality? an analysis in the context of airbnb. *Marketing Science*, 40(5):813–820.
Zhou, E. and Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS nexus*, 3(3):pgae052.
Zhu, K. and Walker, D. (2024). The promise and pitfalls of ai technology in bridging digital language divide: Insights from machine translation on wikipedia. *Available at SSRN 4708614*.

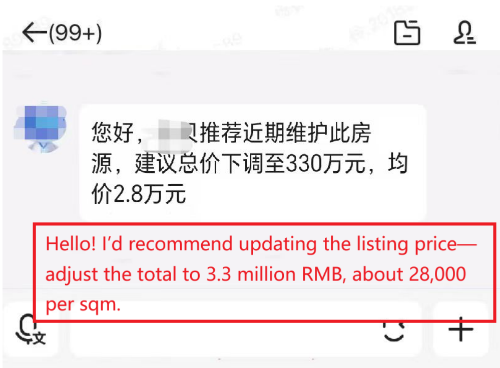
Figures and Tables

Figure 1: Sample AI Output: The App Interface for Agents

(a) AI-Recommended Listings (Functioning for Buyers)



(b) AI-Recommended Dialogue (Functioning for Buyers)

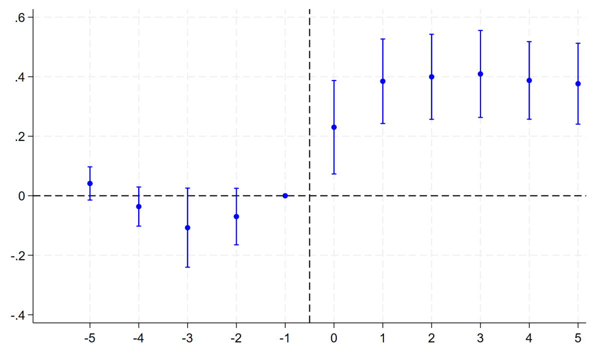


(c) AI-Recommended Price Adjustments (Functioning for Sellers)

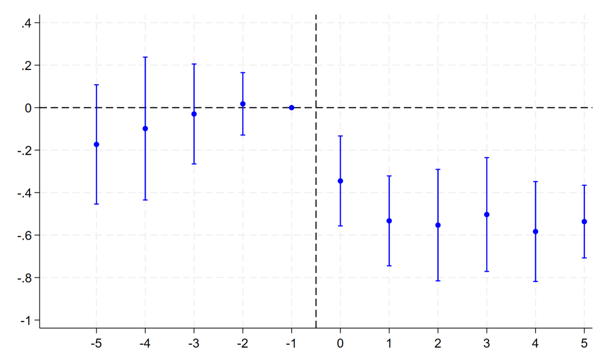
(d) AI-Recommended Selling-Point Descriptions (Functioning for Sellers)

Notes. This figure shows examples of AI-generated recommendations for agents and the corresponding app interface for agents. These recommendations are respectively targeted at buyers and sellers. Panel (a) displays the specific properties recommended by the AI copilot based on buyers' needs. Panel (b) shows an example where the buyer expressed interests in properties in highly ranked school zones; the AI copilot recommended the agent to use a friendly, professional reply highlighting the agent's experience and expertise in the neighborhood. Panel (c) shows the AI copilot's suggestion for a recent price adjustment to the listed property. Panel (d) provides appealing, optimized property title suggestions designed to attract buyers and improve conversion rates.

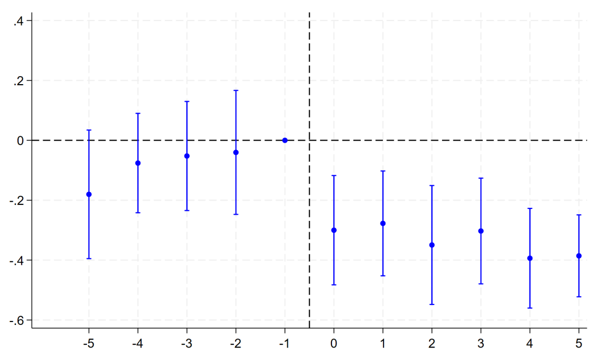
Figure 2: Event Studies for Main Outcomes: Total Effects of AI Adoption



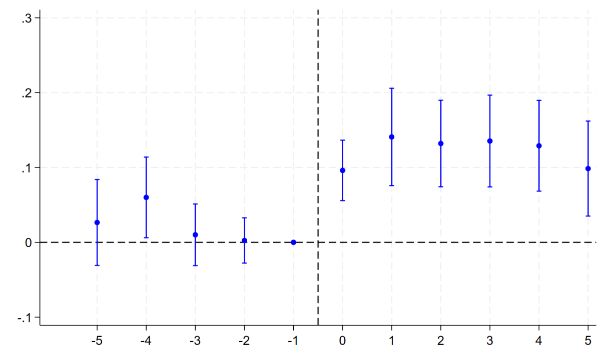
(a) Sold<3month (using Ratio_tran)



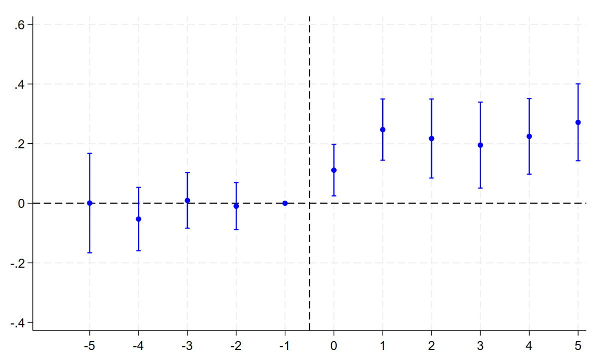
(b) Log_days_seller (using Ratio_tran)



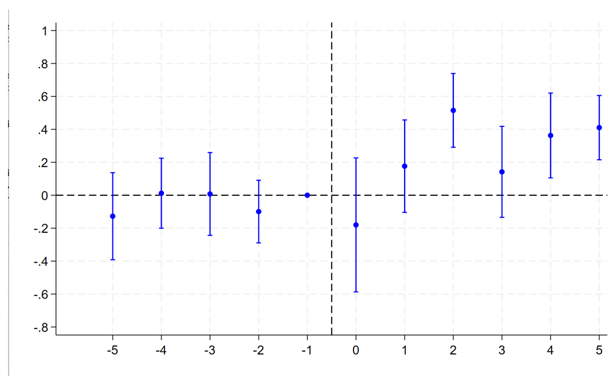
(c) Log_days_buyer (using Ratio_tran)



(d) Log_price (using Ratio_tran)



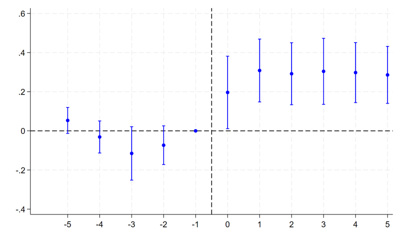
(e) Buyers' ratings (using Ratio_tran)



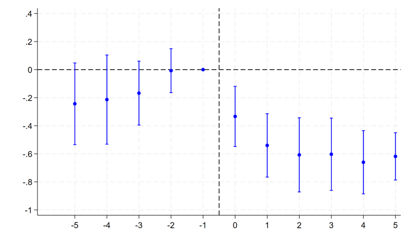
(f) Sellers' ratings (using Ratio_tran)

Notes. This figure plots the bimonthly coefficients and 95% confidence intervals from event study regressions (Equation 3) on the following outcomes: (a) whether the listed property is sold within three months, (b) time on market for sellers, (c) time on market for buyers, (d) unit transaction prices, (e) buyers' ratings, and (f) sellers' ratings. Standard errors are clustered at the district level.

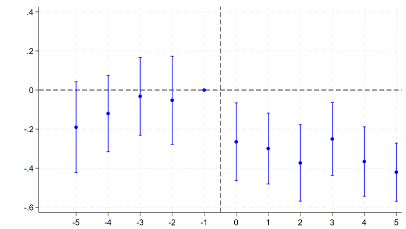
Figure 3: Event Studies for Main Outcomes: Buyer-Side and Seller-Side Effects of AI Copilot Adoption



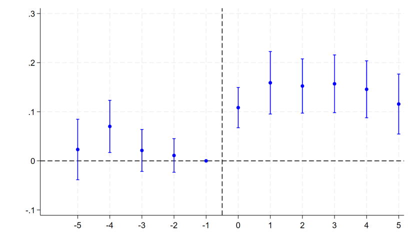
(a) Sold<3month (Ratio_buyer)



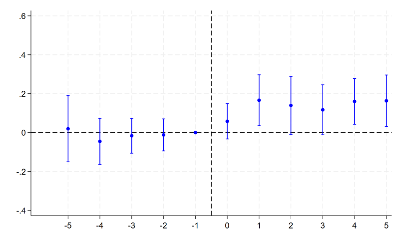
(b) Log_days_seller (Ratio_buyer)



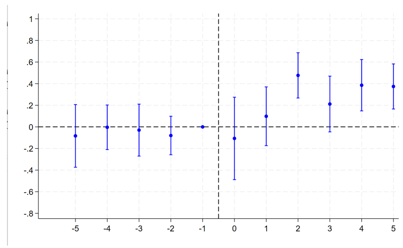
(c) Log_days_buyer (Ratio_buyer)



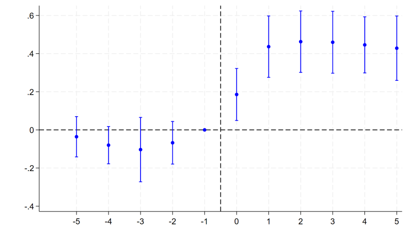
(d) Log_price (Ratio_buyer)



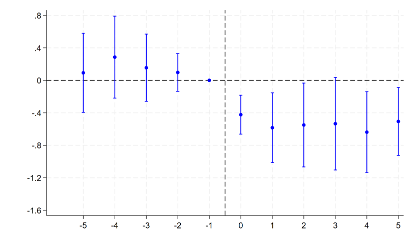
(e) Rating_buyer (Ratio_buyer)



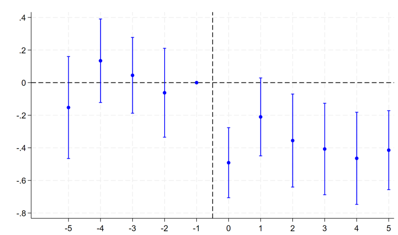
(f) Rating_seller (Ratio_buyer)



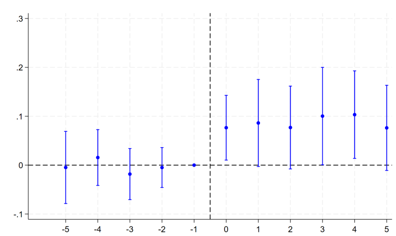
(g) Sold<3month (Ratio_seller)



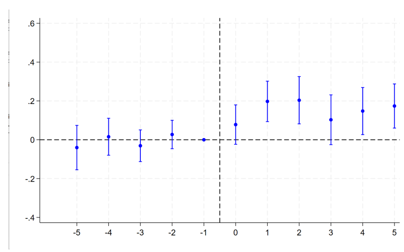
(h) Log_days_seller (Ratio_seller)



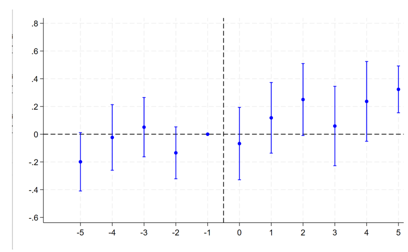
(i) Log_days_buyer (Ratio_seller)



(j) Log_price (Ratio_seller)



(k) Rating_buyer (Ratio_seller)



(l) Rating_seller (Ratio_seller)

Notes. This figure plots the bimonthly coefficients and 95% confidence intervals from event study regressions (Equation 3) regarding buyer-side and seller-side on the following outcomes: whether the listed property is sold within three months, time on market for sellers, time on market for buyers, unit transaction prices, and buyers' and sellers' ratings. Ratio_buyer and Ratio_seller are the pre-treatment district shares of buyers and sellers transacting via the app, respectively. These proxies map to exposure to buyer-facing vs. seller-facing workflows because those features live in the agent–buyer vs. agent–seller chat flows. Standard errors are clustered at the district level.

Table 1: Operational Content of The AI Copilot Rollout

Workflow module	**Underlying inputs**	**Pre-rollout baseline**	**Execution and visibility**
Buyer side: Reply scripts	Ongoing chat text and communication context	Agent manually drafts in-app replies	Advisory. Agent can send, edit, or ignore. Raw suggestion visible only to the agent.
Buyer side: Property recommendation	Buyer browsing history, search criteria, and chat text	Agent manually filters and recommends listings using standard platform tools and judgment	Advisory. Agent chooses which listings to forward. Raw suggestion visible only to the agent.
Seller side: Price adjustment recommendation	Comparable listings and contemporaneous market information	Agent relies on manual pricing comparisons and judgment	Advisory. Agent proposes changes, but the agent or seller controls the actual list-price changes. Raw suggestion visible only to the agent.
Seller side: Title/description generation	Property attributes and listing information	Agent manually drafts titles and descriptions	Advisory. Agent can edit or ignore before publishing. Draft visible only to the agent; adopted listing text becomes public.

Notes: The January 2019 rollout of the AI copilot bundled four module-specific functions embedded in existing agent tasks. Buyer-side modules map to the agent–buyer workflow, and seller-side modules map to the agent–seller workflow. The table describes the operational scope of the intervention, not realized module-level adoption. Our transaction-level data do not allow us to observe feature-specific usage, so the empirical analysis proxies exposure using pre-rollout buyer-side and seller-side app usage.

Table 2: Summary Statistics

Variable	Mean	SD	P10	Median	P90
Sold<3month	0.546	0.498	0	1	1
Log_days_seller	3.982	1.256	2.197	4.094	5.533
Log_days_buyer	2.664	1.656	0	2.639	4.905
Log_price_unit	10.274	0.763	9.326	10.267	11.275
Rating (buyer)	4.884	0.392	5	5	5
Rating (seller)	4.693	0.632	4	5	5
Log_listed_price	10.308	0.767	9.357	10.305	11.315
Ratio_tran	0.446	0.222	0.110	0.39	0.747
Ratio_buyer	0.388	0.208	0.117	0.353	0.716
Ratio_seller	0.265	0.162	0.085	0.276	0.444
Building size	84.28	34.178	49	80	128
Building age	3.347	1.507	1	3	5
Floor	2.118	0.801	1	2	3
Decoration	2.475	0.642	2	3	3
Room	2.183	0.794	1	2	3
School	0.478	0.5	0	0	1
Subway	0.525	0.499	0	1	1
Population	167.032	107.259	59	123	361
Unemployment	2.065	0.889	0.720	2	3.1
Gdp	12.432	7.129	5.268	9.989	19.284
Income	5.49	1.356	3.955	4.978	7.555
Education	0.092	0.288	0	0	0
Gender	0.684	0.465	0	1	1
Performance	3.43	1.36	1.130	4.02	4.71
Tenure	4.879	3.278	1	4	10

Notes. This table presents summary statistics for the 231,680 transacted properties. The variable definitions are as follows: *building size* represents the indoor area in square meters; *building age* is an indicator for the building's age: 1 for 0–5 years, 2 for 5–10 years, 3 for 10–15 years, 4 for 15–20 years, 5 for 20–30 years, and 6 for more than 30 years; $floor$ indicates the property's floor level: 1 for low, 2 for middle, and 3 for high; *decoration* is an indicator for decoration status: 1 for undecorated, 2 for lightly decorated, and 3 for well-decorated; *room* is an indicator of number of rooms: 1 for one bedroom, 2 for two bedrooms, 3 for three bedrooms, 4 for four bedrooms, and 5 for more than four bedrooms; *school* is a dummy variable indicating proximity to a school: 1 for close, 0 otherwise; *subway* is a dummy variable indicating proximity to a subway: 1 for close, 0 otherwise; *Population* is the district's population in 10,000 people; $Income$ is the district-level per capita net income in 10,000 RMB; *Gdp* is the district's per capita GDP in 10,000 RMB; and $Unemployment$ is the district-level unemployment rate. "Rating (buyer)" and "rating (seller)" refer to the 5-point scale ratings (from 1 to 5) provided by buyers and sellers, respectively.

Table 3: Field Experiment: Overall Effects of Copilot Access (ITT)

	(1) Sold< 3 month	**(2)** Log_days _seller	**(3)** Log_days _buyer	**(4)** Log_price _unit	**(5)** Rating (buyer)	**(6)** Rating (seller)
Post×Treat	0.1106***	-0.3704***	-0.4506***	0.0476***	0.1247***	0.2323**
	(0.040)	(0.096)	(0.121)	(0.017)	(0.041)	(0.096)
Implied % change	–	-31.0%	-36.3%	4.8%	–	–
Control mean	0.5267	–	–	–	4.9312	4.7431
Observations	6,748	4,031	4,031	4,031	1,378	561
Adjusted R^2	0.110	0.100	0.033	0.644	0.128	0.031
District×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimation results of the field experiment. The firm randomized 34 districts to a group in which access was maintained and 34 to a group in which access was blocked. Controls include apartment characteristics (e.g., building age, number of rooms) as well as unemployment, population, income, and GDP in the respective region. Implied % change reports $100 \times (\exp(\hat{\beta}) - 1)$ for logged outcomes. The sample size in column 1 (i.e., Sold <3months) is larger than those in columns 2–4 because column 1 is based on all listed properties and measures whether a property is sold within three months of listing, whereas columns 2–4 rely on transaction-level data and therefore include only completed transactions. Standard errors clustered at the neighborhood level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Main Rollout: Overall Effects and Decomposition

	(1) Sold< 3 month	(2) Log_days _seller	(3) Log_days _buyer	(4) Log_price _unit	(5) Rating (buyer)	(6) Rating (seller)
Panel A: Overall Impact of AI Copilot						
Post×Ratio_tran	0.4215***	-0.4434***	-0.2484***	0.1005***	0.2206***	0.3481***
	(0.080)	(0.145)	(0.047)	(0.029)	(0.053)	(0.061)
Effect of +10pp exposure	+4.2 pp	-4.3%	-2.5%	+1.0%	+0.022	+0.035
Control mean	0.546	–	–	–	4.884	4.693
Panel B: Decomposed Effects and Cross-Side Effects						
Post×Ratio_buyer	0.1744*	-0.2376**	-0.1411**	0.1152***	0.1340***	0.2936***
	(0.100)	(0.107)	(0.055)	(0.028)	(0.041)	(0.056)
Post×Ratio_seller	0.4117***	-0.4877**	-0.2550***	-0.0005	0.1525***	0.1354**
	(0.103)	(0.229)	(0.089)	(0.033)	(0.048)	(0.055)
Effect of +10pp buyer exp	+1.7 pp	-2.3%	-1.4%	+1.2%	+0.013	+0.029
Effect of +10pp seller exp	+4.1 pp	-4.8%	-2.5%	0.0%	+0.015	+0.014
Observations	267,882	231,680	231,680	231,680	75,596	20,336
Adjusted R^2	0.222	0.079	0.128	0.948	0.047	0.079
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table presents the estimation results of Equation 1. Panel A shows the overall effects. Panel B shows the impact decomposed by the AI copilot's buyer- and seller-facing exposure. Here, Post × Ratio_buyer and Post × Ratio_seller are pre-period district shares of buyers and sellers transacting via the app, respectively. Controls include apartment characteristics (e.g., building age, number of rooms) as well as unemployment, population, income, and GDP in the respective region. For logged outcomes, Effect of +10pp exposure reports $100 \times (\exp(0.10\hat{\beta}) - 1)$. For binary outcomes, effects are reported in percentage points. For ratings, the effect is reported in rating points. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Heterogeneity by Agent Tenure

	(1) Sold< 3 month	(2) Log_days _seller	(3) Log_days _buyer	(4) Log_price _unit	(5) Rating (buyer)	(6) Rating (seller)
Panel A: Moderation of Overall Effect						
Post×Ratio_tran	0.518***	-0.508***	-0.315***	0.101***	0.2954***	0.5759***
	(0.082)	(0.149)	(0.052)	(0.029)	(0.060)	(0.182)
Post×Ratio_tran×Tenure	-0.0341***	0.0222***	0.0249***	-0.0002	-0.0147**	-0.0683**
	(0.003)	(0.008)	(0.007)	(0.001)	(0.007)	(0.033)
Panel B: Moderation of Decomposed Effects						
Post×Ratio_buyer	0.258**	-0.297***	-0.180***	0.117***	0.2064***	0.6272***
	(0.102)	(0.110)	(0.058)	(0.028)	(0.041)	(0.162)
Post×Ratio_buyer×Tenure	-0.0308***	0.0229**	0.0158	-0.0005	-0.0144**	-0.0838**
	(0.004)	(0.009)	(0.010)	(0.001)	(0.007)	(0.034)
Post×Ratio_seller	0.448***	-0.505**	-0.323***	0.001	0.1420***	-0.0728
	(0.105)	(0.231)	(0.093)	(0.034)	(0.043)	(0.172)
Post×Ratio_seller×Tenure	-0.0109*	0.0053	0.0247*	0.0015	0.0021	0.0309
	(0.007)	(0.009)	(0.013)	(0.002)	(0.006)	(0.037)
Observations	267,882	231,680	231,680	231,680	75,595	20,336
Adjusted	0.224	0.079	0.128	0.948	0.049	0.114
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table presents heterogeneity analysis by agent's tenure (standardized). For brevity, only the highest-order interaction terms and their constituent two-way interactions are shown. All specifications include the full set of main effects and controls based on Equation 1 Standard errors clustered at the district level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Mapping matrix cells to behavioral margins and observed proxies

Matrix cell	**Behavioral margins**	**Observed proxies / evidence**
$\partial Y^B/\partial E^B$ (buyer own-side)	Buyer search and decision efficiency; information processing in the agent–buyer workflow	Buyer time; on-site viewings per transaction; buyer-facing review labels (recommendation quality, communication efficiency, trust) (Table 7)
$\partial Y^S/\partial E^S$ (seller own-side)	Listing and pricing in the agent–seller workflow; speed of convergence to market-clearing price	Price adjustment frequency, direction, magnitude; seller-facing review labels (pricing advice, selling points) (Table 8)
$\partial Y^S/\partial E^B$ (buyer→seller)	Coordination and screening that reduce seller-side time-to-sell; fewer showings / faster buyer decisions	Seller time outcomes loading on E^B (Table 4, Panel B); viewings per transaction (Table 7)
$\partial Y^B/\partial E^S$ (seller→buyer)	Improved listing clarity and pricing that reduce buyer-side time-to-buy; fewer failed negotiations	Buyer time loading on E^S (Table 4, Panel B); seller pricing adjustment dynamics (Table 8)

Table 7: Buyer-Side Workflow: Search Effort and Buyers' Feedback

	(1) Recomm_optimal	(2) Talk_inefficiency	(3) Talk_comfort	(4) Trust	(5) Viewing	(6) Log_listed_price
Panel A: Overall Impact						
Post×Ratio_tran	0.3944***	-0.0683***	0.0688**	0.0575**	-0.0341	0.0963***
	(0.092)	(0.024)	(0.029)	(0.023)	(0.143)	(0.029)
Effect of +10pp exposure	+3.9 pp	-0.7 pp	+0.7 pp	+0.6 pp	-0.003	+1.0%
Control mean	0.368	0.010	0.012	0.042	1.639	–
Panel B: Decomposed Effects						
Post×Ratio_buyer	0.2324**	-0.0608***	0.0714***	0.0591**	-0.3548**	0.1057***
	(0.109)	(0.016)	(0.026)	(0.023)	(0.163)	(0.029)
Post×Ratio_seller	0.2587**	-0.0129	-0.0019	0.0099	0.3094	0.0081
	(0.101)	(0.017)	(0.014)	(0.012)	(0.190)	(0.031)
Effect of +10pp buyer exp	+2.3 pp	-0.6 pp	+0.7 pp	+0.6 pp	-0.035	+1.1%
Effect of +10pp seller exp	+2.6 pp	-0.1 pp	-0.0 pp	+0.1 pp	+0.031	+0.1%
Observations	75,596	75,596	75,596	75,596	13,581,455	231,680
Adjusted R^2	0.201	0.020	0.047	0.044	0.091	0.949
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table reports the regression results from Equation 1, focusing on the buyer-side mechanisms. Columns 1–4 show effects on buyer feedback. Column 5 shows effects on weekly on-site viewing frequency. Column 6 shows effects on the last listed prices of sold properties. For logged outcomes, Effect of +10pp exposure reports $100 \times (\exp(0.10\hat{\beta}) - 1)$. For binary outcomes, effects are reported in percentage points; for viewing, effects are reported in outcome units. Control means are reported in raw outcome units. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Seller-Side Workflow: Pricing Adjustments and Sellers' Feedback

	(1) Pricing_ optimal	**(2)** Sellpoint_ Accuracy	**(3)** Total_adjust _times	**(4)** Total_adjust _mag
		Panel A: Overall Impact		
Post×Ratio_tran	0.2465***	0.3270***	0.0639***	-0.4054**
	(0.065)	(0.087)	(0.013)	(0.155)
Effect of +10pp exposure	+2.5 pp	+3.3 pp	+0.006	-0.041
Control mean	0.255	0.241	0.187	4.151
		Panel B: Decomposed Effects		
Post×Ratio_buyer	0.0041	0.0746	0.0204	0.2791
	(0.070)	(0.051)	(0.027)	(0.204)
Post×Ratio_seller	0.3331***	0.4149***	0.0820***	-0.7933***
	(0.058)	(0.053)	(0.031)	(0.244)
Effect of +10pp buyer exp	+0.0 pp	+0.7 pp	+0.002	+0.028
Effect of +10pp seller exp	+3.3 pp	+4.1 pp	+0.008	-0.079
Observations	20,336	20,336	13,581,455	2,171,883
Adjusted R^2	0.125	0.134	0.027	0.054
Controls	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes

Notes. This table presents regression results based on Equation 1 for the seller-side mechanisms. Columns 1 and 2 show effects on seller feedback labels. Columns 3 and 4 show effects on the frequency and magnitude of weekly price adjustments. All columns include location and City×Month fixed effects. For binary outcomes, effects are reported in percentage points; for price adjustment frequency and magnitude, effects are reported in outcome units. Control means are reported in raw outcome units. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix A. Data, Measures, and Descriptives

This appendix provides additional detail on data construction, variable definitions, and descriptive statistics used in the main analysis and mechanism sections. The key point for the empirical design is that we observe (i) transaction outcomes and post-transaction experience measures at the listing/transaction level, and (ii) weekly process measures (e.g., viewings and price adjustments) at the listing-week level. These complementary panels allow us to link market outcomes to workflow margins.

Figure A1 summarizes how we link the platform's listing inventory, completed transactions, weekly listing activity, and the post-transaction feedback system. The unit of observation differs across modules: the main outcomes are measured at the listing/transaction level, while process outcomes (viewings and price adjustments) are measured at the listing-week level.

Table A1 reports summary statistics for the additional process measures and feedback labels used in Sections 6 and the Online Appendix. The demand- and seller-side weekly process datasets contain 13,581,455 listing-week observations, while the buyer and seller feedback modules contain 75,596 and 20,336 transaction-level evaluations, respectively. We report means and dispersion because many of these variables are sparse indicators (feedback labels) or exhibit right-skew (viewings and adjustment magnitudes).

Figure A1: Flow chart

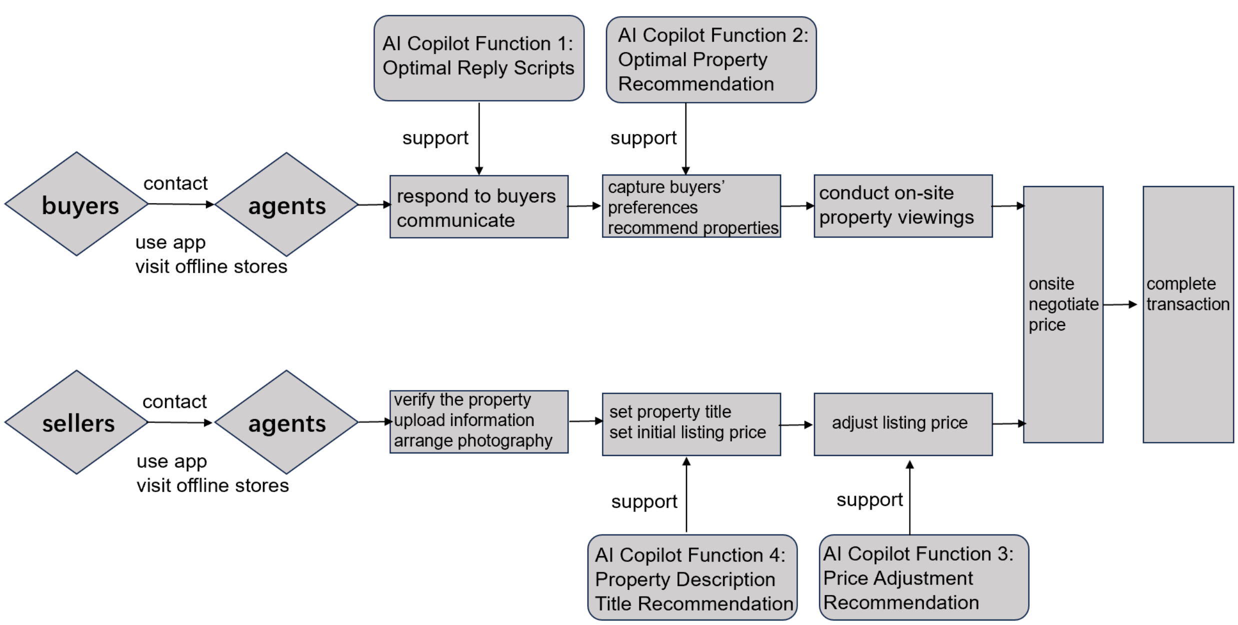


Table A1: Additional Summary statistics

	Mean	SD	P10	Median	P90
Demand side and supply side					
Total_adjust_times	0.187	0.533	0	0	1
Up_adjust_times	0.038	0.206	0	0	0
Down_adjust_times	0.149	0.469	0	0	1
Total_adjust_mag	4.151	3.628	1.150	3.16	8.29
Up_adjust_mag	4.75	4.684	1.120	3.33	10
Down_adjust_mag	3.96	3.325	1.090	3.09	7.8
Viewing	1.639	2.871	0	1	4
Log_volume	1.365	0.63	0.693	1.099	2.303
Num_list	233.495	228.641	42	166	495
Textual feedback (buyer)					
Recommend_optimal	0.368	0.482	0	0	1
Intro_well	0.097	0.291	0	0	0
Talk_inefficiency	0.01	0.097	0	0	0
Talk_comfort	0.012	0.108	0	0	0
Trust	0.042	0.201	0	0	0
Appearance	0.421	0.494	0	0	1
Textual feedback (seller)					
Pricing_optimal	0.255	0.436	0	0	1
Sellingpoint_accuracy	0.241	0.427	0	0	0
Talk_inefficiency	0.006	0.080	0	0	0
Talk_comfort	0.009	0.093	0	0	0
Trust	0.047	0.213	0	0	0
Appearance	0.252	0.434	0	0	1

Notes. This table presents summary statistics for the additional variables in the buyer-side and seller-side datasets (which include 13,581,455 data points), as well as buyers' and sellers' evaluation data in the feedback system (which contain 75,596 and 20,336 data points, respectively).Specifically, Total_adjust_times refers to the number of price adjustments made to a property each week. Up_adjust_times refers to the number of upward price adjustments made to a property in a given week. Down_adjust_times refers to the number of downward price adjustments made to a property in a given week. Total_adjust_mag, Up_adjust_mag, and Down_adjust_mag respectively denote the total magnitude (absolute value) of price adjustments for the property during the week, the magnitude of upward price adjustments, and the magnitude of downward price adjustments. Viewing refers to the on-site viewing frequency of the property during the week. Volume represents the weekly transaction volume at the neighborhood level. Num_list represents the number of listed properties with the same number of rooms as the sold property in the neighborhood during the pre-treatment period. The textual feedback variables from Buyer and Seller indicate whether specific labels were selected in the feedback system. For instance, Recommend_optimal reflects that the property recommendations were precise, while Intro_well suggests that the agent's explanations were clear and effective. Talk_inefficiency denotes that the communication was ineffective, and Trust signals that the agent was perceived as trustworthy. Appearance captures positive evaluations of the agent's personal image and demeanor, whereas Pricing_optimal reflects that the pricing advice was considered appropriate. Moreover, Sellingpoint_accuracy indicates that the agent accurately highlight the key selling points of the property. In addition, we construct an extra variable, Talk_comfort, based on the review text to capture explicit positive evaluations of the agent's communication skills.

Appendix B. Field Experiment Randomization Check

This appendix assesses the baseline balance of the Jinan neighborhood-level experiment used in Section 4. We compare pre-treatment means of key outcomes and housing characteristics between treatment and control neighborhoods. Balance supports interpreting the RCT estimates as causal ITT effects of copilot access.

Table A2 shows that treatment and control neighborhoods are similar along baseline sale probability, time outcomes, prices, ratings, and property characteristics. Differences are economically small and statistically insignificant, consistent with successful random blocking of copilot access.

Table A2: Randomization Check

	Treatment Group	Control Group	p-Value of T-Test
Sold<3 month	0.5369	0.5267	0.28
Log_days_seller	3.7672	3.7777	0.43
Log_days_buyer	2.3362	2.3272	0.45
Log_price	9.8379	9.8473	0.21
Rating (buyer)	4.9261	4.9312	0.85
Rating (seller)	4.7232	4.7431	0.78
Building size	81.9303	82.0814	0.46
Building age	3.4686	3.4120	0.22
Floor	2.1543	2.1381	0.34
Decoration	2.4344	2.4463	0.35
Room	2.1380	2.1287	0.39

Appendix C. Exposure DiD Identification and Robustness

This appendix provides additional evidence supporting the identification strategy and robustness of the multi-city exposure difference-in-differences design in Section 5. Unless stated otherwise, specifications follow Equation 1, include neighborhood fixed effects and city-by-month fixed effects, and cluster standard errors at the district level. The checks below address (i) whether exposure is capturing digitization rather than AI, (ii) interference/spillovers across districts, (iii) overlap between buyer- and seller-side exposure, (iv) differential trends correlated with observables, (v) compositional shifts, (vi) placebo outcomes, (vii) functional form, and (viii) inference.

Exposure stability and digitization concerns. A key concern is that the copilot rollout might coincide with a discrete shift toward app-based transactions, in which case exposure (based on pre-rollout app share) could proxy digitization rather than differential reach of the copilot. Figure A2 plots the app transaction share over time and shows a smooth trend with no visible break at the rollout date. This pattern, together with city-by-month fixed effects, mitigates the interpretation that our estimates are driven by a contemporaneous adoption jump of the app channel.

Differential post trends correlated with observables. Another concern is that districts with higher baseline app penetration may have different post-2019 trajectories for reasons unrelated to the copilot. Table A3 addresses this by allowing the post period to load flexibly on key observables, interacting $Post_t$ with time-varying district-level conditions (e.g., unemployment, population, income, GDP) while maintaining the full set of fixed effects. The main exposure coefficients are stable, alleviating concerns about confounding from observable differential trends.

Alternative exposure based on realized copilot usage. Our baseline exposure measures use pre-period app penetration to proxy the differential reach of the copilot after rollout, which avoids endogeneity from post-treatment adoption. As a supportive robustness check, Table A4 replaces these proxies with district-level copilot *usage intensity* measured in the month immediately following rollout (Usage_tran for overall usage; Usage_buyer and Usage_seller for buyer- and seller-side usage). We then re-estimate Equation 1 using Post$\times$Usage interactions with the same fixed effects and controls as in the baseline specifications. Usage variables are measured after rollout, they may respond endogenously to unobserved shocks or to the perceived returns to adopting the copilot. Nevertheless, the similarity of signs and magnitudes to the baseline results, together with the persistence of cross-side patterns in the decomposed specification, supports the interpretation that the main findings are driven by realized copilot take-up rather than by latent district characteristics correlated with pre-period app penetration.

Composition of listings. The exposure design compares changes across districts with different baseline app penetration. If listing composition shifts differentially with exposure (e.g., higher-exposure districts list systematically different properties post-rollout), outcome changes could reflect selection rather than the copilot. Table A5 tests whether $Post_t \times Ratio\_tran_d$ predicts key listing attributes in the full listing-week panel. Coefficients are small, indicating no meaningful exposure-linked compositional shifts in observable characteristics.

Placebos and additional outcomes. Table A6 reports two sets of auxiliary tests. First, we examine placebo feedback labels that are not targeted by the buyer- or seller-facing workflows (e.g., appearance/temperament or seller-side communication labels); null effects support a workflow-specific interpretation. Second, we show that the effect on *Log_volume*, which captures neighborhood-week transaction volume (a proxy for overall market activity), is not statistically significant.

Robustness to cross-district spillovers. District-level exposure is constructed from district app shares, and agents typically operate within local service boundaries. Neighborhoods that physically straddle district borders are the most likely to generate cross-district handoffs and ambiguous exposure assignment. Table A7 re-estimates the main specifications after excluding the straddling neighborhoods and finds similar overall and decomposed coefficients.

Reducing overlap between buyer- and seller-side exposure. Buyer- and seller-side app shares are correlated, which can raise concerns that decomposed coefficients reflect shared variation rather than distinct workflow-specific exposure. Table A8 trims the sample to limit variation on the "other" side (excluding the top quartile of buyer exposure in Panel A and seller exposure in Panel B). This trimming reduces overlap and provides a robustness check that cross-side patterns are not mechanically driven by joint exposure.

Functional form and treatment intensity. Our baseline specification uses a continuous exposure measure. To reduce reliance on linear functional-form assumptions and to align with the "binned treatment" approach recommended in Callaway et al. (2024), Table A9 replaces continuous exposure with high/low exposure indicators (constructed from exposure quantiles). The resulting patterns are qualitatively consistent with the baseline estimates, supporting robustness to alternative treatment-intensity definitions.

Inference robustness. Baseline standard errors are clustered at the district level to reflect correlation within local markets where exposure is defined. Table A10 reports an alternative, more conservative clustering at the city level. Key coefficients remain similar in magnitude, indicating that the main conclusions are not sensitive to the clustering choice.

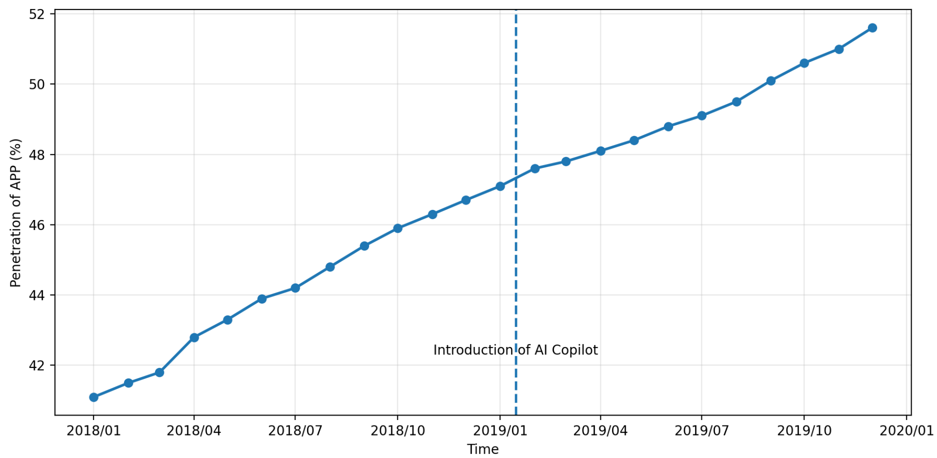


Figure A2: Exposure Measure over Time

Table A3: DiD Estimates: Treatment Interacted with Controls

	(1) Sold< 3 month	**(2)** Log_days _seller	**(3)** Log_days _buyer	**(4)** Log_price _unit	**(5)** Rating (buyer)	**(6)** Rating (seller)
Panel A: Overall Impact of AI Copilot						
Post×Ratio_tran	0.3113***	-0.6810***	-0.3340***	0.1201***	0.1906***	0.3243***
	(0.085)	(0.129)	(0.060)	(0.023)	(0.048)	(0.065)
Panel B: Decomposed Effects and Cross-Side Effects						
Post×Ratio_buyer	0.1209	-0.3880***	-0.1969***	0.1261***	0.1215***	0.2939***
	(0.102)	(0.076)	(0.061)	(0.024)	(0.039)	(0.066)
Post×Ratio_seller	0.3337***	-0.5825***	-0.2745***	0.0104	0.1370***	0.1339**
	(0.093)	(0.189)	(0.087)	(0.028)	(0.049)	(0.060)
Observations	267,882	231,680	231,680	231,680	75,596	20,336
Adjusted R^2	0.229	0.081	0.128	0.948	0.049	0.081
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table presents the estimation results of Equation 1, with additional controls (i.e., the interaction term between *Post* and four district-level variables). Panel A shows the overall effects. Panel B shows the impact decomposed by the AI's buyer- and seller-facing exposure. Post × Ratio_buyer, Post × Ratio_seller are pre-period district shares of buyers/sellers transacting via the app. Controls include apartment characteristics (e.g., building age, number of rooms, etc.) as well as unemployment, population, income, Gdp, and their interaction terms with *Post* in the respective district. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A4: DiD Estimates: Using Post-Treatment Copilot Usage Ratio

	(1) Sold< 3 month	**(2)** Log_days _seller	**(3)** Log_days _buyer	**(4)** Log_price _unit	**(5)** Rating (buyer)	**(6)** Rating (seller)
Panel A: Overall Impact of AI Copilot						
Post	-0.0363	0.4958***	0.2060***	-0.0401***	-0.0658*	-0.1957***
	(0.038)	(0.062)	(0.025)	(0.012)	(0.035)	(0.052)
Usage_tran	-0.0846	0.1485	0.4692***	-0.0673	-0.1391**	-0.3230
	(0.076)	(0.169)	(0.125)	(0.110)	(0.055)	(0.221)
Post×Usage_tran	0.4863***	-0.5018***	-0.2810***	0.1190***	0.2587***	0.4066***
	(0.100)	(0.162)	(0.054)	(0.033)	(0.061)	(0.074)
Panel B: Decomposed Effects and Cross-Side Effects						
Post	-0.0247	0.5206***	0.2194***	-0.0412***	-0.0573*	-0.1921***
	(0.041)	(0.066)	(0.024)	(0.012)	(0.034)	(0.046)
Usage_buyer	0.1329	0.0933	0.3603***	-0.0192	-0.0499	0.0757
	(0.082)	(0.164)	(0.135)	(0.082)	(0.062)	(0.300)
Post×Usage_buyer	0.1901*	-0.2403*	-0.1591**	0.1255***	0.1567***	0.3444***
	(0.114)	(0.129)	(0.064)	(0.032)	(0.046)	(0.063)
Usage_seller	-0.3395***	0.1627	-0.0112	-0.1872	-0.0529	-0.2191
	(0.119)	(0.318)	(0.256)	(0.128)	(0.059)	(0.328)
Post×Usage_seller	0.4845***	-0.6033**	-0.2960***	0.0102	0.1751***	0.1561***
	(0.118)	(0.282)	(0.106)	(0.039)	(0.056)	(0.058)
Observations	267,882	231,680	231,680	231,680	75,595	20,336
Adjusted R^2	0.222	0.079	0.128	0.948	0.047	0.079
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table presents the estimation results of Equation 1, but replaces pre-exposure proxies with district-level copilot *usage intensity* measured in the month immediately following rollout (Usage_tran for overall usage; Usage_buyer and Usage_seller for buyer- and seller-side usage). Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A5: Composition of Listed Properties

	(1) Building size	**(2)** Building age	**(3)** Floor	**(4)** Decoration	**(5)** Room
Post×Ratio_tran	-0.0122	0.0189	0.0013	-0.0093*	0.0050
	(0.008)	(0.015)	(0.005)	(0.005)	(0.003)
Observations	13,581,455	13,581,455	13,581,455	13,581,455	13,581,455
Adjusted R^2	0.646	0.429	0.012	0.088	0.635
Controls	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimation results of Equation 1 results for the composition of listed properties' characteristics. Standard errors clustered at the district level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A6: Placebo Tests and Transaction Volume

	(1) Intro_well (Buyer)	**(2)** Appearance (Buyer)	**(3)** Talk_ineff (Seller)	**(4)** Talk_comfort (Seller)	**(5)** Trust (Seller)	**(6)** Log_volume
Panel A: Overall Impact of AI Copilot						
Post×Ratio_tran	0.0703	0.0408	-0.0080	-0.0026	-0.0176	0.1094
	(0.065)	(0.081)	(0.008)	(0.006)	(0.019)	(0.111)
Panel B: Decomposed Effects and Cross-Side Effects						
Post×Ratio_buyer	0.0348	0.0220	-0.0051	0.0018	0.0091	-0.0010
	(0.065)	(0.063)	(0.007)	(0.006)	(0.023)	(0.115)
Post×Ratio_seller	0.0515	0.0455	-0.0063	-0.0050	-0.0248	0.1660
	(0.052)	(0.052)	(0.006)	(0.006)	(0.022)	(0.118)
Observations	75,596	75,596	20,336	20,336	20,336	58,968
Adjusted R^2	0.168	0.207	0.017	0.049	0.120	0.612
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table reports the estimation results of Equation 1 for placebo tests and transaction volume. Columns (1)-(2) show placebo tests on buyer feedback labels not targeted by the AI. Columns (3)-(5) show placebo tests on seller feedback for communication skills. Column (6) shows the effect on log transaction volume. Standard errors clustered at the district level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A7: DiD Estimates: Dropping Cross-District Neighborhoods

	(1) Sold< 3 month	(2) Log_days _seller	(3) Log_days _buyer	(4) Log_price _unit	(5) Rating (buyer)	(6) Rating (seller)
Panel A: Overall Impact of AI Copilot						
Post×Ratio_tran	0.4315***	-0.4740***	-0.2496***	0.1149***	0.2434***	0.3966***
	(0.083)	(0.159)	(0.051)	(0.026)	(0.052)	(0.073)
Panel B: Decomposed Effects and Cross-Side Effects						
Post×Ratio_buyer	0.1712*	-0.2653**	-0.1588***	0.1249***	0.1535***	0.3653***
	(0.102)	(0.110)	(0.052)	(0.025)	(0.040)	(0.052)
Post×Ratio_seller	0.4212***	-0.4964**	-0.2219***	0.0088	0.1741***	0.1358**
	(0.105)	(0.232)	(0.083)	(0.030)	(0.045)	(0.060)
Observations	225,206	194,133	194,133	194,133	61,012	16,385
Adjusted R^2	0.224	0.079	0.126	0.946	0.050	0.072
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table presents the estimation results of Equation 1 after excluding neighborhoods that span across two districts. Panel A shows the overall effects. Panel B shows the impact decomposed by the AI's buyer- and seller-facing exposure. Post × Ratio_buyer, Post × Ratio_seller are pre-period district shares of buyers/sellers transacting via the app. Controls include apartment characteristics (e.g., building age, number of rooms, etc.) as well as unemployment, population, income, and Gdp in the respective region. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A8: DiD Estimates: Limiting Exposure on the Other Side

	(1) Sold< 3 month	(2) Log_days _seller	(3) Log_days _buyer	(4) Log_price _unit	(5) Rating (buyer)	(6) Rating (seller)
Panel A: Exclude top 25% of buyer exposure						
Post×Ratio_buyer	0.2375	0.0984	-0.1006	-0.0115	0.2089*	0.3766*
	(0.169)	(0.166)	(0.113)	(0.034)	(0.117)	(0.206)
Post×Ratio_seller	0.5645***	-0.3370***	-0.3594***	-0.0090	0.2052***	0.1470
	(0.145)	(0.124)	(0.122)	(0.021)	(0.067)	(0.097)
Observations	197,971	171,981	171,981	171,981	47,076	11,785
Adjusted R^2	0.216	0.082	0.123	0.956	0.041	0.063
Panel B: Exclude top 25% of seller exposure						
Post×Ratio_buyer	0.3781***	-0.2667*	-0.1683*	0.1030***	0.1242**	0.2494***
	(0.088)	(0.140)	(0.092)	(0.032)	(0.056)	(0.086)
Post×Ratio_seller	0.2104	0.0360	-0.2791	-0.1250**	0.2635	0.5185
	(0.258)	(0.282)	(0.174)	(0.054)	(0.190)	(0.343)
Observations	202,520	178,153	178,153	178,153	52,843	13,944
Adjusted R^2	0.208	0.085	0.125	0.948	0.043	0.073
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This trimming reduces overlap in buyer- and seller-side exposure and tests whether cross-side coefficients persist when variation on the other side is limited Post × Ratio_buyer, Post × Ratio_seller are pre-period district shares of buyers/sellers transacting via the app. Controls include apartment characteristics (e.g., building age, number of rooms, etc.) as well as unemployment, population, income, and Gdp in the respective region. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A9: DiD Functional Form: Using Binary Ratios

	(1) Sold< 3 month	(2) Log_days _seller	(3) Log_days _buyer	(4) Log_price _unit	(5) Rating (buyer)	(6) Rating (seller)
Panel A: Overall Impact of AI Copilot						
Post×Ratio_tran_high	0.1030*	-0.2845***	-0.1052***	0.0778***	0.1078***	0.0963***
	(0.059)	(0.087)	(0.037)	(0.015)	(0.015)	(0.027)
Post×Ratio_tran_low	-0.1365***	-0.0127	0.0712**	0.0110	-0.0167	-0.1390***
	(0.041)	(0.052)	(0.031)	(0.010)	(0.044)	(0.050)
Panel B: Decomposed Effects and Cross-Side Effects						
Post×Ratio_buyer_high	-0.0196	-0.1378	-0.0380	0.0569***	0.0338	0.1184**
	(0.052)	(0.088)	(0.058)	(0.017)	(0.023)	(0.049)
Post×Ratio_buyer_low	-0.0302	-0.0364	0.0374	0.0016	-0.0086	-0.0824
	(0.042)	(0.052)	(0.047)	(0.013)	(0.031)	(0.056)
Post×Ratio_seller_high	0.1251**	-0.1566	-0.0429	0.0215	0.0775***	0.0281
	(0.058)	(0.110)	(0.060)	(0.020)	(0.021)	(0.039)
Post×Ratio_seller_low	-0.1205**	0.0027	0.0163	0.0076	-0.0457	-0.0515
	(0.050)	(0.058)	(0.048)	(0.011)	(0.032)	(0.071)
Observations	267,882	231,680	231,680	231,680	75,596	20,336
Adjusted R^2	0.219	0.079	0.128	0.948	0.047	0.079
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table presents the estimation results of Equation 1, with binary ratios based on quintile. Panel A shows the overall effects. Panel B shows the impact decomposed by the AI's buyer- and seller-facing exposure. Post × Ratio_buyer, Post × Ratio_seller are pre-period district shares of buyers/sellers transacting via the app. Controls include apartment characteristics (e.g., building age, number of rooms, etc.) as well as unemployment, population, income, and Gdp in the respective district. Standard errors clustered at the district level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A10: DiD Alternative Clustering at the City Level

	(1) Sold< 3 month	(2) Log_days _seller	(3) Log_days _buyer	(4) Log_price _unit	(5) Rating (buyer)	(6) Rating (seller)
Panel A: Overall Impact of AI Copilot						
Post×Ratio_tran	0.4215***	-0.4434*	-0.2484***	0.1005*	0.2312**	0.2312**
	(0.130)	(0.225)	(0.064)	(0.049)	(0.081)	(0.081)
Panel B: Decomposed Effects and Cross-Side Effects						
Post×Ratio_buyer	0.1744	-0.2376	-0.1411**	0.1152**	0.1511**	0.1511**
	(0.130)	(0.152)	(0.065)	(0.046)	(0.054)	(0.054)
Post×Ratio_seller	0.4117**	-0.4877**	-0.2550***	-0.0005	0.1399**	0.1399**
	(0.172)	(0.216)	(0.086)	(0.027)	(0.062)	(0.062)
Observations	267,882	231,680	231,680	231,680	75,596	20,336
Adjusted R^2	0.222	0.079	0.128	0.948	0.047	0.079
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table presents the estimation results of Equation 1, with standard errors clustered at city level. Panel A shows the overall effects. Panel B shows the impact decomposed by the AI's buyer- and seller-facing exposure. Post × Ratio_buyer, Post × Ratio_seller are pre-period district shares of buyers/sellers transacting via the app. Controls include apartment characteristics (e.g., building age, number of rooms, etc.) as well as unemployment, population, income, and Gdp in the respective district. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix D. Additional Heterogeneity Results

This appendix reports additional heterogeneity analyses that test whether copilot effects are larger when intermediation frictions are plausibly more binding. We examine heterogeneity by (i) agent pre-treatment performance (a standardized score) and (ii) local market thickness, proxied by the number of listings in the neighborhood in the pre-period. For brevity, the tables report the highest-order interaction terms; all specifications include the same fixed effects and controls as in the baseline regressions.

Table A11 shows that the estimated effects attenuate with higher pre-treatment performance, consistent with the copilot providing larger incremental value for less-experienced agents.

Table A11: Heterogeneity in the Effect of AI Adoption by Agent Performance

	(1) Sold< 3 month	**(2)** Log_days _seller	**(3)** Log_days _buyer	**(4)** Log_price _unit	**(5)** Rating (buyer)	**(6)** Rating (seller)
Panel A: Moderation of Overall Effect						
Post×Ratio_tran	0.435***	-0.478***	-0.297***	0.103***	0.3640***	0.6465***
	(0.081)	(0.150)	(0.054)	(0.029)	(0.057)	(0.121)
Post×Ratio_tran×Performance	-0.100***	0.110***	0.0784***	-0.0018	-0.0401***	-0.1010***
	(0.009)	(0.023)	(0.027)	(0.002)	(0.015)	(0.031)
Panel B: Moderation of Decomposed Effects						
Post×Ratio_buyer	0.183*	-0.284***	-0.180***	0.116***	0.2248***	0.5409***
	(0.102)	(0.109)	(0.060)	(0.028)	(0.049)	(0.128)
Post×Ratio_buyer×Performance	-0.0849***	0.0973***	0.0627**	-0.0012	-0.0248**	-0.0830**
	(0.010)	(0.028)	(0.012)	(0.003)	(0.012)	(0.038)
Post×Ratio_seller	0.463***	-0.551**	-0.320***	-0.003	0.2388***	0.1910*
	(0.106)	(0.233)	(0.094)	(0.034)	(0.049)	(0.108)
Post×Ratio_seller×Performance	-0.0427***	0.0725**	0.0473	-0.0017	-0.0243***	-0.0219
	(0.014)	(0.031)	(0.030)	(0.004)	(0.009)	(0.032)
Observations	267,882	231,680	231,680	231,680	75,595	20,336

Notes. This table presents heterogeneity analysis by agent's pre-treatment performance score (standardized). For brevity, only the highest-order interaction terms and their constituent two-way interactions are shown. All specifications include the full set of main effects and controls as in Table 2. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A12 shows that effects are weaker in thicker neighborhoods (more listings), consistent with larger gains where matching and coordination frictions are more binding.

Table A12: Heterogeneity in the Effect of AI Adoption by Market Thickness

	(1) Sold< 3 month	(2) Log_days _seller	(3) Log_days _buyer	(4) Log_price _unit	(5) Rating (buyer)	(6) Rating (seller)
Panel A: Moderation of Overall Effect						
Post×Ratio_tran×Num_list	-0.0002	0.0007**	0.0003	-0.0001***	-0.0002**	-0.0007*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel B: Moderation of Decomposed Effects						
Post×Ratio_buyer×Num_list	0.0000	0.0009***	0.0002	-0.0002***	-0.0002**	-0.0006**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Post×Ratio_seller×Num_list	-0.0002	-0.0000	0.0001	0.0001	-0.0001	-0.0001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	267,882	231,680	231,680	231,680	75,595	20,336

Notes. This table presents heterogeneity analysis by market thickness (the number of listings at the neighborhood level in the pretreatment period). For brevity, only the triple interaction terms are shown. All specifications include the full set of controls and lower-order interaction terms. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Online Appendix E. Additional Mechanism Evidence

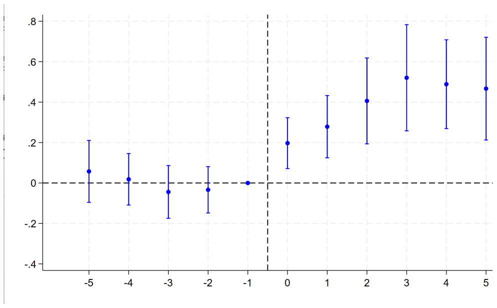
This appendix provides additional event-study evidence on workflow margins and process measures. The figures are based on event-study regressions (Equation 3) using the same clustering as in the main analysis. These plots show that the relevant process measures move sharply after the rollout and show flat pre-trends.

Buyer-side workflow margins. Figure A3 reports event studies for buyer feedback labels and buyer-side process measures (viewings and last listed prices). The key diagnostic is the absence of differential pre-trends and the alignment of post-rollout movements with the buyer-side workflow: improved perceived recommendation/communication/trust and reduced search intensity as proxied by on-site viewings.

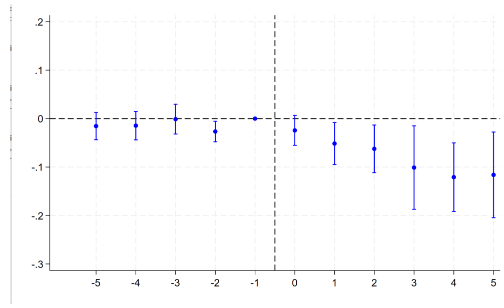
Seller-side workflow margins. Figure A4 reports event studies for seller feedback labels and seller pricing behavior. Post-rollout, price adjustment frequency increases while adjustment magnitudes fall, consistent with faster convergence in listing strategy.

Directional price adjustments. Table A13 decomposes total adjustments into downward and upward components. This decomposition clarifies whether the increase in adjustment frequency is driven by price cuts (consistent with correcting initial overpricing and speeding convergence) vs. upward revisions. The results show that the adjustment response is concentrated in downward adjustments, consistent with a faster price-discovery process.

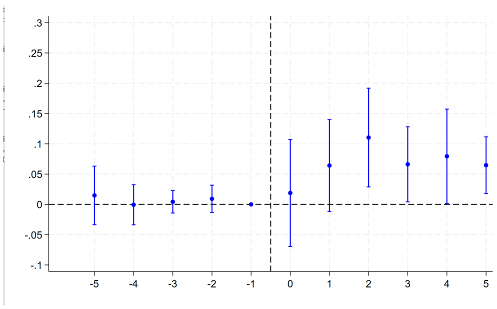
Figure A3: Buyer-Side Event Studies



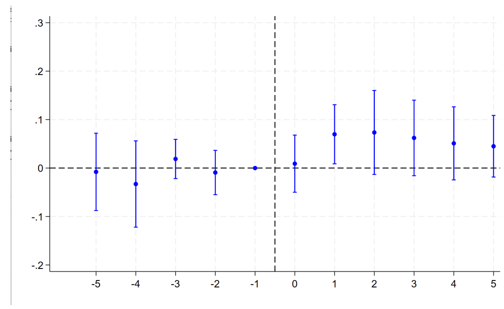
(a) recommend_optimal (using Ratio_tran)



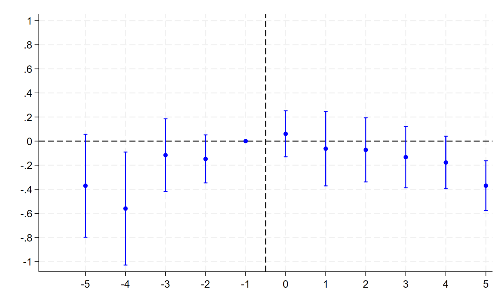
(b) Talk_inefficiency (using Ratio_tran)



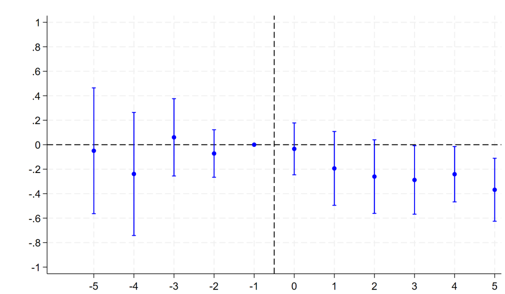
(c) Talk_comfort (using Ratio_tran)



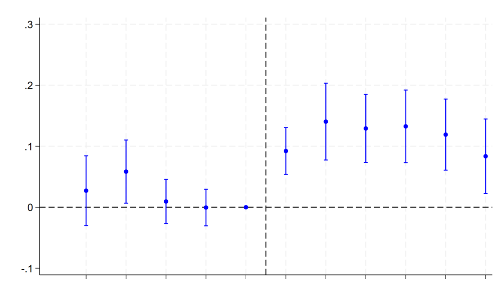
(d) Trust (using Ratio_tran)



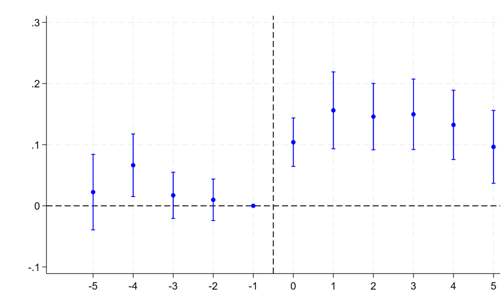
(e) Viewing (using Ratio_tran)



(f) Viewing (using Ratio_buyer)



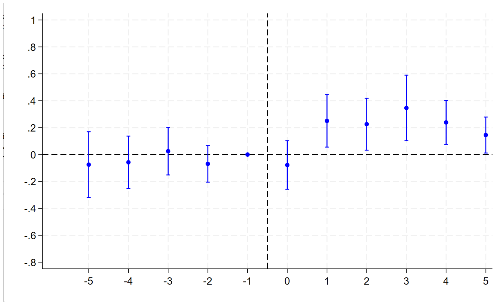
(g) Log_listed_price (using Ratio_tran)



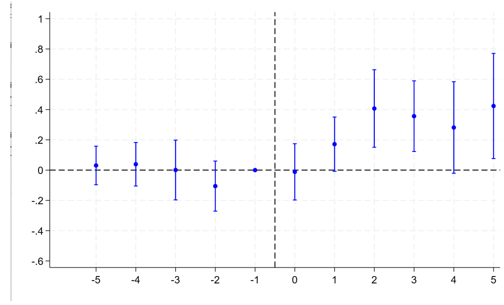
(h) Log_listed_price (using Ratio_buyer)

Notes. This figure plots the bimonthly coefficients and 95% confidence intervals from event study regressions (Equation 2) regarding buyer-side outcomes. Panels (a)-(d) show effects on buyer feedback labels. Panels (e)-(h) show effects on weekly on-site viewing frequency and the last listed prices of sold properties, presenting results for both the overall (Ratio_tran) and decomposed (Ratio_buyer) specifications. *Ratio_buyer* is the pre-treatment district shares of buyers transacting via the app. Standard errors are clustered at the district level.

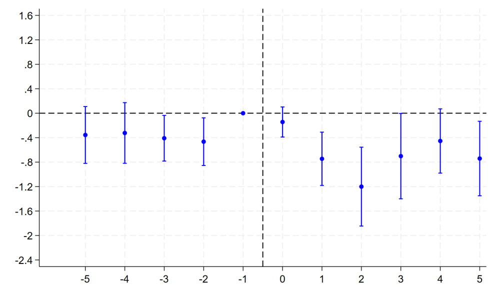
Figure A4: Seller-Side Event Studies: Strategy and Price Adjustments



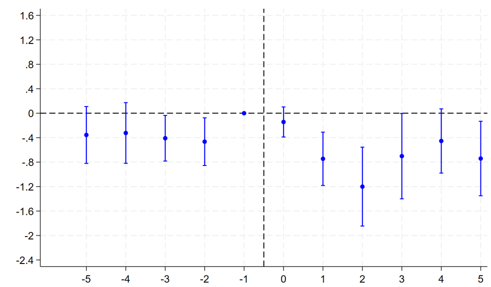
(a) Pricing_optimal



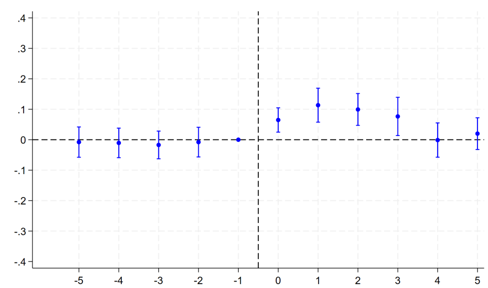
(b) Sellingpoint_accuracy



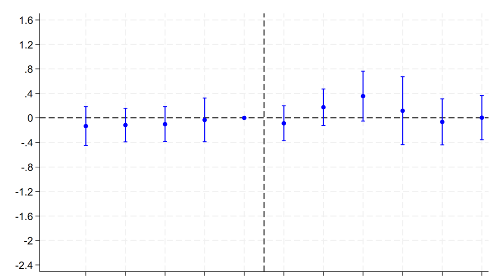
(c) Total_adjust_times



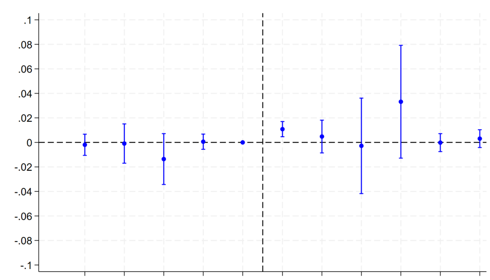
(d) Total_adjust_mag



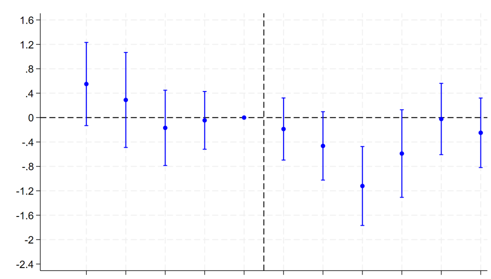
(e) Down_adjust_times



(f) Down_adjust_mag



(g) Up_adjust_times



(h) Up_adjust_mag

Notes. This figure plots the bimonthly coefficients and 95% confidence intervals from event study regressions (Equation 2) regarding seller-side behaviors. All specifications use *Ratio_tran*. Panels (a)-(b) show effects on seller feedback labels. Panels (c)-(d) show total weekly frequency and magnitude of price adjustments. Panels (e)-(h) decompose these adjustments into downward and upward directions. Standard errors are clustered at the district level.

Table A13: Analysis of Upward and Downward Price Adjustments

	(1) Down_adjust _times	**(2)** Down_adjust _mag	**(3)** Up_adjust _times	**(4)** Up_adjust _mag
	Panel A: Overall Impact of AI Copilot			
Post×Ratio_tran	0.0533***	-0.3504**	0.0106	-0.6597***
	(0.014)	(0.152)	(0.011)	(0.232)
	Panel B: Decomposed Effects and Cross-Side Effects			
Post×Ratio_buyer	0.0078	0.2421	0.0126	0.0849
	(0.025)	(0.196)	(0.011)	(0.245)
Post×Ratio_seller	0.0845***	-0.6345**	-0.0026	-0.9195***
	(0.031)	(0.256)	(0.011)	(0.245)
Observations	13,581,455	1,807,297	13,581,455	491,030
Adjusted R^2	0.026	0.059	0.015	0.047
Controls	Yes	Yes	Yes	Yes
City×Month FE	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes

Notes. This table presents the estimation results of Equation 1 results for sellers' price adjustment behaviors. The outcome variables are the number of downward price adjustments, the magnitude of downward price adjustments, the number of upward price adjustments, and the magnitude of upward price adjustments, respectively. Panel A shows the overall effects. Panel B shows the impact decomposed by the AI's buyer- and seller-facing exposure. Post × Ratio_buyer, Post × Ratio_seller are pre-period district shares of buyers/sellers transacting via the app. Controls include apartment characteristics (e.g., building age, number of rooms, etc.) as well as unemployment, population, income, and Gdp in the respective district. Standard errors clustered at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1